

JAMES A. COAN *University of Virginia*

JOHN M. GOTTMAN *Relationship Research Institute**

Sampling, Experimental Control, and Generalizability in the Study of Marital Process Models

The Oregon Youth Study has yielded invaluable insights into the adult antisocial lifestyle (Shortt, Capaldi, Dishion, Bank, & Owen, 2003), deviant interpersonal processes (Dishion, Nelson, Winter, & Bullock, 2004), psychological and physical aggression within romantic couples (Capaldi & Crosby, 1997; Capaldi, Shortt, & Crosby, 2003), and even the influence of male peer groups on aggressive behaviors directed at romantic partners (Capaldi, Dishion, Stoolmiller, & Yoerger, 2001). Kim, Capaldi, and Crosby (this volume) have framed their latest article describing Oregon Youth Study couples as a test of the generalizability of marital process models presented in Gottman, Coan, Carrere, and Swanson (1998, referred to by Kim et al. as *the Newlywed Study*).

Citing Lykken (1968), Kim et al. remind readers that the gold standard of repeatability for any study is *constructive* replication, not literal replication, and especially not statistical significance. They emphasize that constructive replication does not refer to the literal duplication of some earlier study's sampling procedures, research methodology, specific measurement approaches, or

even specific results but rather refer to the replication of *conclusions* about a set of theoretical propositions. We wholeheartedly agree with this perspective but also note that the theoretical propositions that formed the basis of the interest of Kim et al. in constructively replicating the process models described in Gottman et al. (1998) in their sample of lower income, at-risk couples were somewhat unclear. Although they describe their report's purpose as being to examine whether the process models described in Gottman et al. would also be predictive in a sample with differing characteristics from that used in the Newlywed Study, they offer little else in the way of specific rationales for their hypotheses.

In contrast to Kim et al., we would not expect many of the process models described in the Newlywed Study to generalize to the Oregon Youth Study. On the one hand, this is on the basis of our own experience in studying patterns of affect within lower income, at-risk couples (e.g., Babcock, Waltz, Jacobson, & Gottman, 1993; Coan, Gottman, Babcock, & Jacobson, 1997; Jacobson, Gottman, Gortner, Berns, & Shortt, 1996; Jacobson et al., 1994) but also because of the important differences between married and cohabiting couples, differences that are relatively independent of factors such as income (e.g., Dush, Cohan, & Amato, 2003). On the other hand, we would expect to see higher

Department of Psychology, University of Virginia, 102 Gilmer Hall, P.O. Box 400400, Charlottesville, VA 22904-4400 (jcoan@virginia.edu).

*Relationship Research Institute, P.O. Box 15644, Seattle, WA 98115-0644 (john@gottmanresearch.com).

levels of negative affect reciprocity in the Oregon Youth Study and that if any of the process models described in the Newlywed Study would be found to be of consequence in Oregon Youth Study interactions, it would be the male escalation of negative affect. As it happens, and as we describe in greater detail below, these predictions are consistent with what Kim et al. actually report.

In this commentary, we highlight that (a) in contrast to the Newlywed Study, the Oregon Youth Study sought out and utilized a sample intended to describe a unique population, one that is of lower income and at higher risk for antisocial behavior; (b) some of the experimental procedures used in the Oregon Youth Study may limit the generalizability of conclusions that can be drawn from it; and (c) despite these potential difficulties, the findings they report are consistent with other research on at-risk couples, and even to a limited extent the Newlywed Study itself.

SAMPLING

At-Risk Couples

Among the goals of the Newlywed Study was to build an empirically tested process model of marriage from detailed observations of marital interactions in the early stages of normative marital relationships. To be eligible for the Newlywed Study, couples had to have married for the first time within 6 months of participating in the study and they had to be childless. The demographic breakdown of the sample was representative of the greater Seattle area, according to the Seattle City Metropolitan Planning Commission Report (Gottman et al., 1998). This was of course one of the goals of the newlywed sampling procedure, because it was deemed important to recruit a sample that represented a broad population of married individuals.

Kim et al. assert that the Oregon Youth Study sample is similarly comparable to representative national U.S. samples, noting that the overall prevalence of physical aggression among Oregon Youth Study participants at approximately 18 years old was about 36% for women and 31% for men. Earlier reports referencing the Oregon Youth Study, however, have emphasized that the sample was specifically selected to represent a relatively rare population, one of lower income and at higher risk for antisocial behavior. This strategy has been successful. Twenty-five percent

of the boys in the Oregon Youth Study were arrested before age 14, and 29% were arrested between ages 14 and 17 (Shortt et al., 2003). The Oregon Youth Study sample has been explicitly described as consisting of "at-risk young men" (Capaldi et al., 2001; Shortt et al.) and, later, of "at-risk young couples" (Capaldi et al., 2003). Capaldi et al. (2003) have reported that among the young adults in the Oregon Youth Study sample, the frequency of one or more acts *per week* of physical aggression toward a partner in the past year was much higher (9% of the young men and 13% of the young women) than in a national survey sample (Straus, 1990), where less than one tenth of 1% endorsed such behavior. They concluded for this and other reasons that physical aggression was "a substantial problem" (Capaldi et al., 2003, p. 5) among Oregon Youth Study couples. Other characteristics in the report of Kim et al. suggest that the Oregon Youth Study couples are particularly aggressive. For example, 37% of the couples in the Oregon Youth Study had separated by Time 2, nearly triple the proportion of divorced couples in the Newlywed Study, which is characteristic of domestically violent couples (Gortner, Jacobson, Berns, & Gottman, 1997).

The Oregon Youth Study sample has been considered ideal for the study of social processes related to antisocial behavior (Dishion et al., 2004; Shortt et al., 2003), and with good reason, but it is probably less than ideal as a representative of the broader population. Indeed, the Oregon Youth Study sample resembles the more aggressive tail of the Newlywed Study distribution. A lack of broad variability in behavior (relative to the Newlywed Study) may have limited the number of processes observed among the Oregon Youth Study couples.

Cohabiting Couples

The lower income and more aggressive characteristics of the Oregon Youth Study sample set it apart from the newlyweds or restrict its generalizability to the same broad population as the newlyweds, but an additional problem concerns the mixing of cohabiting and married couples. Blumstein and Schwartz (1983) have noted that with respect to conflict, power, sex, and money, married, gay, and lesbian couples appear more similar to one another than any of them are to heterosexual cohabiters. Although the expectation was that the longer the cohabiting heterosexual

couples were together, the more they would resemble married couples, the opposite was true: In extensive interviews, cohabiters emphasized their *lack* of commitment to their partners. The extensive literature review of Waite and Gallagher (2000) has further suggested that young cohabiting couples in the United States can be reliably discriminated from young married couples on, for example, the basis of prior high school records that show at-risk behavior. It also suggested that few if any of the positive protective features of marriage are conferred upon U.S. cohabiters. Many studies confirm or extend these conclusions. Cohabiting relationships are, for example, generally less stable than marriages (Brown, 2003). Cohabiters tend to report lower relationship quality than married individuals (Brown & Booth, 1996), demonstrate more negative and less positive problem-solving and support behaviors than married couples (Cohan & Kleibnaum, 2002), and report lower levels of happiness and general well-being than their married counterparts (Dush & Amato, 2005; Skinner, Bahr, Crane, & Call, 2002). Cohabiters are even more likely to physically harm or kill each other than are married individuals (Shachelford & Mouzos, 2005).

Among the couples in the Oregon Youth Study sample, only about 30% were married at Time 1. By Time 2, about 60% of the couples still together were married. This shift warrants a more detailed look at the changes in the sample from Time 1 to Time 2. At Time 1, the total number of intact couples was 85. By Time 2 that total had shrunk to 54, meaning that 31 couples had separated. The number of married couples from Time 1 to Time 2 increased from 28 to 32, whereas the number of cohabiting couples from Time 1 to Time 2 decreased from 57 to 22, a difference of 35 couples. Information about the number of married couples who eventually divorced is not provided, so the ratio of divorced to separated couples at Time 2 is unknown. Nevertheless, we do know that from Time 1 to Time 2, at least four cohabiting couples married. Indeed, it is reasonable to suppose that all 31 separated couples had been cohabiting at Time 1 and no married couples divorced. Put another way, every one of the couples who were separated by Time 2 may have been (a) from low-income backgrounds, (b) substantially more aggressive with each other than nationally representative samples, and (c) cohabiting. We in fact feel that in all likelihood the divorce rate was substantially lower than the cohabiting-separation

rate. Unfortunately, there is little hope of addressing such a problem with statistical controls. If all separated couples had previously been cohabiting, predictions of relationship status are simply confounded by marital status. Even if the proportion of divorced to separated couples is relatively matched, however, statistical adjustments are unlikely to render marital status differences irrelevant, and in any case, the interpretation of results following such statistical adjustments is exceedingly difficult (cf. Miller & Chapman, 2001).

Shifts between measurement occasions hold implications for those couples who stayed together as well. By Time 2, the whole sample had shed a substantial portion (almost 40%) of its presumably less stable and predominantly (if not entirely) cohabiting couples, leaving 54 couples—now mostly married—with which to predict Time 2 satisfaction scores. It is reasonable to assume that this subsample of 54 couples is a little closer to the center of the newlywed distribution, just as the newlywed divorcees were a little further out on the tails. Indeed, results reported by Kim et al. suggest that the newlywed divorcees are strikingly similar to the Oregon Youth Study sample of stable unsatisfied couples at Time 2 (see Table 1).

GENERALIZABILITY AND EXPERIMENTAL CONTROL

The trade-off between generalizability and experimental control is well known and much discussed, especially in the context of quasi-experimentation (Shadish, Cook, & Campbell, 2002). In the Newlywed Study, conflict discussion topics were selected without specific constraints on whose particular issue (husband's or wife's) was discussed. Instead, the primary considerations were the duration (as well as whether the topic was currently an issue) and severity of potential areas of disagreement. In this context, and in others like it, the vast majority of topics ultimately discussed by couples could nominally be described as the wife's. The obvious potential downside of this approach is that our process models are biased toward wife issues. This may, however, be in itself a highly generalizable finding. Women typically start most marital conflict discussions in laboratories that use observational methods (Ball, Cowan, & Cowan, 1995; Oggins, Veroff, & Leber, 1993). Thus, the frequency of wife issues is likely to be higher than the

Table 1. Comparing Newlywed Study Separation Effects to Oregon Youth Study Satisfaction Effects in Terms of Anger Versus High- and Low-Intensity Negative Affect, Means in Seconds

	Newlywed Status	Oregon Youth Study Satisfaction
Anger		
Men	Not significant	Not significant
Women	Not significant	Not significant
Low negativity		
Men	Not significant	Greater in unsatisfied couples
Women	Greater in divorced couples	Greater in unsatisfied couples
High negativity		
Men	Greater in divorced couples	Greater in unsatisfied couples
Women	Greater in divorced couples	Greater in unsatisfied couples
Total positivity		
Men	Greater in stable couples	Greater in satisfied couples
Women	Greater in stable couples	Greater in satisfied couples
Positivity/(Pos. + Neg.)		
Men	Greater in stable couples	Greater in satisfied couples
Women	Greater in stable couples	Greater in satisfied couples

frequency of husband issues, in most relationships, most of the time. It is important to note that in the Newlywed Study procedure, both husbands and wives were asked about areas of disagreement in their relationship, as they saw them, during the period preceding the conflict discussion. Although husbands and wives were not specifically required to raise an issue of their own, they were each given the opportunity to do so. (See Roberts, Tsai, & Coan, in press, for a methodological discussion of dyadic interaction tasks.)

Interestingly, in the longitudinal study of domestically violent couples of Jacobson and Gottman (1999), most of the selected areas of disagreement were, using the identical procedure, indeed the husband's, and the analysis of Babcock et al. (1993) of these data using the Christensen demand-withdraw coding system (e.g., Christensen & Heavey, 1990) confirmed a preponderance of male demand to female withdraw, the reverse of the typical gender pattern. This illustrates the value of measuring whose issue (the husband's or the wife's) is predominantly associated with the most severe difficulties a couple faces. In any case, who starts the discussion, and the way a conflict discussion starts, is critical in determining its outcome. In one study, the way conflict began determined its subsequent course in 96% of marital conflict interactions (Gottman, 1994; Gottman et al., 1998, p. 7).

From this perspective, it may be more realistic to think of both wife issues and husband issues as *couple* issues.

For the Oregon Youth Study, Kim et al. were careful to avoid biases in the overrepresentation either of husband or of wife issues by requiring couples to talk separately for 7 minutes about each partner's issue. We have no strong complaint with this approach and indeed feel that it can probably yield very important information. Such careful balancing of male and female issues, however, may introduce a threat to ecological validity. It may, for example, result in conflict discussions that are not as likely to reflect what marital therapists, to take one "real-world" example, encounter in their practices. As mentioned above, the primary selection criteria for areas of disagreement in the Newlywed Study were duration and severity, leaving the designation of wife or husband issues measured but uncontrolled. In addition to our ecological validity concerns, we fear that requiring specific husband and wife areas of disagreement in the experimental procedure introduces the danger of diluting the impact of problem duration and severity on the conflict discussions, thereby altering the nature of the interactions. This could happen if, for example, less severe issues are overrepresented inadvertently in the attempt to ensure that both spouses raise concerns designated as their own. In light of these considerations, it would be interesting to know

whether the duration and the severity of the areas of disagreement solicited in the procedure of Kim et al. differed as a function of husband and wife issues.

These potential problems notwithstanding, it remains possible to examine the process models Kim et al. tested in the context of women's issue discussions alone. When this was done, three processes predicted separation (see Table 2). They are male high-intensity negative reciprocity, male negative escalation, and female de-escalation of high-intensity affect. The male escalation of negative affect was identified in the Newlywed Study as one of the most powerful predictors of divorce. It has also been found to distinguish violent from nonviolent couples and to play a role in the severity of violence among domestically violent couples (Coan et al., 1997). Interestingly, male high-intensity negative reciprocity has been found to distinguish violent from nonviolent couples as well (Coan et al.), again suggesting that the Oregon Youth Study sample leans toward the violent end of the couple spectrum.

One of the most interesting wife issue sequences found to distinguish stable from separated couples in the Oregon Youth Study was male negative start-up, higher levels of which were actually predictive of relationship stability. Kim et al. describe this as a counterintuitive finding, and so it is, but the literature on domestically violent couples offers some clues as to what might be going on. For example, married couples with higher levels of husband demand/wife withdrawal are more likely to be violent (Babcock et al., 1993; Sagrestano, Heavey, & Christensen, 1999). This does not in itself explain why couples in the Oregon Youth Study with higher levels of male negative start-up are more stable. Given the high separation rate in the Oregon Youth Study, however, and the fact that few or no couples in the Oregon Youth Study actually divorced, one simple explanation is that married couples in the Oregon Youth Study look like violent married couples among whom male demand/female withdrawal is a defining characteristic (e.g., Babcock et al.). It would in any case be very interesting to learn the difference in the frequency of male negative start-up between stable married and stable cohabiting couples. (Indeed, it would be very interesting to learn the difference between stable married and stable cohabiting couples on any or all of the dimensions Kim et al. sought to investigate.)

Table 2. Comparing Newlywed Study Affect Sequence Effects to Oregon Youth Study Affect Sequence Logistic Regression Effects Culled From Conflict Discussions of the "Woman's Issue"

	Newlywed Status	Oregon Youth Study Woman's Issue Status
Reciprocity of low-intensity negativity		
Men	Not significant	Not significant
Women	Greater in divorced	Not significant
Reciprocity of high-intensity negativity		
Men	Not significant	Greater in separated ^{a,b}
Women	Not significant	Not significant
Negative escalation		
Men	Greater in divorced	Greater in separated ^{a,b}
Women	Not significant	Not significant
Negative start-up		
Men	Not significant	Lower in separated
Women	Greater in divorced	Not significant
De-escalation of low-intensity negativity		
Men	Lower in divorced	Not significant
Women	Not significant	Not significant
De-escalation of high-intensity negativity		
Men	Not significant	Not significant
Women	Not significant	Lower in separated

^aAlso true of the full Oregon Youth Study sample after controlling for demographic variables. ^bThis pattern was also observed among domestically violent couples in Coan et al. (1997).

SEQUENTIAL ANALYSIS

Kim et al. raise an additional concern about statistically controlling for antecedents and consequents in joint frequency analyses. As many readers will know, in sequential analysis, the primary issue concerns how to describe sequential connections,

the fundamental question being whether knowledge of the antecedent behavior reduces uncertainty in our prediction of the consequent behavior in comparison to just predicting the consequent behavior from its unconditional base rate. Of course, sequential analysis does not provide a complete picture. We also want to know the base rates in order to describe how often these events (consequent and antecedent) actually occurred. In the Newlywed Study, we conducted covariance analyses using consequent codes as covariates following our initial joint frequency analyses. This step was in fact statistically conservative. Such statistical adjustments help rule out the effects of relatively rare instances of a given consequent, as might similar adjustments of the antecedent advocated by Kim et al. In any case, there are many ways to do these sequential analyses, and most of them are statistically similar. We refer the reader to the book *Sequential Analysis* of Gottman and Roy (1990); the introduction to sequential analysis of Bakeman and Gottman (1997); the work on identifying sequences of Bakeman and Quera (1995), with their General Sequential Querier program; and the discussion of their log-linear method of Bakeman and Robinson (1994), for more complete discussions of sequential analytic strategies. More recently, these linear analyses have been updated with nonlinear difference equation modeling of connection and influence in the book *The Mathematics of Marriage*, by Gottman, Murray, Swanson, Tyson, and Swanson (2002).

CONCLUSIONS

The particular problems associated with lower income, at-risk early childhood environments in establishing and maintaining satisfying and stable adult relationships are poorly understood and in need of serious and committed study. Few if any individuals are better placed than Kim et al. to make important inroads into the resolution of such problems, and endeavors such as the Oregon Youth Study stand to provide insights that can substantially affect the health and well-being of the population the Oregon Youth Study sample was intended to represent. We feel, however, that the theoretical rationale Kim et al. provide for their attempt to replicate the Newlywed Study process models of Gottman et al. (1998) in this sample was incomplete. Neither the suggestion that the Newlywed Study has been highly influential nor the fact that it may have inspired

controversy provides strong reasons in themselves to expect the models to generalize to the Oregon Youth Study sample. Indeed, it seems to us on a priori theoretical grounds more likely that the newlywed process models should *not* generalize to their sample and that their sample should instead more closely resemble other samples of at-risk couples. The results Kim et al. report suggest that this is in fact the case.

The Oregon Youth Study was designed from the beginning to represent a relatively rare and troubled population of lower income, at-risk individuals. Their study suggests that at-risk couples behave in ways that are reliably distinct from the broader population of married couples. As we have pointed out, other studies have produced similar results. Unfortunately, the fact that many of the couples in the Oregon Youth Study were not actually married and that married and cohabiting couples were unequally distributed across stable and nonstable relationships introduces an additional level of complexity in interpreting the results of Kim et al.. The separated group was likely composed predominantly (perhaps exclusively) of previously cohabiting couples, whereas the stable group provides more of a mixture of cohabiting and married couples. This situation is troubling, because there are known, reliable differences between married and cohabiting couples that may be difficult or impossible to control for statistically.

As for generalizability, one important consideration is the way Kim et al. chose to structure their couples' conflict discussions. In their design, men and women were each required to speak for 7 minutes about their own respective issues. Kim et al. assert that tests of gender differences require experimental conditions such as this, but generalization is about experimental conditions at least as much as it is about sampling, and there is a balance to be struck between the risks Kim et al. are concerned about and the risk of placing procrustean experimental restrictions on otherwise naturally occurring systems. Although such findings may provide glimpses of what couples are, strictly speaking, capable of, there is also the danger that such experimentally controlled conditions will not generalize well outside of the laboratory. A more generalizable task may have allowed areas of disagreement to be selected solely as a function of duration and severity. Had this been done, it would have remained possible to measure whether husband or wife issues were more or less likely to be

discussed, as well as to monitor the implications of any such differences that emerged. Although it is impossible to know what their data would look like had they not structured their couples' conflict discussions in this way, a look at the discussions deriving from only the wife's issue reveals that couples from the Oregon Youth Study look substantially like violent couples studied elsewhere (Babcock et al., 1993; Berns, Jacobson, & Gottman, 1999; Coan et al., 1997; Sagrestano et al., 1999).

Ultimately, we feel the conclusions that can be drawn from the latest work by Kim et al. must be qualified by the fact that their sample consisted of lower income, at-risk, mixed married and cohabiting couples, as well as the fact that their experimental procedures may have limited the generalizability of their results. We are also somewhat skeptical of the theoretical rationale for suggesting that the models discussed in the Newlywed Study should generalize to their at-risk sample. Nevertheless, we recognize the principle underlying the important work of Kim et al. and understand that although constructive replication is very difficult, it is also highly desirable. Their work has added a valuable layer of depth to our understanding of affective process models in married and cohabiting couples.

REFERENCES

- Babcock, J. C., Waltz, J., Jacobson, N. S., & Gottman, J. M. (1993). Power and violence: The relation between communication patterns, power discrepancies, and domestic violence. *Journal of Consulting and Clinical Psychology, 61*, 40–50.
- Bakeman, R., & Gottman, J. M. (1997). *Observing interaction: An introduction to sequential analysis* (2nd ed.). New York: Cambridge University Press.
- Bakeman, R., & Quera, V. (1995). *Analyzing interaction: Sequential analysis with SDIS and GSEQ*. New York: Cambridge University Press.
- Bakeman, R., & Robinson, B. F. (1994). *Understanding log-linear analysis with ILOG*. Hillsdale, NJ: Erlbaum.
- Ball, F. L. J., Cowan, P., & Cowan, C. P. (1995). Who's got the power? Gender differences in partner's perception of influence during marital problem-solving discussions. *Family Process, 34*, 303–321.
- Berns, S. B., Jacobson, N. S., & Gottman, J. M. (1999). Demand-withdraw interaction in couples with a violent husband. *Journal of Consulting and Clinical Psychology, 67*, 666–674.
- Blumstein, P., & Schwartz, P. (1983). *American couples*. New York: William Morrow.
- Brown, S. L. (2003). Relationships quality dynamics of cohabiting unions. *Journal of Family Issues, 25*, 583–601.
- Brown, S. L., & Booth, A. (1996). Cohabitation versus marriage: A comparison of relationship quality. *Journal of Marriage and Family, 58*, 668–678.
- Capaldi, D. M., & Crosby, L. (1997). Observed and reported psychological and physical aggression in young, at risk couples. *Social Development, 6*, 184–206.
- Capaldi, D. M., Dishion, T. J., Stoolmiller, M., & Yoerger, K. (2001). Aggression toward female partners by at-risk young men: The contribution of male adolescent friendships. *Developmental Psychology, 37*, 61–73.
- Capaldi, D. M., Shortt, J. W., & Crosby, L. (2003). Physical and psychological aggression in at-risk young couples: Stability and change in young adulthood. *Merrill-Palmer Quarterly, 49*, 1–27.
- 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.
- Coan, J., Gottman, J. M., Babcock, J., & Jacobson, N. (1997). Battering and the male rejection of influence from women. *Aggressive Behavior, 23*, 375–388.
- Cohan, C. L., & Kleibnaum, S. (2002). Toward a greater understanding of the cohabitation effect: Premarital cohabitation and marital communication. *Journal of Marriage and Family, 64*, 180–192.
- Dishion, T. J., Nelson, S. E., Winter, C. E., & Bullock, B. M. (2004). Adolescent friendship as a dynamic system: Entropy and deviance in the etiology and course of male antisocial behavior. *Journal of Abnormal Child Psychology, 32*, 651–663.
- Dush, C. M. K., & Amato, P. R. (2005). Consequences of relationship status and quality for subjective well-being. *Journal of Social and Personal Relationships, 22*, 607–627.
- Dush, C. M. K., Cohan, C. L., & Amato, P. R. (2003). The relationship between cohabitation and marital quality and stability: Change across cohorts? *Journal of Marriage and Family, 65*, 539–549.
- Gortner, E., Jacobson, N. S., Berns, S., & Gottman, J. M. (1997). When women leave violent relationships: Dispelling clinical myths. *Psychotherapy, 34*, 343–352.
- Gottman, J. M. (1994). *What predicts divorce?* Hillsdale, NJ: Erlbaum.

- Gottman, J. M., Coan, J., Carrere, S., & Swanson, C. (1998). Predicting marital happiness and stability from newlywed interactions. *Journal of Marriage and Family, 60*, 5 – 22.
- Gottman, J. M., Murray, J. D., Swanson, C. C., Tyson, R., & Swanson, K. R. (2002). *The mathematics of marriage: Nonlinear dynamic models*. Cambridge, MA: MIT Press.
- Gottman, J. M., & Roy, A. K. (1990). *Sequential analysis: A guide for behavioral scientists*. New York: Cambridge University Press.
- Jacobson, N. S., & Gottman, J. M. (1999). *When men batter women*. New York: Simon & Schuster.
- Jacobson, N. S., Gottman, J. M., Gortner, E., Berns, S., & Shortt, J. W. (1996). Psychological factors in the longitudinal course of battering: When do the couples split up? When does the abuse decrease? *Violence and Victims, 11*, 371 – 392.
- Jacobson, N. S., Gottman, J. M., Waltz, J., Rushe, R., Babcock, J., & Holtzworth-Munroe, A. (1994). Affect, verbal content, and psychophysiology in the arguments of couples with a violent husband. *Journal of Consulting and Clinical Psychology, 62*, 982 – 988.
- Kim, H. K., Capaldi, D. M., & Crosby, L. (2007). Generalizability of Gottman and colleagues' affective process models of couples' relationship outcomes. *Journal of Marriage and Family, 69*, 55 – 72.
- Lykken, D. T. (1968). Statistical significance in psychological research. *Psychological Bulletin, 70*, 151 – 159.
- Miller, G. A., & Chapman, J. P. (2001). Misunderstanding analysis of covariance. *Journal of Abnormal Psychology, 110*, 40 – 48.
- Oggin, J., Veroff, J., & Leber, D. (1993). Perceptions of marital interactions among black and white newlyweds. *Journal of Personality and Social Psychology, 65*, 494 – 511.
- Roberts, N. A., Tsai, J. L., & Coan, J. A. (in press). Emotion elicitation using dyadic interaction tasks. In J. A. Coan & J. J. B. Allen (Eds.), *The handbook of emotion elicitation and assessment*. New York: Oxford University Press.
- Sagrestano, L. M., Heavey, C. L., & Christensen, A. (1999). Perceived power and physical violence in marital conflict. *Journal of Social Issues, 55*, 65 – 79.
- Shachelford, T. K., & Mouzos, J. (2005). Partner killing by men in cohabiting and marital relationships: A comparative, cross-national analysis of data from Australia and the United States. *Journal of Interpersonal Violence, 20*, 1310 – 1324.
- Shadish, W. R., Cook, T. D., & Campbell, D. T. (2002). *Experimental and quasi-experimental designs for generalized causal inference*. Boston: Houghton-Mifflin.
- Shortt, J. W., Capaldi, D. M., Dishion, T. J., Bank, L., & Owen, L. D. (2003). The role of adolescent friends, romantic partners, and siblings in the emergence of the adult antisocial lifestyle. *Journal of Family Psychology, 17*, 521 – 533.
- Skinner, K. B., Bahr, S. J., Crane, D. R., & Call, V. R. A. (2002). Cohabitation, marriage and remarriage: A comparison of relationship quality over time. *Journal of Family Issues, 23*, 74 – 90.
- Straus, M. A. (1990). Injury and frequency of assault and the "representative sample fallacy" in measuring wife beating and child abuse. In M. A. Straus & R. G. Gelles (Eds.), *Physical violence in American families: Risk factors and adaptations to violence in 8,145 families* (pp. 75 – 91). New Brunswick, NJ: Transaction.
- Waite, L., & Gallagher, M. (2000). *The case for marriage*. New York: Doubleday.