Journal of Family Psychology
1999, Vol. 13, No. 4, 663-668

0893-3200/99/$3.00

The Time Has Come to Talk of Many Things: A Commentary on Kurdek (1998) and the Emerging Field of Marital Processes in Depression

Steven R. H. Beach and Adam Davey
University of Georgia

Frank D. Fincham
State University of New York at Buffalo

In an important investigation of the longitudinal relation between marital adjustment and depressive symptoms, L. A. Kurdek (1998) reported an apparent failure to replicate prior findings reported by F. D. Fincham, S. R. H. Beach, G. T. Harold, and L. N. Osborne (1997). He attributed the diverging results to differences between the analytic approach taken in each of the studies. The authors reanalyzed F. D. Fincham et al.'s data using L. A. Kurdek's analytic procedures and found no change in results. A more parsimonious assessment of the diverging results is that L. A. Kurdek failed to generalize F. D. Fincham et al.'s findings to a longer time lag and a sample characterized by a rather different pattern of attrition. The authors highlight several important issues that require greater attention in developing a cumulative body of knowledge on the longitudinal relation between marital processes and depression.

> "The time has come," the Walrus said, "To talk of many things: Of shoes- and ships- and sealing wax- Of cabbages- and kings- And why the sea is boiling hot- And whether pigs have wings."
>
> —Lewis Carroll, *Through the Looking-Glass*

The robust association between depression and marital distress raises the question of whether a causal relationship exists between these two conditions. Two recent longitudinal studies examined this important question but found quite different results. Fincham, Beach, Harold, and Osborne (1997) found that husbands' depressive symptoms predicted their marital distress both concurrently and over an 18-month period, whereas wives' marital distress predicted their current and future depressive symptoms. In contrast, Kurdek (1998), using a similar sample of newlywed couples, failed to replicate these gender-linked relationships and found little evidence of a longitudinal relationship between levels of depressive symptoms and marital distress over a 4-year period. Such discrepant findings provide a shaky foundation for understanding the longitudinal relationship between depression and marital distress. We therefore address this apparent failure to replicate and, in doing so, identify factors that might provide a stronger foundation for the emergence of a cumulative body of knowledge on the potential causal relationship between depression and marital distress.

Of Dependence

Data obtained from couples pose interesting analytic choices (e.g., what is the appropriate unit of analysis and how should spouse scores be combined?). Kurdek (1998) partly attributed his failure to replicate Fincham et al.'s (1997) study to one such analytic choice: his modeling of husbands' and wives' data in a single causal model that controlled for interdependencies between their scores.

To examine the possibility that the results

Stephen R. H. Beach, Department of Psychology, University of Georgia; Adam Davey, Department of Child and Family Development, University of Georgia; Frank D. Fincham, Department of Psychology, State University of New York at Buffalo.

We each contributed equally to this article and are listed in alphabetical order. The preparation of the manuscript was supported by National Science Foundation Grant SBR-9511385 and by a grant from the Templeton Foundation.

Correspondence concerning this article should be addressed to Frank D. Fincham, Department of Psychology, Park Hall, State University of New York at Buffalo, Buffalo, New York 14260-4110. Electronic mail may be sent to fincham@buffalo.edu.

differed due to the different analytic approaches used, we reanalyzed the Fincham et al. (1997) data in the manner proposed by Kurdek (1998). Using a single model in which error terms for outcome variables (Time 2 depressive symptoms and marital distress) were allowed to correlate both within spouses and across spouses, we found that a cross-lagged stability model yielded parameter estimates for the cross lags that were almost identical to those originally reported (for depression leading to later marital distress, for husbands, −.28 vs. original −.31, for wives, .02 vs. original .03; for marital distress leading to later depression, for husbands, −.19 vs. original −.19, for wives, −.29 vs. original −.29). Unlike Kurdek's results, constraining the corresponding cross-lag parameters to be equal for husbands and wives yielded a significant deterioration in model fit, $\chi^2(2, N = 116) = 7.48$, $p < .05$, confirming that gender differences do exist in these data. Further reanalysis showed that the path from depressive symptoms to later marital distress was significantly stronger in husbands than in wives, $\chi^2(1, N = 116) = 6.95$, $p < .01$. Finally, allowing all cross-spouse variables to covary did not change these findings.

In his analysis of bidirectional or synchronous effects between depressive symptoms and marital distress, Kurdek (1998) found evidence that later marital quality influenced later depressive symptoms among husbands and that later depressive symptoms influenced later marital quality among wives. However, these apparent gender effects did not prove to be significant. In contrast, reanalysis of the Fincham et al. (1997) data, taking into account dependency between husbands' and wives' data, showed a robust gender difference. A model constraining the corresponding paths to be equal across gender differed significantly from one in which they were free to vary, $\chi^2(2, N = 116) = 6.48$, $p < .05$ (again the marital distress leading to depression path was significantly larger for wives than for husbands, $\chi^2[1, N = 116] = 6.48$, $p < .01$).

Finally, because Kurdek (1998; see Figures 2 and 3) constrained correlated errors for several cross-spouse correlations to be zero, we reanalyzed the data releasing these constraints. We found that freeing the constraints on cross-spouse effects reduced the previously significant cross-lagged effect to nonsignificance. Accordingly, it appears that when all cross-spouse effects were estimated freely, there was no evidence in Kurdek's data of any longitudinal relationship between marital quality and depressive symptoms.

In summary, Kurdek's (1998) improvement on Fincham et al.'s (1997) data analysis does not account for the discrepant longitudinal or concurrent connections between depression and marital quality. When all disturbance terms were allowed to correlate freely across spouses, there was evidence of a gender difference in the Fincham et al. data, but no evidence of any cross-lagged effects in the Kurdek data. This raises the question of what factors, other than data-analytic decisions, may account for the difference. Understanding this difference may help address problems common in longitudinal investigations and so begin to provide a foundation for the accumulation of findings across longitudinal data sets. Continuing our comparison of the Kurdek and Fincham et al. studies, we therefore identify several important factors that might influence the ability to detect any potential causal relationship between marital distress and depression.

Of Causal Lags

An important difference between the two studies is the time lag used to examine the potential prospective effect of each variable on the other. As noted by Fincham et al. (1997), we simply do not know the lag at which marital adjustment influences depression to the greatest extent or the lag at which depression influences marital adjustment to the greatest extent. However, in the Fincham et al. study, a comparison of simple recursive models and nonrecursive models suggested that the effect of marital adjustment on depression for wives might unfold relatively rapidly and at a lag of less than the 18 months used in that investigation. Conversely, the effect of depression on satisfaction for husbands appeared to be more adequately captured with an 18-month lag. At a minimum, this discussion underscores the fact that we have no a priori reason to expect the optimal lag time to be the same regardless of direction of effect, and there is no solid theoretical basis for hypothesizing a particular lag time as optimal.

It may be tempting in such a situation to assume that a longer lag is better or more

powerful than a shorter lag in that it allows more of the mediating events that connect the two constructs to unfold. Although Kurdek (1998) did not directly address the issue, the Fincham et al. (1997) comparison of synchronous and lagged models suggests that examination of relatively shorter lagged relationships may be more appropriate, particularly if one is attempting to capture the lagged effect of marital adjustment on depression. Kurdek indirectly addressed this issue in a footnote (p. 508) by noting that the use of a 2-year lag did not appreciably change the results of his cross-lag analysis over a 3-year period. However, because a 2-year lag is probably longer than is optimal, and longer than the time frame used by Fincham et al., it does not adequately respond to current concerns.

As noted earlier, our reestimation of parameters in Kurdek's (1998) model after freeing correlated error in disturbance terms resulted in no significant cross-lagged effects. Accordingly, it seems possible that the small cross-lagged effect from husbands' marital satisfaction to husbands' depression reported by Kurdek resulted from his having constrained the correlation between the disturbance terms for wives' Year 4 marital quality and husbands' Year 4 depressive symptoms to be zero despite an observed simple correlation of $-.37$ ($p < .01$). The reestimated model, with its uniformly null lagged effects, suggested the parsimonious conclusion that if there is a causal relationship between depression and marital satisfaction at work in the Kurdek data, it must act over an interval of less than 36 months (and probably less than 24 months). This would represent a point of convergence rather than a point of disagreement between the Kurdek study and the Fincham et al. (1997) study.

Of Participant Recruitment

The recruitment of participants may also be quite important in helping reconcile differing patterns of longitudinal findings across studies (cf. Glenn, 1998; Karney & Bradbury, 1995). The primary difference between Kurdek's (1998) recruitment and the Fincham et al. (1997) recruitment is that Kurdek used a mass-mailing approach to recruitment, whereas Fincham et al. used a telephone recruitment procedure after *identifying* potential participants. This may account in part for the very different rates of response in the two studies (in Kurdek's study 18% of couples indicated an interest in the study, and just under 7% of the original sample returned initial questionnaires; in the Fincham et al. study over 80% of couples approached about the study agreed to participate). Accordingly, whereas the initial pool of couples approached may have been similar, there is reason to suspect that the samples may have already diverged somewhat by the time of the first assessment.

Recruitment of a broadly representative sample is important if one plans to compare or combine results across samples. Although it is possible that a sample with a low response rate might, nonetheless, be representative in important respects, there are many variables that could influence both tendency to participate and marital quality and tendency to participate and rate of change in marital quality. For example, Krokoff (1990) demonstrated that husbands who were prone to avoidance were likely to be underrepresented in samples of married couples responding to requests to participate in studies of marriage. Presumably, avoidance is an example of the type of characteristic that could lead both to lower likelihood of responding to a letter soliciting participation and to behavior that would be problematic in a marriage. If so, a sample with a very low response rate may be relatively homogenous on a variety of risk factors and so tend to distort both concurrent and longitudinal relationships. This may be of less concern when the focus is on theoretically specified comparisons within the sample. It becomes more central, however, when the focus is on comparisons across samples.

Of Participant Attrition

Another aspect of longitudinal research that can matter a great deal in terms of observed relationships is level of participant attrition. In marital research, some loss of participants because of divorce or separation is inevitable, and this loss becomes greater as the time frame of the investigation increases. Especially in the early years of marriage, such losses are likely to reflect loss of the most discordant couples, as Karney and Bradbury (1995) and Kurdek (1998) have found that couples in early marriage who ultimately separate or divorce show a more rapid decline in marital satisfaction. Such losses can profoundly alter observed cross-sectional relationships at later lags, leading even to

counterintuitive reversals of effect (Glenn, 1998), and there will be at least some attenuation in the magnitude of longitudinal relationships. Loss of couples because of nonresponse is also a problem but one that is especially difficult to interpret. Such data loss may be random, but it may also be that important predictors discriminate those dropped from those retained in the sample.

In this regard, Kurdek (1998) reported 22 couples divorced between entry into the study and the end-point assessment, with an overall attrition of 45%. In contrast, Fincham et al. (1997) reported just 8 couples divorcing and an overall attrition of 22.6%. Accordingly, it is quite possible for substantively different effects to emerge across studies as a result of differential removal of weak marriages from the analyses. Because participant attrition is likely to attenuate correlations, such a difference would be most likely to manifest itself as weaker cross-sectional and longitudinal relationships between marital adjustment and depression in the Kurdek sample relative to the Fincham et al. sample.

Attenuated Correlations

Comparisons of the zero-order correlations reported in each study are consistent with our above concern. Time 1 marital adjustment correlated with Time 2 depression −.43 versus −.16 for the Fincham et al. (1997) and Kurdek (1998) samples, respectively. The corresponding correlation for husbands was −.33 versus −.25. It should be noted that with weaker cross-lagged correlations, it becomes increasingly difficult to reject the null hypothesis of no lagged effect of marital adjustment on depression and especially difficult to demonstrate the presence of significant gender differences. Accordingly, greater attrition related to the removal of weak relationships should tend to work against finding lagged effects from marriage to depression, and this may account, in part, for the divergence between the Kurdek and the Fincham et al. analyses.

Diverging Levels of Depressive Symptoms

Participant attrition due to loss of the weakest marriages could also result in relatively lower depressive symptoms over time. In this regard, it is interesting to note that depressive symptoms scores in Fincham et al.'s (1997) study showed a nonsignificant increase (from 4.1 to 4.5 for husbands and from 5.9 to 6.8 for wives on the Beck Depression Inventory; Beck, Steer, & Garbin, 1988), whereas in Kurdek's (1998) study, depressive symptoms showed a nonsignificant decrease over time (from 6.27 to 5.18 for husbands and from 8.63 to 7.91 for wives using a scale drawn from the Symptom Checklist; Derogatis, 1983). The attrition hypothesis might also help explain the negative stability coefficient found for marital satisfaction in Kurdek's growth-curve analysis. If there is selective attrition with those correctly assessed as dissatisfied dropping out but those falsely assessed as distressed remaining in the sample, there would be an exaggerated regression to the mean effect that could produce a negative correlation between initial value and slope. Consistent with this interpretation, mean levels of husbands', $t(135) = -6.92$, and wives', $t(135) = -5.51$, depressive symptoms showed significant declines from Wave 1 to Wave 4 for the sample of 136 individuals who completed all waves of Kurdek's study.

Dealing With Participant Attrition

One important implication of our discussion of attrition and its possible effects is that researchers in the marital area must come to grips with the problem of sample attrition in longitudinal designs. Fortunately, social scientists are rapidly finding new methods for addressing selective nonresponse. In particular, full information maximum-likelihood and multiple imputation procedures have recently been proposed as methods for estimating and correcting for patterns of missingness (e.g., Arbuckle, 1996; Davey, Halverson, McCrae, Zonderman, & Costa, 1999; Little & Rubin, 1989; Schafer, 1997). As these methods of dealing with nonrandom missingness gain more widespread use, researchers may be better positioned to accumulate knowledge across samples even when they vary widely in sample biases.

Of Correlated Change

Kurdek's (1998) use of growth-curve modeling is an important contribution to the study of marital processes and depression. It permits the observation that there is a correlation between the trajectory of change in marital quality and

the trajectory of change in depressive symptoms. On the basis of this observation, and in the absence of significant cross-lagged effects, he concluded that his findings support

> a "doubly-developmental" approach in which attention is directed to the trajectory of change in individual differences variables known to account concurrently for variability in marital quality, the trajectory of change in marital quality, and the link between these two trajectores (p. 508).

However, to the extent that the lack of lagged effects are due to the use of an overly long lag or problems in participant retention, this conclusion may be premature.

At a conceptual level it should be noted that the doubly developmental suggestion is a retreat from the attempt to establish patterns of causal connection rather than an advance in this direction. In particular, the doubly developmental approach identifies only correlated patterns of change without explicating the causal basis of such correlations. This may indeed be useful as a descriptive step that provides a foundation for causal theorizing, but it does not provide a substitute for causal models and the useful inferences such models allow.

Of Growth-Curve Modeling Through Hierarchical Linear Models (HLM)

Kurdek's (1998) use of HLM to investigate change is a novel contribution to the study of marital processes and depression and focuses attention on rates of change, in addition to absolute levels, over time. Individual or couple differences in such rates can also be examined along with variables that might predict rates of change. For example, specifying rate of change in marital satisfaction as a random effect in the Kurdek data would allow initial level of depression to predict individual differences in rate of marital change. However, even with careful parameterization, the use of version 4.03 of Bryk, Raudenbush, & Congdon's (1996) HLM software has certain important limitations.

One limitation of the Bryk et al. (1996) software is that it does not provide an absolute index of fit. For this reason it is not possible using HLM software to identify and reject models that may fit the data poorly. An index of fit is a standard part of the output when structural equation modeling (SEM) software is used. Recently, it has been noted by several authors that SEM software can be used to fit growth-curve models (e.g., Muthén & Curran, 1997; Raudenbush, in press; Willett & Sayer, 1994, 1996). Using LISREL to generate the implied covariance matrix for Kurdek's (1998) model provides an assessment of its implications for relationships among variables over time. In turn, even in the absence of an absolute index of fit, this procedure may raise questions about the plausibility of certain of these implied relationships. Therefore, we ran the HLM model proposed by Kurdek using LISREL 8.30 (Jöreskog & Sörbom, 1999).

The estimated model was found to imply a substantial negative correlation between Time 1 marital satisfaction and Time 4 marital satisfaction ($r = -.84$ and $-.32$, for husbands and wives, respectively). This result suggests poor fit of the model to the data both because negative stability coefficients for marital quality are unlikely and because the observed stability correlations (see Kurdek's, 1998, Table 4) are positive (.56 for wives, .57 for husbands). Accordingly, it appears likely that the HLM model estimated by Kurdek does not provide a good fit to the observed data.

The fact that it provides estimates that are a relatively poor fit to the data without a clear mechanism for alerting the researcher to this fact is one drawback of version 4.03 of Bryk et al.'s (1996) HLM software. In particular, because this software does not define a baseline saturated model against which to compare models of interest, it provides only a logliklihood function rather than a goodness-of-fit index, which could quickly identify problematic solutions. In the current case, a poor fit index might have suggested the possibility that the estimated HLM model needed to be modified and might have prompted a search for any unlikely values that were being generated.

Conclusion

The study reported by Kurdek (1998) is a laudable attempt to replicate and extend a prior finding. However, Kurdek's improved data-analytic procedures do not account for the apparent failure to replicate, as their application to the Fincham et al. (1997) data set did not alter the originally reported findings. This conclusion begs the question as to why the results of the two studies differ. Close examination showed important differences between the studies and raised

the question of whether the Kurdek study can be considered a replication of the Fincham et al. study. Specifically, the Kurdek study used a different time frame, which is probably too long for observing lagged effects, and a sample in which the magnitude of the effects might be expected to be somewhat attentuated. Rather than viewing the two studies in terms of nonreplicated findings, they might be viewed as confirming the need to pay closer attention to a number of factors that are likely to influence longitudinal findings. Prominent among these factors is the need to investigate shorter rather than longer lagged effects and to obtain and retain larger percentages of couples asked to participate in the research. Shorter lags may help address the problem of recruitment and attrition, but it is likely that researchers in the field will also need to begin considering the utility of new methods to correct for loss of data and so enhance our chances of obtaining replicable results across longitudinal samples.

References

Arbuckle, J. L. (1996). Full information estimation in the presence of incomplete data. In G. A. Marcoulides & R. A. Schumacker (Eds.), *Advanced structural equation modeling: Issues and techniques* (pp. 243–277). Hillsdale, NJ: Erlbaum.

Beck, A. T., Steer, R. A., & Garbin, M. G. (1988). Psychometric properties of the Beck Depression Inventory: Twenty five years of evaluation. *Clinical Psychology Review, 8,* 77–110.

Bryk, A., Raudenbush, S., & Congdon, R. (1996). *HLM: Hierarchical linear and nonlinear modeling with HLM/2L and HLM/3L programs.* Chicago: Scientific Software International.

Davey, A., Halverson, C. F., Jr., McCrae, R. R., Zonderman, A. B., & Costa, P. T., Jr., (1999). *Stability and change in depressive symptoms over time: Evidence from the Baltimore Longitudinal Study of Aging.* Manuscript in preparation.

Derogatis, L. (1983). *SCL-90–R: Administration, scoring, and procedures manual.* Towson, MD: Clinical Psychometric Research.

Fincham, F. D., Beach, S. R. H., Harold, G. T., & Osborne, L. N. (1997). Marital satisfaction and depression: Different causal relationships for men and women? *Psychological Science, 8,* 351–357.

Glenn, N. D. (1998). Problems and prospects in longitudinal research on marriage: A sociologist's perspective. In T. N. Bradbury (Ed.), *The Developmental course of marital dysfunction.* Cambridge, England: Cambridge University Press.

Joreskög, K. G., & Sörbom, D. A. (1999). LISREAL (Version 8.30) [computer software]. Chicago: Scientific Software International.

Karney, B. R., & Bradbury, T. N. (1995). The longitudinal course of marital quality and stability: A review of theory, method, and research. *Psychological Bulletin, 118,* 3–34.

Krokoff, L. J. (1990). The relationship of the husband's emotional involvement to participation in research on marital relationships. *Journal of Family Issues, 11,* 182–190.

Kurdek, L. A. (1998). The nature and predictors of the trajectory of change in marital quality over the first 4 years of marriage for first-married husbands and wives. *Journal of Family Psychology, 12,* 494–510.

Little, R. J. A., & Rubin, D. B. (1989). The analysis of social science data with missing values. *Sociological Methods and Research, 18,* 292–326.

Muthén, B. O., & Curran, P. J. (1997). General longitudinal modeling of individual differences in experimental designs: A latent variable framework for analysis and power estimation. *Psychological Methods, 2,* 371–402.

Raudenbush, S. (in press). Are hierarchical models commensurate with mean and covariance structure models? In A. G. Sayer & L. M. Collins (Eds.), *New methods for the analysis of change.* Washington, DC: American Psychological Association.

Schafer, J. L. (1997). *Analysis of incomplete multivariate data.* London: Chapman & Hall.

Willett, J. B., & Sayer, A. G. (1994). Using covariance structure analysis to detect correlates and predictors of individual change over time. *Psychological Bulletin, 116,* 363–381.

Willett, J. B., & Sayer, A. G. (1996). Cross-domain analyses of change over time: Combining growth modeling and covariance structure analysis. In G. A. Marcoulides & R. A. Schumacker (Eds.), *Advanced structural equation modeling: Issues and techniques* (pp. 125–157). Hillsdale, NJ: Erlbaum.

Received April 16, 1999
Revision received May 17, 1999
Accepted July 23, 1999 ■