

Linguistic Predictors of Adaptive Bereavement

James W. Pennebaker
Southern Methodist University

Tracy J. Mayne
University of California, San Francisco

Martha E. Francis
Southern Methodist University

The words people use in disclosing a trauma were hypothesized to predict improvements in mental and physical health in 2 studies. The first study reanalyzed data from 6 previous experiments in which language variables served as predictors of health. Results from 177 participants in previous writing studies showed that increased use of words associated with insightful and causal thinking was linked to improved physical but not mental health. Higher use of positive relative to negative emotion words was also associated with better health. An empirical measure that was derived from these data correlated with subsequent distress ratings. The second study tested these models on interview transcripts of 30 men who had lost their partners to AIDS. Cognitive change and empirical models predicted postbereavement distress at 1 year. Implications of using computer-based text analyses in the study of narratives are discussed.

Traumatic experiences, such as the death of a primary partner, can affect psychological functioning on many levels. It can influence how people think about themselves, their past and future, their relationships with others, and the meaning of life. Talking with others about traumatic events is an important way in which many people confront and resolve these issues (cf. Rimé, 1995). What people say about loss not only reflects their psychological and emotional state, but may also aid them in coping with the event.

The goal of this research was to examine the words people use to discuss the recent death of a loved one and to predict their subsequent physical and mental health. In the first of two studies, writing samples from participants in six existing writing studies were computer analyzed in an attempt to predict physical health and psychological well-being. From these data, several competing models were developed to specify exactly which aspects of language were associated with each of several outcomes. In the second study, the models developed in Study 1 were used to predict the psychological functioning of 30 men

whose primary partners had died on the basis of two interviews, performed 2 and 4 weeks after the deaths.

There is mounting evidence that the disclosure of emotional experiences can have positive effects. In a series of laboratory studies, individuals have been asked to write about traumatic or other emotional experiences for 3–5 consecutive days, for 15–30 min per day. Participants who have written about traumatic events have shown improved health compared to control participants who have written about superficial topics. Those who have written about traumatic events have been less likely to visit physicians because of illness in the subsequent 2–14 months (Pennebaker & Beall, 1986; Pennebaker, Kiecolt-Glaser, & Glaser, 1988; Greenberg & Stone, 1992; Pennebaker, Colder, & Sharp, 1990; Pennebaker & Francis, 1996; Krantz & Pennebaker, 1995; Greenberg, Wortman, & Stone, 1996; Richards, Pennebaker, & Beal, 1995). Experimental participants who have written about traumas have also demonstrated greater immunocompetence, as measured by heightened blastogenic responses to mitogens (Pennebaker, Kiecolt-Glaser, & Glaser, 1988), responses to Epstein-Barr antibodies (Esterling, Antoni, Fletcher, Margulies, & Schneiderman, 1994), and higher antibody responses to hepatitis B vaccinations (Petrie, Booth, Pennebaker, Davison, & Thomas, 1995). Writing about trauma has also been found to be associated with positive behavioral outcomes, including faster reemployment among laid-off engineers (Spera, Buhrfeind, & Pennebaker, 1994) and higher grades among college students (Pennebaker, Colder, & Sharp, 1990; Pennebaker & Francis, 1996; Krantz & Pennebaker, 1995).

On the basis of these and several other writing studies, we have developed a general theory of disclosure and language. The original theory posited that confronting upsetting topics reduced the constraints or inhibitions associated with not talking about the events. This was based on the assumption that inhibition causes or exacerbates stress-related disease processes (Pennebaker, 1989). In recent years, our working theory has evolved

James W. Pennebaker and Martha E. Francis, Department of Psychology, Southern Methodist University; Tracy J. Mayne, Center for AIDS Prevention Studies, University of California, San Francisco.

This is one of four studies analyzing from different theoretical perspectives the bereavement narratives of 30 men whose partners had died of AIDS. The data were from a longitudinal study: the University of California, San Francisco, Coping Project.

Preparation of this article was made possible by National Science Foundation Grant SBR-9411674 and National Institutes of Health Grants MH52391 and MH440045. We are indebted to Diane Berry, Susan Folkman, and Tom Trabasso for their comments on earlier versions of this article.

Correspondence concerning this article should be addressed to James W. Pennebaker, Department of Psychology, Southern Methodist University, Dallas, Texas 75275. Electronic mail may be sent via the Internet to pennebaker@mail.smu.edu.

by taking into account the central role of language. Specifically, we hypothesized that linguistically labeling an event and its emotions forces the experience to be structured. This structure promotes the assimilation and understanding of the event and reduces the associated emotional arousal. In other words, translating traumas and their accompanying images and emotions into language demands that all features of the experience be encoded and stored in a more organized, coherent, and simplified manner.

In previous studies, two components emerged that appear to be important in writing or talking about traumatic experiences. The first is the construction of an organized and coherent explanation or story surrounding the trauma. When individuals disclose a trauma to others, they must narrate an understandable account (Clark, 1993). The construction process emerges over time with repeated writing or telling, often moving from a vague and disorganized account of the trauma to a coherent and insightful explanation of events and feelings. The second component concerns the labeling of emotions. Individuals appear to benefit from describing their own feelings (Pennebaker, 1993; Pennebaker & Francis, 1996). Verbally labeling an emotional response may help to reduce the associated arousal (cf. Berkowitz & Troccoli, 1990; Keltner, Locke, & Audrain, 1993) and also aid the person in integrating emotional reactions into their general understanding of the traumatic event.

From our previous work, we hypothesized that cognitive change and emotional expression could be measured by counting relevant words in written (or by extension, spoken) text. Cognitive change was defined as the use of words in two general text dimensions: self-reflective thinking and causal thinking. The self-reflection category includes words such as *realize*, *understand*, *think*, and *consider*. The causal thinking category includes words such as *cause*, *effect*, *reason*, and *because*. We predicted that individuals whose use of these words in writing or speaking increased (i.e., they used a higher proportion of insight or causal words to total words on the last day relative to the first) would show health improvements. Indeed, two studies have found support for the cognitive change hypothesis (Pennebaker, 1993; Pennebaker & Francis, 1996).

We also hypothesized that writing is beneficial to the degree to which people are able to express their emotions in words. On the basis of one set of analyses, Pennebaker (1993) proposed a *differential emotion* model. This model assumed that individuals who use more negative emotion words and fewer positive emotion words would show health improvements. By using a more sophisticated analytic strategy (but on a more restricted sample), Pennebaker and Francis (1996) suggested that the use of both negative and positive emotion words will be associated with improved health outcomes. Indeed, this *summed emotion* model was based on the observation that positive emotion words were better predictors of improved health in the study than negative emotion words—a finding that contradicted the original differential emotion hypothesis.

We also posited that different models would vary in their ability to predict different outcome measures. An implicit assumption made in psychology is that physical and mental health outcomes are positively correlated. Recent research findings have begun to question this idea (cf. Shedler, Mayman, & Manis, 1993; Watson & Pennebaker, 1989). To the degree that self-

reported psychological distress, symptom reporting, illness, and behaviors reflect different processes, we should expect that different linguistic conventions may uniquely predict specific outcomes. Most of our earlier research has focused on physical rather than mental health measures. This study allowed us to apply each model to each of these domains of measurement.

The present experiment sought to test the cognitive change, differential emotion, and summed emotion hypotheses by using computerized text analyses with physical and mental health outcome measures. The first step (Study 1) involved analyzing all the writing samples of people in the experimental conditions (writing about trauma and emotionally upsetting experiences) of the writing studies that we have conducted in the last few years. We sought to determine how change in peoples' use of words over time was related to changes in physical and mental health. These analyses allowed us to gauge the power of the various hypotheses and, more important, to learn if more efficient models could be created. The second step (Study 2) was to apply the models developed in Study 1 to the interviews of the bereaved men collected as part of the University of California, San Francisco (UCSF), Coping Project, to test which model best predicted mental health 1 year after bereavement.

Study 1

Since 1986, Pennebaker and his colleagues have been involved in several writing studies that have examined the potential benefits of writing about personal and emotional topics. Outcomes have included measures of physical health, mental health, and selected behaviors. The data from six of the most recent studies have been used to devise and test the viability of several theoretical and statistical models of text analysis. As summarized below and in Table 1, the studies have varied considerably in sample characteristics, size, and experimental format.

Method

Samples

Immune-I (Pennebaker et al., 1988). Fifty college students were randomly assigned to write about either "the most traumatic experience of your entire life" in emotional ways (experimental condition), or, in the control condition, about relatively superficial and nonemotional topics (e.g., describe the room or plans for the day). Participants were asked to write for 4 consecutive days, 20 min per day. For the present studies, only data for the experimental participants were included. Analyses compared physician visits to the student health center from the 4 months before writing to the 6 weeks after. A single self-reported sadness question was also given at the 6 week follow-up: "Since the experiment, to what degree have you been sad or depressed?" Analyses controlled for preexperimental negative affectivity using the Negative Affectivity (NA) scale from the Positive and Negative Affect Schedule (PANAS; Watson, Clark, & Tellegen, 1988).

Immune-II (Petrie et al., 1995). Forty medical students attending Auckland University Medical School in New Zealand were randomly assigned to write about traumas or superficial topics for 4 consecutive days, 20 min per day. Analyses compared measures taken before and 4

Table 1
Summary of Six Writing Studies

Study	Exper. <i>N</i>	Mean age (years)	% Men	Description
Immune-I	25	19.8	28.0	SMU undergraduates
Immune-II	20	21.5	45.0	New Zealand medical students
College-I	37	17.9	48.6	SMU first year students
College-II	37	18.1	43.2	SMU first year students
Prison	38	36.9	100.0	Maximum security prisoners
Layoff	20	53.6	95.0	Unemployed engineers

Note. Exper. *N* refers to the sample size of the participants in the experimental condition of the respective study on which the analyses for the present study were based. SMU = Southern Methodist University.

months after writing in the trauma condition.¹ Measures included self-reported symptoms (from the Pennebaker Inventory of Limbic Languidness, or PILL; Pennebaker, 1982) and life satisfaction, which is inversely related to measures of NA or distress.

College-I (Pennebaker et al., 1990). One hundred and thirty first year college students were randomly assigned to write about their deepest thoughts and feelings about coming to college, or, in the control condition, to write about superficial topics. Participants were asked to write for 20 min on each of three consecutive days. The data of 37 students in the experimental condition with complete data sets were retained for the present analyses. Measures included change in health center visits for illness from before to the 4–6 months after writing, change in adjusted grade point average (GPA; adjusted for college entrance exam scores and hours attempted) from the semester of writing to the semester after, change in self-reported adjustment to college (College Adjustment Test, or CAT; Pennebaker et al., 1990), and self-reports of sadness adjusted for NA (as in the Immune-I study).

College-II (Pennebaker & Francis, 1996). Seventy-two first year college students were randomly assigned to write about their deepest thoughts and feelings about coming to college ($N = 35$) or about superficial topics ($N = 37$). Participants wrote for 3 days, 15 min per day. Measures included change in health center visits from 2 months before to 6 months after writing, changes in adjusted GPA from the semester of writing to the following semester, and changes in self-reports of college adjustment from the CAT.

Prison (Richards et al., 1995). Ninety-eight maximum security male inmates in a large state penitentiary were randomly assigned to one of three conditions. For the present analyses, the data of 38 men who were assigned to the trauma writing condition were retained. Men wrote about their deepest emotions and thoughts about the most traumatic experiences of their entire lives for 30 min on three occasions. Changes in infirmary visits and self-reports of anxiety (Cognitive and Somatic Anxiety Questionnaire) and symptoms (PILL) from the 6 months before to the 6 months after writing were analyzed.

Layoff (Spera et al., 1994). Sixty-three recently unemployed male professionals were randomly assigned to write about their thoughts and feelings about having been laid off or about superficial topics. Participants wrote for 5 consecutive days, 30 min per day. The present analyses included the data of those asked to write about being laid off ($N = 17$). Dependent measures included number of months to reemployment and change in self-reported anxiety from the weeks before writing to the 2–6 months after writing.

Procedure

All essays in the experimental conditions of each of the above studies were analyzed with the computerized text analysis program, Linguistic Inquiry and Word Count (LIWC; Francis & Pennebaker, 1993; Penne-

baker & Francis, 1996). The LIWC program individually searches text files (or, in the present study, files of each person's typed essays) and computes the percentage of words that were earlier judged to reflect negative emotion, positive emotion, causation, and insight or self-reflection. These and other LIWC dimensions of words are part of an extensive dictionary file that is composed of over 2,100 words or word stems. LIWC calculates the total number of words, sentences, percentages of unique words, and dictionary words. The sums of each of the scales are converted to a percentage of total words to correct for differences in text length between participants.

Each of the content categories captured by LIWC is composed of groups of category-relevant words initially generated by groups of judges, *Rogel's Thesaurus*, dictionaries, emotion and other types of questionnaires, as well as analyses of words used by previous samples of participants writing about emotional and control topics. After lists of words were compiled, at least three judges independently determined if each word should go into each category. A word was retained if two or more judges agreed on its inclusion. On the second wave of word evaluations, at least three new judges evaluated each of the words within a broader category (e.g., all negative words or all cognitive strategy words) that had previously been agreed on during the first wave. Reliabilities among judges were computed on each pass. Percentage agreement among the three judges was 93.1% for the first pass and 98.1% for the second pass.

Although the original LIWC program included 61 analysis dimensions, primary analyses focused only on the following categories: overall negative emotions, overall positive emotions, insight words, and causal words. Construction of the empirical model (see below) was based on the full range of LIWC dimensions. Information concerning these LIWC categories can be seen in Table 2.

LIWC measures. In each of the six preliminary studies, participants wrote on 3–5 occasions about emotional topics. Two sets of LIWC-based scores were computed for each participant. First, mean scores for each text category were computed over all writing days (e.g., mean percentage of negative emotion words, insight words, etc.). Second, change scores from the first to last day of writing were computed for each of the pertinent categories. Change scores were based on linear orthogonal polynomials. For example, in those studies in which people wrote three times, the first day's writing scores were subtracted from the last; in those with four writing days, the following algorithm was

¹ In some of the previous studies (e.g., Pennebaker, Colder, & Sharp, 1990), physician visit data were available for up to 9 months after the experiment. The strongest effects of writing, however, typically occur in the 2 to 4 months after writing. Because the UCSF Coping Project sought to predict mental health 1 year after the interviews, the dependent measures furthest in time after writing were used.

Table 2
LIWC Output Variable Information for Selected Variables Used in Studies 1 and 2

LIWC text dimensions	Example scale words	No. of words in scale	Judges' scale agreement (%)		Correlation between LIWC scale and judges' ratings
			Pass 1	Pass 2	
Negative emotionality	angry, sad, wrong	541	99	99	.68
Positive emotionality	happy, elegant, joyful	328	97	97	.64
Insight	realize, see, understand	116	87	93	.77
Causation	because, infer, thus	52	88	97	.41
Death ^a	coffin, dead, suicide	51	99	100	
Past tense verbs ^b	did, felt, rushed	83			.78

Note. Judges' scale agreement refers to the percentage of words that independent judges agreed were appropriate to each text category. Correlations between LIWC scale and judges' ratings were based on simple correlations between the relevant scale and a group of judges independently rating 105 trauma-relevant essays on the same general dimension (from Pennebaker & Francis, 1996). All correlations are significant ($p \leq .01$). LIWC = Linguistic Inquiry and Word Count.

^a Death words were not rated by judges.

^b Past tense verbs were not evaluated by judges in the construction of the LIWC.

computed: $(\text{Day4} \times 3) + (\text{Day3} \times 1) - (\text{Day2} \times 1) - (\text{Day1} \times 3)$. In sum, all participants had two sets of LIWC scores: mean and change.

Data from all six of the writing studies were used to create and test the theories previously mentioned (cognitive change, differential emotion, summed emotion) plus an empirically derived model described in the following section. Within each study, LIWC mean and change scores were standardized. In addition, all six studies required participants to complete a self-report measure of anxiety or distress prior to writing and again at follow-up. In all studies, the longest term follow-up distress measure that was available was subjected to a regression analysis controlling for prewriting distress, that is, analyses were performed on the residuals after regressing baseline anxiety-distress onto follow-up anxiety-distress. The residual distress measure was then standardized within each study. Comparable analyses were conducted for studies that incorporated before and after measures of objective health as measured by physician visits (four studies), GPA (two studies), symptom reports based on the PILL (two studies), and success in finding a job following a layoff (one study). Thus, the relevant postexperimental dependent measures were standardized and corrected for preexperimental levels in all studies.

All the standardized LIWC and residual outcome measures were incorporated into one data file with 177 participants. The number of participants differed as a function of study, as well as for attrition for studies with long-term follow-ups. Sample sizes for each of the outcome measures were as follows: distress measures ($N = 149$ for all studies), physician visits ($N = 132$ for 4 studies, Immune-I, College-I, College-II, and Prison), PILL ($N = 53$ for two studies, Immune-II and Prison), GPA ($N = 65$ for two studies, College-I and College-II), and job loss ($N = 16$ for one study, Layoff).

Hypotheses. The cognitive change model predicts that those who increase their use of insight or causal words from the first day of writing to the last day will show health improvements. This model was tested by using the standardized change scores from the orthogonal polynomials described earlier. A higher score indicates that participants used more insight and causal words on the last day of writing than they did on the first. The differential emotion model predicts that individuals who use more negative emotion words and fewer positive emotion words will show health improvements. As a test of the differential emotion hypothesis, the standardized positive emotion LIWC scores were subtracted from the negative emotion LIWC scores. The summed emotion model predicts that individuals who use both negative and positive emotion words will show improved health outcomes. As a test of the summed emotion

hypothesis, the standardized positive emotion LIWC scores were added to the negative emotion LIWC scores.

Results

The aim of Study 1 was to develop and test the cognitive change, differential emotion, summed emotion, and empirical models against the six writing studies conducted over the last few years on various outcome measures. As can be seen in Table 3, none of the predicted hypotheses were found to be related to the distress measures. However, interesting and reasonably consistent effects emerged for the other outcome measures. Particularly striking are the results based on the cognitive change hypothesis. In general, as participants' use of cognitive dimensions increased from the first days of writing to the last, their number of physician visits decreased, their rate of physical symptom reports dropped, college students' grades improved, and unemployed engineers found new jobs faster. It is interesting that neither the additive nor interactive models (wherein emotion variables were either added to or multiplied with the cognitive variables) increased the simple correlations for any factor.

Contrary to one of our earlier findings, the summary results were consistent with the differential emotion model in the direction opposite of what we predicted. That is, the use of more negative emotion words and fewer positive emotion words was linked with more negative outcomes in the months after writing. Expressing positive emotions in writing, then, was tied to better health.² Further, as depicted in Table 3, the summed emotion model was unrelated to all outcome measures.

² Although not a feature of this study, exploratory analyses included tests of curvilinear relationships between negative and positive emotion word use and the various outcome measures. Curvilinear emotion word variables were computed by using the absolute value of the standardized negative emotion LIWC and the positive emotion LIWC variables. Interestingly, the curvilinear negative emotion word use variable was positively related to illness change ($r = .27, p < .05$), whereas curvilinear positive words were not related ($r = -.14, ns$). In other words, individuals who used a moderate number of negative emotion words benefited more than did those who either had very high or very low negative emotion word usage.

Table 3
Tests of Emotion and Cognitive Change Hypotheses: Correlations Between LIWC Dimensions and Primary Outcome Measures From Six Previous Studies

Dependent measure	Cognitive change hyp.	Differential emotion hyp.	Summed emotion hyp.	Additive model	Interactive model	N
Distress	-.05	.01	.11	.04	-.05	149
Illness	-.16†	.17*	.03	-.09	-.04	132
Symptoms	-.31*	.28*	-.03	-.25†	-.06	53
GPA	.28*	.12	.07	.21†	-.01	65
New job	.67*	-.14	-.16	.52*	-.14	16

Note. The dependent measures of distress (based on self-reports), illness (physician visits), symptoms (self-reports from the Pennebaker Inventory of Limbic Languidness), and grade point average (GPA) were residual scores controlling for prewriting levels. The new job variable refers to number of months to reemployment after the writing manipulation. The cognitive change variables were based on the difference of the last days of writing minus the first days (see text for description); hence, a positive correlation would indicate relatively greater use of insight and causal words on the last day relative to the first day of writing. The differential emotion hypothesis was computed negative minus positive: Positive correlations indicate relatively greater use of negative emotion words relative to positive words. The additive model was based on the sum of both emotions and both cognitive change variables. The interactive model was the product of the summed emotions and the summed cognitive change variables. LIWC = Linguistic Inquiry and Word Count; hyp = hypothesis.

† $p < .10$. * $p < .05$.

The outcome measures in the bereavement project (Study 2) are essentially distress, or NA, measures. The inability to find linguistic correlates of self-reported distress within the theoretical models from which we have worked is both puzzling and distressing. Because the theoretically based cognitive change and emotion models were unrelated to the distress measures, we sought to construct a statistically based model to predict distress by using a regression solution. If we counted all the LIWC mean and change variables, over 100 variables could be forced into a regression equation to determine the best empirical model to predict distress. Such a strategy, of course, violates all assumptions of statistics as we know them. To construct a relatively simple predictive model, a stepwise regression procedure was computed on the standardized LIWC data for the 146 participants pooled across the six preliminary studies, with the distress outcome measures as the dependent variables. A small number of mean and change LIWC categories were not included because their meaning was quite different for the preliminary studies and the bereavement study (e.g., references to weight and illness categories).

The results of the regression analysis with adjusted distress as the outcome measure retained only four variables: use of past tense verbs ($\beta = -.23$), death-related words ($\beta = +.25$), change in unique words ($\beta = -.22$; unique words, also known as type-token ratios, refer to the percentage of different words used in a given essay), and change in use of positive emotion words ($\beta = +.15$). Thus, participants who reported greater distress at follow-up tended to (a) use fewer past tense verbs, (b) use more death-related words, (c) use fewer unique words over time, (d) use more positive emotion words over time. The combined four variables yielded a highly significant adjusted R^2 of .14, $F(4, 144) = 7.21$, $p < .001$. Not surprisingly, the empirically derived variable (based on the sum of the four LIWC scores multiplied by their beta weights) was not significantly correlated with changes in illness, $r(130) = .02$, symptom re-

porting, $r(51) = .08$, GPA, $r(63) = -.12$, or ability to find a new job, $r(14) = -.37$.

Study 2: UCSF Coping Project

The rationale for analyzing the essays in the first study was to apply the findings to a sample of men who had recently suffered the deaths of their lovers due to AIDS. Using the transcripts of two interviews collected within a month of their lover's death, we sought to predict the men's mental health 1 year later.

Method

Participants

This study examined data from the UCSF Coping Project, a longitudinal study of the effects of caregiving and bereavement on the mental and physical health of 86 HIV-positive and 167 HIV-negative partners of men with AIDS. Participants were recruited between April 1990 and June 1992 from the San Francisco Bay area. To be included in the cohort of bereaved caregivers selected for the Rashomon study, the participant had to be an HIV-negative caregiver whose partner died during the study and who had been interviewed at least once prior to his partner's death, twice within approximately 1 month following the partner's death, and 12 months following the partner's death as of the date of the commencement of this analysis, February 1993. Thirty participants who met these criteria were selected. The average length of time caregivers were involved with their partners was 6.27 years. The median educational level was college graduate, and 97% (29) were Caucasian.

Procedure

Participants in the UCSF Coping Project were interviewed bimonthly for 2 years. The bimonthly interviews included paper-and-pencil measures of positive and negative mood, stress, coping, social support, and physical health. Participants whose partners died during the study were interviewed at approximately 2 weeks and 4 weeks following the death.

These bereavement interviews began with paper-and-pencil assessments of psychological well-being. At the first of the two interviews, the interviewer asked the participant to describe what had happened at the time of his partner's death, what he was feeling, what he was thinking, what helped, and what made things more difficult. The second interview sometimes covered the same material as the first, with an update; at other times, the second interview covered only the interim between the first and second interviews. The open-ended sections of these interviews were tape-recorded and transcribed verbatim. Transcriptions were checked for accuracy. The 12-month follow-up interview was a structured interview that included paper-and-pencil measures of psychological well-being. Each participant had the same interviewer throughout the study.

Measures

The following four measures of psychological well-being were used:

1. Depressive mood was assessed bimonthly with the 20-item Centers for Epidemiological Studies Depression measure (CES-D; Radloff, 1977) of cognitive, affective, and vegetative symptoms of depression. Participants indicated on a 4-point scale that ranged from *rarely or none of the time* (0) to *most or all of the time* (3) how frequently each of 20 symptoms had occurred during the previous week (Cronbach's $\alpha = .89$).

2. Positive morale was assessed with a modified version of the positive morale subscale from Bradburn's (1969) Affect Balance Scale. The measure consists of eight items assessing positive mood states that are rated on a 4-point Likert scale that ranges from *none* (0) to *often* (3) for how often the person felt a particular way during the previous week (Cronbach's $\alpha = .90$). Examples of items include feeling "on top of the world," "optimistic," "pleased about having accomplished something," and "cheerful."

3. Positive states of mind were assessed with the Positive States of Mind Scale (Horowitz, Adler, & Kegeles, 1988), a six-item scale that is a measure of positively valued states such as pleasure, productivity, focused activity, and connection to others. Participants rated on a 4-point Likert scale that ranged from *unable to have it* (0) to *have it easily* (3) how much trouble, if any, the person had in having the state of mind during the previous week (Cronbach's $\alpha = .81$).

4. Impact of death was measured with the Impact of Event Scale (Horowitz, Wilner, & Alvarez, 1979), a 15-item measure that includes subscales of intrusive and avoidant thoughts. The measure was administered with respect to the partner's death. A summary score was used in the present analysis (Cronbach's $\alpha = .81$).

For the present study, we used two scores for each of the four measures: A score at the time of the partner's death that was an average of the two assessments made at 2 and 4 weeks following the partner's death and the score obtained 12 months following the partner's death. All 12-month measures were then adjusted for the 2-4 week levels by using regression procedures. For each of the adjusted measures, higher scores represented greater amounts of the psychological state being assessed.

Results

The goal of Study 2 was to apply the models developed and tested in Study 1 to the UCSF Coping Study data to predict psychological functioning 1 year after bereavement on the basis of interviews conducted 2 and 4 weeks after the death of a partner. Table 4 depicts the means of the primary LIWC variables of interest in the current project from Interview 1 (2 weeks postbereavement) and Interview 2 (4 weeks postbereavement). Virtually all changes were in face-valid directions. Over time, participants tended to increase in positive affect and insight. They also used fewer words associated with death. From the

Table 4
Mean Word Usage in Study-Related LIWC Variables During Interviews 1 and 2 ($N = 30$)

Variable	Interview 1		Interview 2		$t(29)$	p
	M	SD	M	SD		
Insight	2.82	0.47	3.22	0.77	2.37	.01
Cause	0.87	0.27	0.92	0.35	0.68	.50
Positive emotion	2.26	0.49	2.79	0.78	4.30	.01
Negative emotion	1.35	0.33	1.34	0.46	-0.13	.90
Death	0.52	0.28	0.37	0.27	-2.85	.01
Past tense	6.32	1.33	3.62	1.90	-6.27	.01
Unique words	18.90	6.52	24.43	7.96	5.60	.01

Note. Means and standard deviations reflect percentage of total words from the respective category. The t values are based on paired two-tailed t tests. LIWC = Linguistic Inquiry and Word Count.

first to the second interview, the interviewees also decreased in past tense verbs. Finally, the men tended to use more unique words in their second interview than they did in the first. It is interesting that these categories coincided with those found empirically to be related to improvements in psychological function in the previous writing studies (past tense, positive emotion, death, unique word).

On the basis of the preliminary studies, four different approaches or models were used to predict the outcomes for the 30 bereaved men in the UCSF sample. For all measures, the 12-month scores were adjusted for the levels at the time of bereavement by using regression solutions. The adjusted measures were then correlated with the various predictive models. The cognitive change model was based on the difference in insight and causal words from Interview 1 to Interview 2 (all LIWC means and changes were standardized). The differential emotion model was a single variable computed by subtracting mean positive emotion LIWC words from mean negative emotion LIWC words. The summed emotion variable was computed by summing the LIWC negative and positive emotion words. Finally, the empirical model variable was computed from the beta weights of the regression equation derived from the preliminary studies by using the following variables: (+) mean death words, (-) mean past-tense verbs, (+) change in positive emotion words, and (-) change in unique words. Note that higher numbers indicate greater predicted distress.

As depicted in Table 5, a number of intriguing results emerged. First, the cognitive and emotion model variables yielded unexpected effects. Contrary to our expectations, the cognitive change and the summed emotion models produced some significant correlations with the distress outcome measures. Particularly perplexing was that the cognitive change variable, which reflects greater insight and causal language use on the second interview, was negatively correlated with Positive States of Mind a year later, $r(28) = -.40, p = .03$. Increased cognitive change was marginally associated with a decrease in ruminations a year later as measured by the Impact of Events Scale, $r(28) = -.31, p = .10$, although the effect did not reach statistical significance. Neither the differential nor summed emotion model was significantly associated with any of the outcome measures.

Table 5
*Correlations of Clinical Ratings and Various Predictive Models With
 Primary Outcome Measures for UCSF Sample (N = 30)*

Variable	Cognitive change model	Differential emotion model	Summed emotion model	Empirical model
Self-reports				
Depression, CES-D	.08	.05	.20	.37*
Impact of Events	-.31†	.20	.17	-.11
Positive States of Mind	-.40*	.02	-.32†	-.41*
Bradburn Positive Morale	-.29	-.05	-.23	-.35†
Intercorrelations				
Cognitive change model	—			
Differential emotion model	.03	—		
Summed emotion model	.14	.00	—	
Empirical model	.22	-.12	.36*	—

Note. Higher numbers reflect increased use of causal and insight words on the second relative to the first interview (cognitive change model), relatively more negative emotion word use relative to positive (differential emotion model), and greater distress (empirical model). UCSF = University of California, San Francisco; CES-D = Center for Epidemiological Studies Depression measure.

† $p < .10$. * $p < .05$.

Most promising were the associations between the empirical model and the outcome measures associated with negative and positive affect. Recall that the empirical model was derived from six previous writing samples and was based on only four variables. As seen in Table 5, high predicted distress was correlated with the CES-D, $r(28) = .37$, $p < .05$, negatively correlated with Positive States of Mind, $r(28) = -.41$, $p < .05$, and marginally correlated with the Bradburn Positive Morale subscale, $r(28) = -.35$, $p < .06$. Separate simple correlations of the four variables comprising the empirical model indicated that the mean death word use category was most consistently related to the distress measures (r s with CES-D = .37, with Impact of Events = .38, with Positive States of Mind = -.38, and with Bradburn = -.38, all p s < .05).

General Discussion

The present studies have been important in examining word usage as it relates to physical and mental health across a variety of samples. In many respects, the analyses of the six preliminary studies yielded some of the most valuable data. Most striking was the consistent impact of participants' use of two word categories suggesting cognitive processing: insight and causation words. Specifically, participants' physical health and adaptive behaviors improved with increased use of these categories across the days of writing. Cognitive processing through writing, then, may be critical to coping with traumatic experiences. Indeed, these findings are remarkably consistent with those of Nolen-Hoeksema and her colleagues concerning the problems associated with rumination (Nolen-Hoeksema, McBride, & Larson, 1997; Lyubomirsky & Nolen-Hoeksema, 1995). The present analyses indicated that changes in thinking patterns—as opposed to static thinking patterns, which do not change over time—predict improved health. Static thinking patterns, as measured by the use of the same linguistic forms over time, could be viewed as rumination. In short, when individuals change in the ways they talk about a trauma, they are probably ruminating less and evidencing better health.

Contrary to our expectations, two of the emotion models that we have supported in previous work were not confirmed. The differential emotion hypothesis, which argues that expressing negative emotions is more beneficial than expressing positive emotions, did not receive any support in either study. Indeed, some results indicated that the expression of negative emotions over positive emotions may, in fact, be associated with higher rates of illness and symptom reports. The summed emotion model, which suggests the open expression of any emotions is beneficial, received only modest support in the bereavement project. All things considered, the linguistic expression of high rates of negative emotions, low rates of positive affect, or both, may not be conducive to physical or psychological health. Note that internal analyses suggested that maximal health benefits may be linked to moderate levels of negative emotion and high rates of positive emotion word usage.

Neither cognitive nor emotion word use was related to measures of self-reported distress for the six preliminary studies. This failure, which may reflect the domain to which cognitive or emotion models may predict, prompted the development of a nontheoretical empirical model with multiple regression. Overall, the regression solution pointed to four linguistic categories associated with distress in the preliminary studies: high use of words suggestive of death, low use of past tense verbs, an increasing use of the number of positive emotion words from the first to the last day of writing, and a decreasing use of unique words. When these variables were added together, they were unrelated to any of the health or behavioral measures in the preliminary studies. When applied to the bereavement project, the empirical model based on these variables was the only variable to reliably predict long-term adjustment.

It is interesting to note that these four LIWC dimensions were also the ones which changed significantly from Interview 1 to Interview 2 in face-valid directions for the UCSF Coping Project participants. Results showed that the interviewees increased in positive affect over time and in cognitive processing. Further, the two factors associated with worse outcomes (increasing

positive affect and an increase in unique words) were also normative for this population over this period. One could well imagine that a disproportionate increase in positive affect from 2 to 4 weeks following a partner's death may indicate other psychological processes (such as denial), whereas a slow recovery of positive affect may indicate a more normative process. By the same token, individuals who used the fewest past tense verbs in both postbereavement interviews evidenced the greatest distress 1 year later. These results hint that the failure to look back and confront a recent significant life event is a harbinger of poor coping. Further exploration of the empirical model is clearly warranted.

The goal of the UCSF Coping Project was to discover which factors were related to long-term psychological adjustment on the basis of only two interviews within a month after the death of the participants' partners. As suggested in the preliminary studies (and virtually all studies published by James W. Pennebaker dealing with disclosure), our models have been developed to predict physical rather than mental health. Nevertheless, the results from the analyses were promising. For example, the more that people increased in their use of insight and causal words, the less likely they were to ruminate about the death a year later. Interestingly, the increased use of these cognitive words was also correlated with lower levels of positive states of mind 1 year later.

The results of the present study highlight some important theoretical and practical problems for psychologists interested in adjustment. First, psychologists tend to assume that all outcome measures are equally valid and reflect the same underlying processes. Indeed, there is mounting evidence to contradict this assumption (e.g., Watson & Pennebaker, 1989). In looking at the results of the preliminary studies and the UCSF project, it is clear that factors that may correlate with objective health or other behaviors may be uncorrelated with self-reports of distress. Not only are the outcome measures unrelated, but different conceptual variables appear to selectively predict different and specific domains of mental and physical health.

A related issue concerns the links between the outcome measures and the context of the interview itself. Qualitative data depend on the richness of the interview, the questions asked, and the interviewer-interviewee relationship. We have made comparisons with earlier studies that were standard writing projects. That is, volunteers knew that they would be writing for a fixed number of days in relatively anonymous surroundings. It is quite promising that several of the linguistic factors that are used in solitary writing and that predict health outcomes apparently generalized to language used in social interactions. It is also encouraging that the empirical model, which was constructed to predict emotional distress in Study 1, significantly predicted emotional distress as measured by quite different measures in Study 2 (CES-D and Positive States of Mind). Although these variables were clearly related to the construct of emotional distress, the nature of the relationship merits further investigation. A related contextual issue is that the writing studies were intended to bring about a change in the physical and mental health of the participants. The interviews in the UCSF project were thought to be data-gathering techniques rather than interventions. Unknown at this point is the degree to which the interviews themselves were therapeutic and may have affected

change. An alternative explanation is that the interviews may simply reflect the cognitive changes that were occurring naturally.

Finally, the results of the present study (and those within the current series) point to the potential value of examining qualitative data with a computer-based text analysis procedure. That the information from two relatively brief interviews can predict self-reports and behaviors almost a year later is truly remarkable. Further, a computerized text analysis method, such as LIWC, is fast and efficient. Although a word counting approach cannot easily detect subtleties of language, such as irony or metaphor, it offers a method that can be easily adapted to a broad range of social-psychological and clinical paradigms.

References

- Berkowitz, L., & Troccoli, B. T. (1990). Feelings, direction of attention, and expressed evaluations of others. *Cognition and Emotion, 4*, 305-325.
- Bradburn, N. M. (1969). *The structure of psychological well-being*. Chicago: Aldine.
- Clark, L. F. (1993). Stress and the cognitive-conversational benefits of social interaction. *Journal of Social and Clinical Psychology, 12*, 25-55.
- Esterling, B. A., Antoni, M. H., Fletcher, M. A., Margulies, S., & Schneiderman, N. (1994). Emotional disclosure through writing or speaking modulates latent Epstein-Barr virus reactivation. *Journal of Consulting and Clinical Psychology, 62*, 130-140.
- Francis, M. E., & Pennebaker, J. W. (1993). *LIWC: Linguistic Inquiry and Word Count*. (Tech. Rep.). Dallas, Texas: Southern Methodist University.
- Greenberg, M. A., & Stone, A. A. (1992). Emotional disclosure about traumas and its relation to health: Effects of previous disclosure and trauma severity. *Journal of Personality and Social Psychology, 63*, 75-84.
- Greenberg, M. A., Wortman, C. B., & Stone, A. A. (1996). Emotional expression and physical health: Revising traumatic memories or fostering self-regulation? *Journal of Personality and Social Psychology, 71*, 588-602.
- Horowitz, M. J., Adler, N., & Kegeles, S. (1988). A scale for measuring the occurrence of positive states of mind: A preliminary report. *Psychosomatic Medicine, 50*, 477-483.
- Horowitz, M. J., Wilner, N., & Alvarez, W. (1979). Impact of Events Scale: A measure of subjective stress. *Psychosomatic Medicine, 41*, 209-218.
- Keltner, D., Locke, K. D., & Audrain, P. C. (1993). The influence of attributions on the relevance of negative feelings to personal satisfaction. *Personality and Social Psychology Bulletin, 19*, 21-29.
- Krantz, A., & Pennebaker, J. W. (1995). *Bodily versus written expression of traumatic experience*. Manuscript submitted for publication.
- Lyubomirsky, S., & Nolen-Hoeksema, S. (1995). Effects of self-focused rumination on negative thinking and interpersonal problem solving. *Journal of Personality and Social Psychology, 69*, 176-190.
- Nolen-Hoeksema, S., McBride, A., & Larson, J. (1997). Ruminative coping and depressed mood among bereaved partners: The Rashomon Project. *Journal of Personality and Social Psychology, 72*, 855-862.
- Pennebaker, J. W. (1982). *The psychology of physical symptoms*. New York: Springer-Verlag.
- Pennebaker, J. W. (1989). Confession, inhibition, and disease. In L. Berkowitz (Ed.), *Advances in experimental social psychology* (Vol. 22, pp. 211-244). New York: Academic Press.
- Pennebaker, J. W. (1993). Putting stress into words: Health, linguistic, and therapeutic implications. *Behaviour Research and Therapy, 31*, 539-548.

- Pennebaker, J. W., & Beall, S. K. (1986). Confronting a traumatic event: Toward an understanding of inhibition and disease. *Journal of Abnormal Psychology, 95*, 274-281.
- Pennebaker, J. W., Colder, M., & Sharp, L. K. (1990). Accelerating the coping process. *Journal of Personality and Social Psychology, 58*, 528-537.
- Pennebaker, J. W., & Francis, M. E. (1996). Cognitive, emotional, and language processes in disclosure: Physical health and adjustment. *Cognition and Emotion, 10*, 601-626.
- Pennebaker, J. W., Kiecolt-Glaser, J. K., & Glaser, R. (1988). Disclosure of traumas and immune function: Health implications for psychotherapy. *Journal of Consulting and Clinical Psychology, 56*, 239-245.
- Petrie, K. J., Booth, R. J., Pennebaker, J. W., Davison, K. P., & Thomas, M. G. (1995). Disclosure of trauma and immune response to a hepatitis B vaccination program. *Journal of Consulting and Clinical Psychology, 63*, 787-792.
- Radloff, L. S. (1977). The CES-D Scale: A self-report depression scale for research in the general population. *Journal of Applied Psychological Measurement, 1*, 385-401.
- Richards, J. M., Pennebaker, J. W., & Beal, W. E. (1995, May). *The effects of criminal offense and disclosure of trauma on anxiety and illness in prison inmates*. Paper presented at the annual meeting of the Midwest Psychological Association, Chicago.
- Rimé, B. (1995). Mental rumination, social sharing, and the recovery from emotional exposure. In J. W. Pennebaker (Ed.), *Emotion, disclosure, and health* (pp. 271-291). Washington, DC: American Psychological Association.
- Shedler, J., Mayman, M., & Manis, M. (1993). The illusion of mental health. *American Psychologist, 48*, 1117-1131.
- Spera, S., Buhrfeind, E., & Pennebaker, J. W. (1994). Expressive writing and job loss. *Academy of Management Journal, 37*, 722-733.
- Watson, D., Clark, L. A., & Tellegen, A. (1988). Development and validation of brief measures of positive and negative affect: The PANAS scales. *Journal of Personality and Social Psychology, 54*, 1063-1070.
- Watson, D., & Pennebaker, J. W. (1989). Health complaints, stress, and distress: Exploring the central role of negative affectivity. *Psychological Review, 96*, 234-254.

Received March 22, 1996

Revision received October 30, 1996

Accepted November 2, 1996 ■