FIGURE 5.6. Intimacy as a function of daily conflict: Raw data and fitted regression lines for participants at the 5th, 25th, 50th, 75th, and 95th slope percentiles for the low- and high-relationship quality groups. Pairs of dashed lines show 95% confidence intervals. (This figure has been duplicated as Figure 2 of the example write-up.)

The Write-Up

As in Chapter 4, we begin the write-up with a statement of the hypothesis to be tested, as it might appear at the end of an Introduction section of an APA-style empirical paper.

In sum, the key hypothesis to be tested is whether global relationship quality is a moderator of the impact of daily conflicts on intimacy.

METHOD

Sample

The sample consisted of 66 women in a cohabiting intimate relationship. Fifty-eight of these were in an opposite-sex relationship and eight were in a same-sex relationship. The women had a mean age of 26.2 years (SD = 3.1). The racial breakdown in order of size was white (n = 38, 58%), Asian (n = 12, 18%), Hispanic (n = 8, 12%), and black (n = 8, 12%).

Measures

Daily conflict. Conflict was measured each day by the following diary question: "Did you experience any tension or disagreement with your partner today?" A "No" answer was coded 0, and a "Yes" answer was coded 1. To facilitate interpretation of the intercept in the analyses, we grand-mean centered scores on this variable by subtracting the mean conflict across subjects and time points (M = 0.22) from each score. In addition, because conflict varied both between- and within-subjects and because we wished to focus on within-subjects variation, we created a between-subjects and within-subjects version of the variable, *between conflict* and *within conflict*, respectively.

Daily intimacy. The six-item Reis and Shaver Intimacy Scale was used (Reis & Shaver, 1988). Raw scores were rescaled to a 0–10 interval, such that 0 was the lowest possible score and 10 was the highest possible score. Summary statistics for intimacy over participants and time were M = 4.8, SD = 2.3, range = 0–10.

Relationship quality. Participants completed the single-item measure of how globally satisfied they felt about their relationship. The question was "Overall, how satisfied are you with your relationship with [intimate partner]?" The response options were (with n's in parentheses): 5 = best I could ever imagine a relationship being (n = 10), 4 = extremely satisfied (n = 18), 3 = moderately satisfied (n = 12), 2 = a bit dissatisfied (n = 19), 1 = very dissatisfied (n = 7). A dichotomous (0, 1) version used in the analyses was created by combining codes 3, 4, and 5 into a high-RQ group (n = 40, 61%) and codes 1 and 2 into a low-RQ group (n = 26, 39%).

Procedure

Participants were recruited using flyers on a university campus. They responded to the flyers by e-mail or phone. Following recruitment, they

completed an online background questionnaire and online nightly dia-
ries for 28 consecutive days. Research assistants verified via time-date
stamps that daily entries were indeed completed each evening.

RESULTS

Preliminary Analyses

The analysis dataset consisted of 66 (subjects) × 28 (days) = 1,848
observations. Inspection of scatterplots, subject by subject, indicated
that four of the subjects did not report any relationship conflicts on any
of the 28 diary days. These subjects, therefore, could not contribute to
the estimation of within-subject reactivity to conflicts nor to relation-
ship quality differences in this within-subject relation. The scatterplots
did not reveal any outliers in the dataset. Although there were no appar-
ent time trends in the data, we included in the model a centered version
of day (divided by 7 to put the units in weeks). We observed no missing
data. Finally, although there were too few same-sex couples to examine
them separately, when the analyses reported below were rerun with
women in opposite-sex couples only, the results were substantively
unchanged.

Main Analyses

We analyzed our data using a multilevel model that specified a within-
subject process of reactivity to daily conflicts that we predicted would
be stronger for those in low-quality as opposed to high-quality relation-
ships. The results are presented in Figure 1 and Table 1. Figure 1 shows
fitted regression lines for each subject in the low- and high-RQ groups
together with thick fitted lines for the average subject in each group.
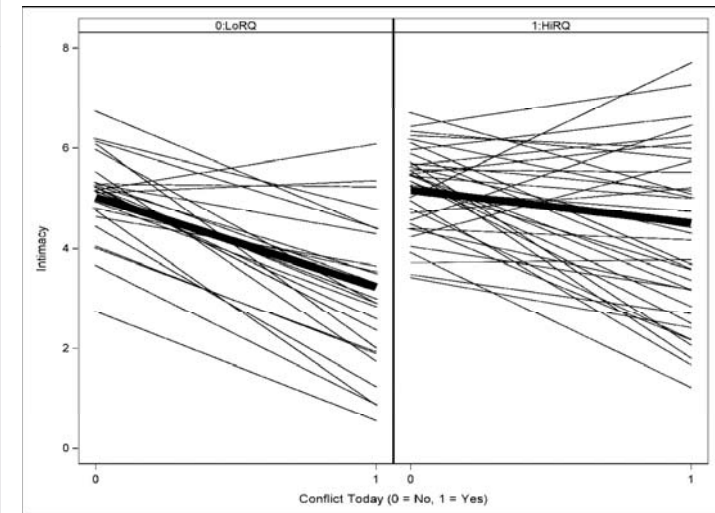Visual inspection of the thick lines suggests that our hypothesis was

Figure 1. Spaghetti plots of average (thick) and subject-specific (thin) regression lines for
female daily intimacy as a function of conflict for low- (left) and high- (right) relationship
quality (RQ) groups.

supported: The slope for the high-RQ group is less steep than it is for
the low-RQ group.

For a statistical test of the hypothesis, we turn to the upper panel
of Table 1, labeled *fixed effects*, and to the results for the RQ by *within-
conflict* interaction. Note that because the analysis model included time
in weeks, and because the within- and between-subjects forms of con-
flict were also included, results for the focal moderation hypothesis can-
not be artifacts of temporal changes or between-subjects differences in
average number of conflicts over the diary recording period. The esti-
mate of 1.02 units means that the women in the high-RQ group were
1 unit (on a 0–10 scale) less reactive to (within-person variation in)

Table 1. Parameter Estimates for Multilevel Model of Female Daily Intimacy as a Function of Daily Conflict and Relationship Quality

Fixed effects (intercept, slopes)	Estimate	(SE)	t^a	p^b	CI_{95}	
					Lower	Upper
Intercept	4.53	(0.22)	20.48	<.001	4.09	4.97
Time (per 7 days)	−0.03	(0.04)	−0.75	.45	−0.10	0.05
Within conflict	−2.01	(0.37)	−5.42	<.001	−2.75	−1.27
Relationship quality^c	0.65	(0.28)	2.29	.025	0.83	1.21
Within conflict by RQ	1.02	(0.49)	2.06	.044	0.03	2.00
Between conflict	−0.84	(1.10)	−0.76	.45	−3.05	1.49
Between conflict by RQ	2.53	(1.68)	1.50	.14	−0.83	5.89

Random effects ([co-]variances)	Estimate	(SE)	z	p^b	$CI_{95}{}^d$	
					Lower	Upper
Level 2 (between-person)						
Intercept	0.81	(0.17)	4.87	<.001	0.56	1.27
Within conflict	2.71	(0.67)	4.07	<.001	1.76	4.69
Intercept and within conflict	0.40	(0.24)	1.65	0.098	−0.07	0.88
Level 1 (within-person)						
Residual	3.58	(0.12)	29.32	<.001	3.35	3.83
Autocorrelation	−0.046	(0.025)	−1.82	0.068	−0.095	0.003

Note: $N = 66$ persons, 28 days, 1,848 observations.
^{a}We took a conservative approach to specifying degrees of freedom, such that these were based on the number of subjects ($N = 66$) rather than the total number of observations ($N = 1,848$).
^{b}All *p*-values are two-tailed except in the case of variances, where one-tailed *p*-values are used (because variances are constrained to be non-negative).
^{c}Relationship quality (RQ) is coded 0 for those in low-quality relationships and 1 for those in high-quality relationships.
^{d}Confidence intervals for variances were computed using the Satterthwaite method (see Littell et al., 2006).

daily conflicts than women in the low-RQ group, $t(62) = 2.06$, $p = .044$, $CI_{95} = 0.03$, 2.00.

   Important ancillary hypotheses are tests of the simple slopes for daily conflict for the low- and high-RQ groups. Given the coding of

relationship quality, and given the presence of the conflict-by-relationship quality interaction term in the model, for the low group this is the coefficient for *within conflict*. It is −2.01 units, $t(62) = −5.42$, $p < .001$, $CI_{95} = −2.75$, −1.27. For the high group the simple slope is half that amount: −2.01 + 1.04 = −0.97 units, $t(62) = 3.06$, $p = .003$, $CI_{95} = −1.6$, −0.3. (Note that this coefficient was obtained by rerunning the analysis model with the 0,1 coding of relationship quality reversed and reading off the coefficient for *within conflict*.)

   As we saw in Figure 1, there was substantial between-subjects variability in slopes and intercepts in the both the low- and high-RQ groups. The lower panel of Table 1 presents numerical estimates and statistical tests of this variability. These are reported as variances and covariances. Expressed as a standard deviation, the variation for the conflict slopes is $\sqrt{2.7} = 1.6$ units, which, assuming a normal distribution in the population, implies that approximately 95% of the population are within ±2 * 1.6 = ±3.2 units of the typical value for their relationship-quality subpopulation. Accordingly, the low-RQ group is predicted to have a 95% range of slopes of −2.0 ± 3.2 = −5.2 to 1.2. The high-RQ group is predicted to have an equivalent range from −1.0 ± 3.2 = −4.2 to 2.2.

   Missing from Figure 1 are the raw data that were used to obtain the fitted lines for each person. Figure 2 shows the raw data and the fitted lines for five subjects selected from each RQ group. These subjects represented the 5th, 25th, 50th, 75th, and 95th percentiles for the conflict reactivity slopes for their group. Also shown are the 95% confidence intervals for each slope. Not surprisingly, the slopes for the low-RQ group are more negative than those for the high-RQ group. However, it can also be seen that because those in the high-RQ group experience fewer conflicts, the uncertainty as to their true slope is greater than it is for those in the low-RQ group. (This difference in uncertainty is
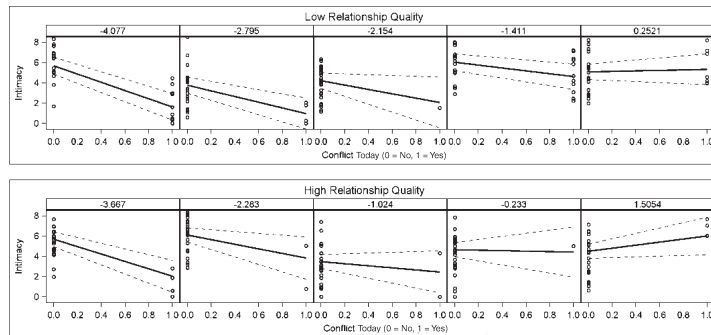
Figure 2. Intimacy as a function of daily conflict: Raw data and fitted regression lines for participants at the 5th, 25th, 50th, 75th, and 95th percentiles for the low- and high-relationship quality groups. Pairs of dashed lines show 95% confidence intervals.

balanced out by the greater number of people in the high-RQ group [40 vs. 26], such that the uncertainty about the typical person in the high group is very similar to that of the low group [*SE* of 0.33 for the high vs. 0.37 for the low].)

At the bottom of Table 1 is an estimate of the size of the residual variance at level 1 (also known as the level-1 random effect). This represents the deviations of the actual intimacy scores at level 1 from the predicted values obtained from the model. A common value is assumed for conflict and no-conflict days and for all subjects. Finally, as shown by the size and *p*-value of the coefficient, there is no evidence of auto-correlation in the level-1 (within-subject) residuals.

5.6   CHAPTER SUMMARY

Testing a causal hypothesis is difficult under any circumstances, but it is especially difficult when one wishes to observe a process unfold