

Behavior Exchange Theory and Marital Decision Making

John Gottman, Cliff Notarius, Howard Markman,
Steve Bank, and Bruce Yoppi
Indiana University

Mary Ellen Rubin
South Central Indiana Mental Health Foundation, Bloomington, Indiana

Distressed and nondistressed couples in two studies made decisions on high- and low-conflict tasks. They continuously coded both the intended impact of their own behavior and the impact of their spouse's behavior. In Study 1 distressed couples did not differ from nondistressed couples on how they intended their behavior to be received. However, the behavior of distressed spouses was actually received more negatively by their partners than the behavior of their nondistressed counterparts. The couples in Study 2 also behaved in a way consistent with a communication deficit explanation of distressed marriages; that is, distressed couples' behavior was likely to be coded as more negative than they intended. Task effects and a reciprocity hypothesis were also tested. Data from Study 1 showed no conflict effect, but the results of Study 2 suggested that high-conflict tasks may be a better means for discriminating distressed from nondistressed couples than low-conflict tasks. The data on reciprocity indicate only minimal support for the view that distressed marriage is characterized by less positive or more negative reciprocity than nondistressed marriage.

Recently, two different hypotheses have been proposed to describe conflict resolution in distressed and nondistressed marriages. One hypothesis, based on behavior exchange theory, is that it is more likely that nondistressed couples will produce behaviors coded as positive by observers than will distressed couples. This hypothesis has recently received some support (Birchler, Weiss, & Vincent, 1975).

A second hypothesis is that there is a greater *reciprocity* of positive exchange in nondistressed than in distressed marriages. Reciprocity of positive exchange has been suggested as the central characteristic of successful marital interaction in the clinical literature on marriage counseling (Azrin, Naster, &

Jones, 1973; Lederer & Jackson, 1968; Rapaport & Harrell, 1972; Stuart, 1969; Weiss, Hops, & Patterson, 1973).

The two hypotheses are independent. High base rates of positive behaviors do not imply reciprocity. For example, Birchler et al.'s (1975) finding of higher rates of positive behavior in nondistressed compared with distressed couples does not imply that reciprocity is different for the two groups of couples. Although nondistressed couples may seem to be reciprocating positive behavior more frequently than distressed couples, that may only be an artifact of the higher probability of positive behaviors in nondistressed couples. By emitting more positive responses, nondistressed couples increase the probability that one partner's positive response will be followed by the other partner's positive response. What needs to be demonstrated is that a particular consequent code in a sequence can be predicted from a particular antecedent code. The test for reciprocity is that the diagonal elements of the first-order Markov matrix of conditional probabilities must be significantly greater than the relative frequency of the consequent code (Patterson, 1974). For example, the conditional probability (p) of a conse-

This research was partially supported by Grant R01 MH 24594 SP from the National Institute of Mental Health. The authors wish to thank the Mental Health Center, the Center for Human Growth, and the Center for Counseling and Psychological Services for their cooperation, and Steve Asher, Jack Bates, Fran Cherry, and Alice Eagly for their comments on the manuscript.

Requests for reprints should be sent to John Gottman, who is now at the Department of Psychology, University of Illinois, Children's Research Center, 51 E. Gerty Drive, Champaign, Illinois 61820.

quent positive wife code ($W+$) given an antecedent positive husband code ($H+$) must be significantly greater than the unconditional probability of occurrence of $W+$. This means that a knowledge of the antecedent code, $H+$, adds significantly to the ability to predict the occurrence of a $W+$ code, over and above prediction, based simply on knowing the relative frequency of occurrence of $W+$. Symbolically, what must be demonstrated is that $p(W+/H+) > p(W+)$ (Raush, Barry, Hertel, & Swain, 1974). In the present investigation these information analyses were performed to assess the ability of reciprocity variables to discriminate distressed from nondistressed marriages.

The distinction between high rates of positive codes and reciprocity has been ignored in the clinical literature on family interaction. For example, Alexander (1973) found that the correlation across families between parent-to-child supportive behavior and child-to-parent supportive behavior was significantly different from zero (father-son $r = .69$, $p < .05$; mother-son $r = .59$, $p < .05$). Contrary to Alexander's conclusions, these correlations do not imply reciprocity. A family with high rates of supportiveness could be distributing these behaviors *noncontingently* throughout a discussion. In this case the correlations obtained would be high due to different base rates across families, but there would still be no evidence of reciprocity.

The base rate-reciprocity issue is also important for distinguishing between behavior exchange and social learning theories of marital conflict resolution. Birchler et al. (1975) used a mix of language from both theories. They demonstrated that distressed couples emit lower rates of positive codes and higher rates of negative codes than do nondistressed couples. This does not justify their view of the positive codes as "social reinforcements." It would be necessary to show that a particular code on the part of one spouse affected the probability of occurrence of a particular code of the other spouse.

Reciprocity hypotheses, therefore, have yet to be adequately tested. Behavior exchange hypotheses of marital conflict resolution have also been ambiguous in making specific predictions. There has, in fact, been little re-

search on marital interaction from a behavior exchange standpoint. For example, standard game theory experiments, which are commonly associated with behavior exchange theory, were not applied to the study of the marital dyad until Speer (1972). Speer studied 60 distressed couples seeking counseling in a clinic and 60 nondistressed couples in an extension of the Prisoner's Dilemma Game. None of the four response variables in Speer's study discriminated the two groups on four forms of the game.

Perhaps two reasons for the lack of success of Speer's measures are that (a) behavior exchange experiments reduce the behavioral repertoire of the couple and (b) they externally control the payoffs of specific behaviors that they exhibit (Gergen, 1969). Another potential problem with game theory experiments applied to the marital dyad has to do with how the payoff matrix is interpreted. There are many alternative ways of interpreting a payoff matrix. Behavior exchange theory proposes that a dyad engaged in a mutually satisfying relationship will exchange behaviors that have low cost and high reward to both members (Thibaut & Kelley, 1959). One interpretation of the reward-cost notion is that *producing* a specific behavior has a specific reward and a specific cost to the person producing the behavior (Gergen, 1969, p. 36). An alternative interpretation of the payoff matrix is that in a satisfying dyadic relationship, a person will receive behaviors that have high reward for himself or herself. This implies that a person will perceive a relationship as satisfying to the extent that he or she codes the behaviors *received* as a positive.

Another issue regarding the way payoffs are assigned is the comparison level for alternative relationships. Presumably, a behavior that will be coded as positive in one relationship could be coded as negative in another relationship in which the receiver's comparison level is much higher. The fact that we do not know a spouse's comparison level for alternatives suggests a phenomenological method for measuring payoff. To test this interpretation of the payoff matrix, then, it is necessary to design a measurement procedure in which the positivity of the behavior received would be coded directly by the receiver. In this pro-

cedure, it is the couple who codes the behaviors exchanged, not independent observers. It is possible that what is coded as a warm smile by two observers is perceived as a sarcastic smirk to a spouse, and conversely, a "put-down" or "interruption" that is coded as negative in Birchler et al. (1975, p. 352) is not always seen as negative by the couple.

The concept of the couple coding their own behavior raises the question of whether distressed couples differ from nondistressed couples in the *intended* positivity of their messages. According to the widely accepted communication deficit explanation of marital distress, distressed couples are presumed to intend their messages to be received as far more positive than they are in fact received. However, it could be that distressed couples *intend* their messages to be more negative, or less positive, than nondistressed couples. If the messages were then received as more negative, or less positive, distressed couples would be communicating well; their spouses would receive the messages as intended. A communication deficit explanation of marital distress would seem more appropriate if intended positivity of messages differed markedly from the coding of the messages received. In this case messages received would not be coded as the sender intended them to be coded.

In the present investigation, couples interacted using a "talk table," a device we constructed in which only one person could speak at a time. After speaking, the speaker coded the "intended impact" of his or her message. Before speaking, the listener coded the "actual impact" of the message. The talk table need not constrain the behavioral repertoire of the interacting dyad as in Speer (1972); the talk table is apparently an inexpensive, simple way to study behavior exchange in marriages, especially compared with detailed ratings of videotapes (as in Birchler et al., 1975). However, before this can be confidently asserted, the talk table must demonstrate its ability to elicit different patterns of exchange in different kinds of couples.

One purpose of the present investigation was to assess the effect of the experimental task on the coding of the interaction. The effect of situational context on marital and family interaction has been either ignored

or minimized (Riskin & Faunce, 1972). For example, Haley (1964, 1967) placed no importance on the tasks he chose for families to work on. Jacob and Davis (1973) reported considerable stability across experimental tasks to the structure of talk and interruptions in father-mother-child interactions. One potentially important dimension of contextual variation that has been ignored is the degree to which the decision-making task induces conflict. For example, all of Jacob and Davis' (1973) tasks were low-conflict tasks. It would seem theoretically important to ascertain whether the variables derived from behavior exchange theory can discriminate nondistressed couples from distressed couples in low-conflict as well as high-conflict tasks. The present investigation employed two sets of tasks that induced either high or low conflict. The tasks ranged in content from a consensual ranking of a list of preferred dog breeds to the discussion of an actual unresolved marital problem.

STUDY 1

Method

Subjects and selection. Thirty couples responded to an advertisement asking for couples who either classified their relationship as "mutually satisfying" or "experiencing marital difficulties." All couples were offered \$10 for participating, and it was made clear that no therapy would be offered. In addition, other distressed couples were recruited from local campus and community mental health centers. Of the 15 distressed couples who participated in the study, 11 were referred from clinical sources and 4 responded to the advertisement.

Two major definitions of marital distress have been used, self-report of satisfaction and whether or not the couple comes to the attention of public agencies such as marriage counselors or divorce courts. Diverse measures of marital satisfaction have been shown to tap the same basic factor (Burgess, Locke, & Thomes, 1971). The Locke-Wallace Marital Relationship Inventory (MRI) is the most widely used of these inventories; it has excellent discriminative validity in cross-sectional studies (Navran, 1967) and good predictive validity in longitudinal studies (Terman & Wallin, 1949). Using the cutoff scores on the MRI recommended by Burgess et al. (1971, pp. 330-331), only high-scoring nondistressed couples and low-scoring distressed couples were included in the analyses. Specifically, distressed couples in which at least one spouse's MRI score was less than 85 were included, and nondistressed couples in which both husband and wife MRI scores were greater than 102 were included. A distressed couple was selected for analysis only if at least one spouse was dissatisfied

with the marriage (low MRI), and a nondistressed couple was selected for analysis only if both spouses were satisfied with the marriage (high MRI). Using these cutoff scores, in Study 1 10 distressed couples and 6 nondistressed couples were selected for analysis from the sample of 30 couples, all of whom had completed the experimental procedures.¹ Coefficient alpha for the sample of Study 1 on the MRI was .88.

The two groups of couples did not significantly differ in age, educational level, or the number of years married. Couples were, on the average, 24.95 years old and had been married an average of 3.22 years.

Procedure. Interview teams were composed of one male graduate student in clinical psychology and one female undergraduate. There were four such teams, and each couple was interviewed by one of these teams. Spouses individually filled out a problem inventory, a demographic information sheet, and the MRI. The problem inventory asked each spouse to rate the perceived severity of a list of problem areas. The couple was then asked to agree jointly on the three most salient current problem areas or areas of disagreement in their marriages. Next, the couple was asked to describe each of the problem areas separately and to provide a specific description like a play-by-play account of an incident, providing situational context and a typical conversation illustrating each problem. This was done in order to establish a rapport with the interviewing team and to provide specificity to the problem to be discussed on the talk table. The talk table is a double sloping table. A toggle switch on the side of the table operated by the couple lit a button on the side of the spouse who had the floor to speak. There were two rows of five buttons. The five buttons on the left were used by the speaker to code the "intended impact" of his or her message; the five buttons on the right were used by the receiver to code the "impact" of the message received. The buttons were labeled "super negative," "negative," "neutral," "positive," and "super positive." Although partners could see one another, a metal shield blocked the buttons from the view of the other person so that neither spouse could see the codes assigned by their partners throughout the experiment.

In order to familiarize themselves with its use, each couple was then asked to converse for a few minutes using the talk table. When the couple indicated they were comfortable with the talk table, the experimental tasks began.

Each couple then completed three low-conflict tasks: (a) The choice questionnaire (Haley, 1964) is a consensus decision-making task that requires an agreed-upon ranking of personal preferences using lists such as new cars or breeds of dogs. (b) Three Thematic Apperception Test cards (Locke, 1961) required the couple to create jointly one story for each set of cards. (c) One of two tasks was administered first individually and then again for consensual ranking; one task (called NASA) involved rank ordering 15 items for their survival value for a life-and-death trip to the moon. Correct answers

had been supplied by the National Aeronautics and Space Administration (Hall, 1971). The other task designed for this study (called the food task) involved rank ordering 10 items according to their nutritional value. Correct answers had been supplied by a nutrition specialist at Indiana University. A pilot study with 147 undergraduates, conducted to determine whether there was a sex bias to the tasks, revealed no significant sex differences in knowledge on either of the tasks. The couple also completed two high-conflict tasks: (a) The Inventory of Marital Conflict (IMC) (Olson & Ryder, 1970) has been frequently used in marital research. The IMC is a high-conflict consensus decision-making task, in which the couple is presented with three short vignettes of marital conflict and required to agree on which spouse in the vignette is primarily responsible for the problem. (b) The couple was also asked to discuss a current problem that they had agreed in the interview was most salient and asked to try to come to a mutually satisfactory resolution of this problem. Within each session the five tasks were randomly ordered. Distressed couples were told that a report would be written that they could read and discuss with the interviewing team and that could be sent to the couple's therapist with their written permission.

Results and Discussion

Multivariate analysis of variance of spouses' coding of their partners' behavior (impact), combined over all five tasks, with distressed-nondistressed as a between-subjects factor and husband-wife as a within-subjects factor resulted in a significant effect for the distress factor, Wilks-Lambda $F(4, 25) = 4.54$, $p < .01$, and no significant sex or Sex \times Distress interaction. Table 1 presents the univariate F ratios that contributed to the distress multivariate F ratio. Nondistressed spouses were more likely to code their partners' behavior as positive and super positive and less likely to code their partners' behavior as negative than their distressed counterparts.

Multivariate analysis of variance of spouses' intent resulted in no significant distress, sex, or sex \times distress F ratios. There were no significant differences in intention for either husbands or wives on any category. For comparison with the impact data, Table 2 pre-

¹ Analyses on couples were also done regardless of their Marital Relationship Inventory score to study the effect of the convergent operations procedure for defining the two groups of couples. The same pattern of results emerge from these analyses; a similar pattern of F ratios were significant but were not as large. This suggests that multiple criteria improved discriminability for classifying couples.

TABLE 1

UNIVARIATE *F* RATIOS AND MEANS FOR RELATIVE FREQUENCY OF SPOUSE'S CODING OF PARTNER'S BEHAVIOR, STUDY 1

Variable ^a	<i>F</i> ^b	Distressed average	Nondistressed average
Super negative	2.06	.06	.01
Negative	10.66**	.24	.11
Positive	7.47*	.34	.50
Super positive	7.36*	.02	.17

^a Neutral is not included in the analysis of multivariate *F*; all five variables form a linearly dependent set, since they sum to 1.0.

^b degrees of freedom = 1, 28.

* $p < .05$.

** $p < .01$.

sents means and univariate *F* ratios for distressed and nondistressed couples.

These results are consistent with the data obtained by Birchler et al. (1975) for observers' coding of the couple's behavior. Birchler et al. found that distressed couples were more likely to emit negative and less likely to emit positive codes than nondistressed couples. Furthermore, the results support a communication deficit model of marital distress. By looking at Table 1 and Table 2 in combination, it can be seen that the distributions of intent and impact are similar for nondistressed couples. This was not the case for distressed couples. Both groups of couples were nearly identical in the way they intended their behavior to be coded. Therefore, for distressed couples, their behavior was coded as far less positive than was their intention.

Conflict effect. To check the low-high conflict manipulation, the videotapes of four couples, two distressed and two nondistressed, were randomly selected and coded for agreement and disagreement. The first 2 minutes of each task were coded, with coders tallying each occurrence of an agreement or disagreement. There were four pairs of coders; each pair coded only one couple. Coders were blind to the task hypothesis and as to whether the couple whose tape they were coding was distressed or nondistressed. The total number of agreements and disagreements were calculated separately for each coder, summing over high- and low-conflict tasks. Coders were first trained using an interaction category system developed in our laboratory (Gottman, Notarius, & Markman, Note 1) and reached

criterion on an achievement test, which tested their knowledge of agreement and disagreement codes. One coder was designated the coder and the other the reliability checker. Correlations between coder and reliability checker over the four couples were .93 for agreement and .93 for disagreement. In the low-conflict tasks, there were averages of 10.33 agreements and 6.92 disagreements; in the high-conflict tasks, there were averages of 5.38 agreements and 8.00 disagreements, $\chi^2(1) = 4.23$, $p < .05$. In research on family interaction, agreement and disagreement have been shown to be the best consistent discriminators between distressed and nondistressed families (Riskin & Faunce, 1972). Agreement-disagreement ratios less than one are characteristic of distressed families, and ratios greater than one are characteristic of nondistressed families. For the low-conflict tasks, the agreement-disagreement ratio was 1.49, whereas for the high-conflict tasks it was .67.

To assess the effect of high- or low-conflict tasks, the five received impact codes were considered on a 5-point Likert scale ranging from 1 (super negative) to 5 (super positive). The same Likert scale was constructed for the intent codes. Scale scores were separately averaged for the three low-conflict tasks and the two high-conflict tasks. An analysis of variance was performed on the impact data and on the intent data, with two levels of distress (distressed-nondistressed), two levels of sex (husband-wife), and two levels of conflict (high-low). As expected, there was a significant *F* ratio for the impact data on the distress factor, $F(1, 56) = 15.37$, $p < .001$. Distressed couples averaged 3.04 and nondistressed couples averaged 3.52 over all tasks.

TABLE 2

UNIVARIATE *F* RATIOS AND MEANS FOR RELATIVE FREQUENCY OF SPOUSES' INTENT, STUDY 1

Variable	<i>F</i> ^a	Distressed average	Nondistressed average
Super negative	.18	.00	.00
Negative	1.72	.09	.05
Positive	.53	.49	.55
Super positive	.12	.05	.03

^a Degrees of freedom = 1, 28. All *F*'s are nonsignificant. Neutral is not included in the analysis of multivariate *F*; all five variables form a linearly dependent set, since they sum to 1.0.

TABLE 3

RECIPROCITY ANALYSES, STUDY 1

Variable	Distressed	Nondistressed	Reciprocity X Groups ^a
Positive reciprocity			
$p(W+)$.340	.541	4.68*
$p(W+/H+)$.430	.561	
$p(H+)$.387	.635	.192
$p(H+/W+)$.350	.687	
Negative reciprocity			
$p(W-)$.298	.161	.26
$p(W-/H-)$.450	.167	
$p(H-)$.291	.063	3.04
$p(H-/W-)$.398	.075	

^a $p(W+)$ = the probability of a positive wife code, $p(W-)$ = the probability of a negative wife code, $p(H+)$ = the probability of a positive husband code, $p(H-)$ = the probability of a negative husband code.

^b Calculated by covariance analysis, with unconditional base rate as the covariate.

* $p < .05$.

when unconditioned base rates are initially low, greater gain in prediction may be more likely.

STUDY 2

Method

Subjects and selection. An independent sample of 14 clinic and 16 nonclinic couples was recruited for participation in the research project. All clinic couples were seeking therapeutic assistance for marital problems at either of two community mental health centers; nonclinic couples responded to a press release that briefly described the project. Both clinic and nonclinic couples were paid \$10 for participation. Precisely the same cutoff scores on the Locke-Wallace MRI as in Study 1 were used to select 12 distressed and 8 nondistressed couples for analysis; all 30 couples completed the experimental procedures. Husbands and wives in the two groups did not significantly differ in age, educational level, number of years married, or salary. Couples in the present study were on the average 32.50 years old, had been married an average of 9.44 years, had completed an average of 13.95 years of schooling, and earned an average salary of 11.44 thousand dollars. Thus, the couples in Study 2 had been married approximately 6 years longer than the couples in Study 1 and had approximately 2 years less of schooling. The major difference between the samples in the studies seems to be the length of time the couples were married.

Procedure. Based upon the results of Study 1, the procedure was modified as follows: First, two tasks were selected: the Inventory of Marital Conflict (IMC) as a high-conflict task and the food task as a low-conflict task. The IMC was selected because it is a standardized task, it appeared to induce a high degree of conflict in Study 1 in both groups (nearly all couples fought considerably on this task), and it could be expanded to include additional vignettes to ensure a sufficient sample of interactions

There was not a significant conflict effect, $F(1, 56) = 3.86$, $p > .05$, although the means were in the predicted direction (high-conflict $M = 3.17$; low-conflict $M = 3.39$). There was no significant sex main effect or any significant interactions. On the intent variable, there were no significant effects for any factor or for the interactions.

It was thus possible to discriminate distressed from nondistressed husbands and wives on how positively they coded their spouse's behavior regardless of the level of conflict the task induced. These results support the Jacob and Davis (1973) findings and extend these findings to high-conflict tasks.

Reciprocity. Four analyses of covariance were performed to compare $p(W+/H+)$ with $p(W+)$, $p(H+/W+)$ with $p(H+)$, $p(W-/H-)$ with $p(W-)$, and $p(H-/W-)$ with $p(H-)$ for distressed and nondistressed couples. Probabilities were separately computed for each couple. In each case the unconditional probability was the covariate. Because of the relatively low frequencies of the super-positive and super-negative codes, these were combined with positive and negative codes, respectively, for the sequential analysis. The test of the reciprocity hypothesis of distressed versus nondistressed marriage is the significance of the group F ratio on that portion of the conditional probability that cannot be predicted by the unconditional probability. For example, the residual in regressing $p(W+/H+)$ on $p(W+)$ must be greater in nondistressed couples than in distressed couples. This means that knowing that the husband has just coded the wife's behavior as positive increases the prediction over the base rate of $W+$ that the wife will subsequently code the husband's behavior as positive. The direction of the prediction is obviously reversed for negative reciprocity; that is, there should be more negative reciprocity in distressed than in nondistressed couples.

Table 3 presents the results of the analyses of covariance. There is some evidence that there is more positive reciprocity in nondistressed couples than in distressed couples. Though the degree of gain in prediction from conditional to unconditional probabilities is small for nondistressed couples, the analysis of covariance controls for the possibility that

TABLE 4

UNIVARIATE *F* RATIOS AND MEANS FOR RELATIVE FREQUENCY OF SPOUSE'S CODING OF PARTNER'S BEHAVIOR, STUDY 2

Variable ^a	<i>F</i> ^b	Distressed average	Nondistressed average
Super negative	2.98	.05	.00
Negative	3.82	.18	.11
Positive	19.87*	.30	.55
Super positive	.21	.12	.15

^a Neutral is not included in the analysis of multivariate *F*, all five variables form a linearly dependent set, since they sum to 1.0.

^b degrees of freedom = 1, 36.

* $p < .001$

for stable estimates of conditional probabilities. The food task was selected because it induced little conflict in either group in Study 1. The IMC was expanded to include three additional vignettes

All sessions were conducted at one of two community mental health centers nearest the couple's residence. It may be the case that the lack of differences on intent variables in Study 1 may have resulted from the couples not understanding that they were to code how they intended their messages to be taken by their spouses. More explanation and a specific example was provided for the intended positivity codes of the talk table than had been provided in Study 1.

Results and Discussion

Multivariate analysis of variance of impact data, with distressed-nondistressed as a between-subjects factor and husband-wife as a within-subjects factor resulted in a significant effect for the distress factor, Wilks-Lambda $F(4, 33) = 7.01$, $p < .001$, and no significant sex or Sex \times Distress interaction. Table 4 presents the univariate *F* ratios that contributed to the distress multivariate *F* ratio. Nondistressed spouses were more likely to code their partner's behavior as positive than their distressed counterparts. These results replicate those of Study 1. A comparison of the multivariate *F* ratios suggests that the results are even stronger here. This may be a function of the differences in the length of marriage in the two studies, since marital happiness is known to decrease with length of time married within the age ranges of the current two studies (Burgess et al., 1971).

Multivariate analysis of variance of spouses' intent resulted in no significant distress, sex, or Sex \times Distress *F* ratios. There were no significant differences in intention for either

husbands or wives on any category. This pattern of no differences is similar to that obtained in Study 1. For comparison with Study 1, Table 5 presents the means and univariate *F* ratios for distressed and nondistressed couples.

These results lend strong support to a communication deficit explanation of marital distress: In the two studies, although distressed and nondistressed couples do not significantly differ in the way they intend their messages to be received by their spouses, they do significantly differ in how the messages are actually received.

Conflict effect. Conflict effects were assessed in the same manner as in Study 1. An analysis of variance on the impact data found no significant main effect for conflict, $F(1, 72) = 2.29$, $p > .05$, although, again, results were in the predicted direction (high-conflict $M = 3.48$, low-conflict $M = 3.57$). This result is consistent with that of Study 1. As expected, there was a significant *F* ratio for the impact data on the distress factor, $F(1, 72) = 14.57$, $p < .001$. Distressed couples averaged 3.33, and nondistressed couples averaged 3.72 over all tasks.

There was a significant Distress \times Conflict interaction for the impact data, $F(1, 72) = 8.44$, $p < .01$. It was easier to discriminate distressed from nondistressed couples on the high-conflict task than on the low-conflict task. Distressed couples averaged .70 lower on impact scores than nondistressed couples on the high-conflict task; distressed couples were .13 higher on impact scores than nondistressed couples on the low-conflict task. There were no significant main effects or interaction effects on the intent variable.

TABLE 5

UNIVARIATE *F* RATIOS AND MEANS FOR RELATIVE FREQUENCY OF SPOUSES' INTENT, STUDY 2

Variable ^a	<i>F</i> ^b	Distressed average	Nondistressed average
Super negative	.66	.00	.00
Negative	.80	.04	.02
Positive	5.43*	.41	.60
Super positive	.23	.14	.10

^a Neutral is not included in the analysis of multivariate *F*, all five variables form a linearly dependent set, since they sum to 1.0.

^b Degrees of freedom = 1, 36

* $p < .05$.

The presence of Group \times Conflict interaction effects in the second study and their absence in the first study is interesting. It may be that the food task is a different kind of task for the couples of Study 2, since the husbands in this study are working and hence are less likely to have a role in food preparation than husbands in Study 1, who are more likely to be in school and share in the housework. The pilot study with both tasks, mentioned in the Procedure section of Study 1, found no significant sex differences in knowledge on either of the food or NASA tasks; however, a subsequent pilot study with college students found significant sex differences in self-ratings of confidence (on a 9-point scale) in the solution for the NASA task, $t(68) = 2.16$, $p < .04$, with females less confident than males. It may be that working husbands in Study 2 have as little confidence in their knowledge of food as the female undergraduates had of their knowledge of space. Both groups of husbands may have yielded more on this task and consequently been perceived as more positive by their wives. It may also be the case that the length of marriage factor was related to the obtained interaction effects. Perhaps couples who have been married longer (Study 2) have learned to exchange more positive messages in low-conflict situations and to avoid situations of high conflict. This is consistent with the clinical picture of the "stable-unsatisfactory" marriage (Lederer & Jackson, 1968) or the "united-front" couple (Kramer, 1968), who can present a satisfactory image of themselves to the public in social situations but are unable to hide behind this image when discussing real marital issues.

Overall, there seems to be evidence that the conflict level of the task influences how spouses code each other's behavior. In the first study it was equally easy to discriminate distressed from nondistressed couples whether they were discussing an unresolved personal issue or rank ordering a list of dog breeds. In the second study couples could be discriminated only on the high-conflict tasks. These results may account, in part, for the difficulty of obtaining consistent differences across studies between distressed and nondistressed families (Jacob, 1975). It would be wise to sample systematically from low- and

TABLE 6
RECIPROCITY ANALYSES, STUDY 2

Variable ^a	Distressed	Nondistressed	Reciprocity \times Groups F^b
Positive reciprocity			
$p(W+)$.448	.700	2.22
$p(W+/H+)$.536	.787	
$p(H+)$.405	.701	.58
$p(H+/W+)$.512	.756	
Negative reciprocity			
$p(W-)$.223	.141	.61
$p(W-/H-)$.354	.195	
$p(H-)$.235	.079	.15
$p(H-/W-)$.331	.121	

^a $p(W+)$ = probability of a positive wife code, $p(W-)$ = probability of a negative wife code, $p(H+)$ = probability of a positive husband code, $p(H-)$ = probability of a negative husband code.

^b Calculated by covariance analysis, with unconditional base rate as the covariate.

high-conflict tasks in future investigations of couples and family interaction. The results of the present investigation would suggest that high-conflict tasks are more reliable for discrimination between distressed and nondistressed couples.

Reciprocity. Table 6 presents the results of the reciprocity analyses. Reciprocity variables did not discriminate between distressed and nondistressed couples. In addition, the one significant positive reciprocity effect of Study 1 was not replicated.

The results of these two studies must bring into question the current marriage counseling image of positive reciprocity as characteristic of nondistressed marriages and not characteristic of distressed marriages. The high base rate of behavior that is positively coded by observers or by spouses may *seem* to be equivalent to reciprocity, but it is not. The data from the present investigation support a "bank account" model of nondistressed marriage rather than a reciprocity model. In a bank account model, a nondistressed marriage differs from a distressed marriage in that there are more positive "deposits" than negative "withdrawals." In a nondistressed marriage the consequent positive impact codes are not contingent upon the spouse's antecedent coding. Perhaps it is precisely this *lack* of reciprocity in a context of high positive exchange that characterizes stable positive interaction in nondistressed couples.

The results of the present investigation have demonstrated the usefulness of the talk table as a means for operationalizing behavior exchange theory. Results obtained are generally consistent with more expensive coding studies. Furthermore, the power of variables obtained by the couple coding their own behavior to discriminate between distressed and nondistressed couples is somewhat greater than that obtained by observer-coding studies. This result may have some theoretical import. Perhaps the differences between observer and spouse coding could account for some of the unreliability inherent in present coding systems. If one observer focuses on the facial expressions of the sender of a message, a positive code could result from a reading of intention; if the other coder focuses on the message's impact on the receiver, a negative code could result from a reading of the impact of the message. It would be fruitful to control the cues that observers use and to study those instances when observers and spouses are discrepant. This would be one way to study the development of a private message system in close relationships in which many messages do not mean to the stranger what they do to the person for whom they were intended. Work along these lines is in progress in our laboratory.

One limitation of the present investigation deserves mention. Because of the arrangements with referring agencies and the nature of the advertisement, both couples and interviewers were aware of which couples were seeking or starting marriage counseling. The MRI scores were, however, not available to either the interviewers or the couples. This aspect of the design of the present investigation introduces an experimenter expectancy criticism and the possibility that couples may be presenting themselves in order to be consistent with their classification. Subsequent investigations could improve on the current procedures by advertising for couples without recruiting couples from therapeutic agencies. Couples could be split on MRI scores, and the interviewers could remain blind to a couple's classification. Such a procedure, however, may draw different samples of couples than the present investigation, which used a convergence of two definitions of marital distress.

REFERENCE NOTE

1. Gottman, J., Notarius, C., & Markman, H. *The couples interaction scoring system (CISS)*. Unpublished manuscript, 1976. (Available from John Gottman, Department of Psychology, University of Illinois, Children's Research Center, 51 E. Gerty Drive, Champaign, Illinois 61820.)

REFERENCES

- Alexander, J. F. Defensive and supportive communications in family systems. *Journal of Marriage and the Family*, 1973, 35, 613-617.
- Azrin, N. H., Naster, B. J., & Jones, R. Reciprocity counseling: A rapid learning-based procedure for marital counseling. *Behavior Research and Therapy*, 1973, 11, 365-382.
- Birchler, G. R., Weiss, R. L., & Vincent, J. P. Multi-method analysis of social reinforcement exchange between maritally distressed and nondistressed spouse and stranger dyads. *Journal of Personality and Social Psychology*, 1975, 31, 349-360.
- Burgess, E. W., Locke, J. J., & Thomas, M. M. *The family*. New York: Van Nostrand Reinhold, 1971.
- Gergen, K. *The psychology of behavior exchange*. Reading, Mass.: Addison-Wesley, 1969.
- Haley, J. Research on family patterns: An instrument measurement. *Family Process*, 1964, 3, 41-65.
- Haley, J. Speech sequences of normal and abnormal families with two children present. *Family Process*, 1967, 6, 81-97.
- Hall, J. Decisions, decisions, decisions. *Psychology Today*, November, 1971, 51-88.
- Jacob, T. Family interaction in disturbed and normal families: A methodological and substantive review. *Psychological Bulletin*, 1975, 82, 33-65.
- Jacob, T., & Davis, J. Family interaction as a function of experimental task. *Family Process*, 1973, 12, 415-427.
- Kramer, C. H. The theoretical position: diagnostic and therapeutic implications. In Charles Kramer, Bernard Liebowitz, Robert Phillips, Sylvia Schmidt, & James Gibson (Eds.), *Beginning phase of family treatment*. Chicago: Family Institute of Chicago, 1968.
- Lederer, W. J., & Jackson, D. D. *The mirages of marriage*. New York: Norton, 1968.
- Locke, H. J. *Prediction adjustment in marriage: A comparison of a divorced and happily married group*. New York: Holt, 1961.
- Navran, L. Communication and adjustment in marriage. *Family Process*, 1967, 6, 173-180.
- Olson, D. H., & Ryder, R. G. Inventory of marital conflicts (IMC). An experimental interaction procedure. *Journal of Marriage and the Family*, 1970, 32, 443-448.
- Patterson, G. R. A basis for identifying stimuli which control behaviors in natural settings. *Child Development*, 1974, 45, 900-911.
- Rappaport, A. F., & Harrell, J. A behavioral-exchange model for marital counseling. *Family Coordinator*, 1972, 21, 203-212.

- Raush, H. L., Barry, W. A., Hertel, R. K., & Swain, M. A. *Communication, conflict and marriage* San Francisco: Josey-Bass, 1974.
- Riskin, J., & Faunce, E. E. An evaluative review of family interaction research. *Family Process*, 1972, 11, 365-455.
- Speer, D. C. Marital dysfunctionality and two-person non-zero-sum game behavior. *Journal of Personality and Social Psychology*, 1972, 21, 18-24.
- Stuart, R. B. Operant interpersonal treatment for marital discord. *Journal of Consulting and Clinical Psychology*, 1969, 33, 675-682.
- Terman, L. M., & Wallin, P. The validity of marriage prediction and marital adjustment test. *American Sociological Review*, 1949, 14, 497-504.
- Thibaut, J. W., & Kelley, H. H. *The social psychology of groups*. New York: Wiley, 1959.
- Weiss, R. L., Hops, H., & Patterson, G. R. A framework for conceptualizing marital conflict: A technology for altering it, some data for evaluating it. In Leo A. Hamerlynck, Lee C. Handy, & Eric J. Mash (Eds.), *Behavior Change: The Fourth Banff Conference on Behavior Modification*. Champaign, Ill: Research Press, 1973.

(Received April 27, 1975)