Journal of Consulting and Clinical Psychology
1998, Vol. 66, No. 1, 174–184

0022-006X/98/$3.00

Written Emotional Expression: Effect Sizes, Outcome Types, and Moderating Variables

Joshua M. Smyth
State University of New York at Stony Brook

A research synthesis was conducted to examine the relationship between a written emotional expression task and subsequent health. This writing task was found to lead to significantly improved health outcomes in healthy participants. Health was enhanced in 4 outcome types—reported physical health, psychological well-being, physiological functioning, and general functioning—but health behaviors were not influenced. Writing also increased immediate (pre- to postwriting) distress, which was unrelated to health outcomes. The relation between written emotional expression and health was moderated by a number of variables, including the use of college students as participants, gender, duration of the manipulation, publication status of the study, and specific writing content instructions.

Emotional expression has long been given a central role in the study and practice of psychology. Both historically and recently, psychologists have cited the expression of emotions as vital for good mental and physical health, although the inhibition of emotion was considered deleterious (e.g., Breuer & Freud, 1895/1966; Grinker & Spiegel, 1945; Rachman, 1980; Scheff, 1979). More recently, there has been a growing body of literature suggesting that emotional expression has salutary health effects (e.g., Esterling, Antoni, Kumar, & Schneiderman, 1990; Fawzy et al., 1993; Mumford, Schlesinger, & Glass, 1983; Murray, Lamnin, & Carver, 1989; Pennebaker & O'Heeron, 1984; Spiegel, Bloom, Kraemer, & Gottheil, 1989), whereas emotional inhibition has detrimental health effects (e.g., Florin, Freudenberg, & Hollander, 1985; Goldstein, Edelberg, Meier, & Davis, 1988; Jamner, Schwartz, & Leigh, 1988; Jensen, 1987; Larson, 1990).

Expressive writing specifically has been used to promote good health in a number of controlled studies (Donnelly & Murray, 1991; L'Abate, 1992; L'Abate, Boyce, Fraizer, & Russ, 1992; L'Abate, Boyce, Russ, & Bird, in press; Lange, 1994; Murray & Segal, 1994). A brief written emotional expression task developed by Pennebaker (e.g., Pennebaker & Beall, 1986) calls for experimental participants to write an essay that expresses their feelings about a traumatic experience in their life (e.g., "write about your deepest thoughts and feelings about a trauma"), whereas control participants write about innocuous topics (e.g., "write about your plans for the day"). Studies using this paradigm have examined differences between control and experimental participants across a wide range of outcomes including health center visits, affect, immune measures, grade point average, and re-employment status. That a brief, written emotional expression intervention can impact overall health (including psychological well-being, physical health, and general functioning) over a number of months is certainly a controversial finding. Interest in the topic has resulted in numerous articles in prestigious journals (e.g., Esterling, Antoni, Fletcher, Marguiles, & Schneiderman, 1994; Greenberg & Stone, 1992; Pennebaker, Colder, & Sharp, 1990) and lay publications (e.g., Pennebaker, 1990).

This emerging area seems especially important in light of the fact that this finding has been applied on the basis of a small number of studies. For instance, entering students at a major university are routinely asked to perform this writing task (J. Pennebaker, personal communication, October 2, 1995). Any number of people may be "self-prescribing" the writing task on the basis of lay reports and publications. *American Health,* for instance, published articles titled "Writing your wrongs" (Pennebaker, 1991) and "Writing off the unemployment blues" (Willensky, 1993), both lauding the benefits of emotional writing. Furthermore, people may have an intense desire to express or discuss traumatic events, yet experience social constraints that force them not to talk about it (Lepore, Silver, Wortman, & Wayment, 1996; Pennebaker & Harber, 1993). Written expression may thus fill a very important niche—providing a mechanism of emotional expression in circumstances where interpersonal expression is not viable.

Originally, writing was conceptualized as allowing individuals to confront upsetting topics, reducing the constraints or inhibitions associated with not talking about the event. The work of inhibition (i.e., the cumulative physiological drain) was considered to cause and/or exacerbate stress-related disease processes (Pennebaker, 1989). Although early work focused on the central role of emotional expression (cf. Scheff, 1979), evidence suggests that emotional expression may be necessary, but is not sufficient, to produce positive change (Murray, Lamnin, & Carver, 1989; Pennebaker & Beall, 1986). More recently, Pennebaker and others have suggested that emotional expression facilitates cognitive processing of the traumatic memory, which

I thank Paul Wortman, Arthur Stone, and Camille Wortman for their many insightful comments on this article.

Correspondence concerning this article should be addressed to Joshua M. Smyth, Department of Psychiatry, South Campus, Putnam Hall No. 133, State University of New York, Stony Brook, New York, 11794-8790. Electronic mail may be sent to jsmyth@mail.psychiatry.sunysb.edu.

leads to affective and physiological change (Pennebaker, 1989, 1993). Specifically, written emotional expression leads to the transduction of the traumatic experience into a linguistic structure that promotes assimilation and understanding of the event, and reduces negative affect associated with thoughts of the event (Pennebaker, Mayne, & Francis, 1997).

Past reviews of studies using this written emotional expression task (e.g., Pennebaker, 1993) generally concluded that experimental participants are superior to controls on a variety of measures over the next few months. These reviews, however, have relied on the narrative method. Studies are grouped, the direction and significance of findings is noted, and overall conclusions are subjectively drawn from the number and consistency of the findings. The recent advent of statistical methods for the aggregation and examination of research literature, meta-analysis (Glass, 1976), or research synthesis (Cooper & Hedges, 1994) provides a more objective process for evaluating the size and significance of an effect. Research synthesis consists of statistical methods for generating an effect size for each observed between-group difference, classifies those effect sizes by domains (e.g., moderating variables), and quantitatively combines and compares effect sizes across studies by domain (Cooper & Hedges, 1994; Hedges & Olkin, 1985; Rosenthal, 1984). The use of research synthesis to evaluate new interventions, such as this writing task, has been advocated (see Yeaton, Langenbrunner, Smyth, & Wortman, 1995) and is especially warranted because there is now a sufficient number of experimental studies to conduct such a synthesis instead of using subjective judgment (Wortman, Smyth, Langenbrunner, & Yeaton, in press). The advantages of research synthesis over traditional narrative reviews in evaluating the efficacy of psychological, educational, and behavioral treatments have also been noted (Lipsey & Wilson, 1993).

Although prior reviews concluded that written emotional expression produces positive outcomes across a variety of measures, a number of questions remain unanswered. What is the effect size produced by the writing task? Although the diversity of outcomes studied suggests the possible breadth of the impact of the writing task, it is not clear if the effect size of the written emotional expression task is clinically relevant. That is, does this manipulation have the potential to meaningfully affect well-being, health, or general functioning? The role of moderating factors in the apparently salutary effects of written emotional expression has also not been examined.

The first goal of this article is to conduct a meta-analytic review of the written emotional expression literature to evaluate the overall significance (either beneficial, no effect, or harmful) and effect size of the brief writing task. Furthermore, the effect size will be examined across various outcome measures: psychological well-being, physical health, and more general functioning. The second goal of this article is to determine the moderating factors through which the effect of this writing task may be attenuated or enhanced. For example, the effect of written emotional expression may vary across sample characteristics (college students vs. adult community sample), outcome type (e.g., self-report vs. objective data), or by dose (how much writing was done). The specific potential moderating variables evaluated in this research synthesis are participant characteristics, dose, essay content instructions, outcome type, and publication status (each of these is discussed below). Should the writing task prove effective at producing positive change, promising research areas may be suggested by significant moderating variables. Such information can help determine the mechanism(s) of action and will be important in the development of emotional writing interventions to be tested in future research.

Method

Literature Search

Relevant articles were located through a computer search of *Psychological Literature, PsycINFO* (*Psychological Abstracts*), and *Citation Index*. Various permutations of keywords were used from the following: *emotion, expression, health, trauma, written,* and *writing*. All articles, as well as a recent review of this literature (Pennebaker, 1993), were used to perform a backward search of the references until no new articles were found. This generated 11 articles. First authors of the published studies were requested to supply information on any other published or unpublished articles on written emotional expression. Responses to these letters generated an additional 8 articles: 5 unpublished manuscripts, 2 dissertations, and 1 article in press. The total group of 19 articles was then examined to determine if they met the necessary inclusion criteria for this review.

Inclusion Criteria

As this review specifically focused on the impact of the written emotional expression manipulation developed by Pennebaker (Pennebaker & Beall, 1986), all studies had to contain a variant on his original task. Only randomized experiments were included to achieve a more stable estimate of overall effect size (Shadish & Ragsdale, 1996). Accordingly, each study had to meet the following criteria: (a) It had to contain an experimental manipulation of written emotional disclosure; (b) experimental participants had to write about traumatic topics, whereas control participants had to write about neutral topics; (c) the study had to contain some outcome measure of health, defined as mental, physical, or general functioning; and (d) the study had to contain statistical information necessary to calculate an effect size. Following these criterion, 13 studies were included for use in this review. Among those excluded were 5 articles typically cited in literature reviews and introduction sections of articles in this area. Esterling et al. (1990) was excluded because it had no control group. Four articles by Pennebaker were excluded: 2 because they did not involve written emotional expression (Pennebaker & Chew, 1985; Pennebaker, Hughes, & O'Heeron, 1987), and 2 because they lacked an experimental manipulation (Pennebaker & O'Heeron, 1984; Pennebaker & Susman, 1988). One unpublished article examining the effects of expressive movement and writing (Krantz & Pennebaker, 1995) was excluded because it lacked an adequate control group and confounded written and physical expression. One study included in the research synthesis (Greenberg, Wortman, & Stone, 1996) had one experimental group who wrote about imaginary traumas (i.e., traumas they did not actually experience). As this was not representative of the original written emotional expression task, this particular group was excluded from analyses. (It should be noted that all included studies used participants that were both physically and psychologically healthy.)

Coding Variables

The following variables were extracted from each study, on the basis of guidelines provided by Stock (1994): (a) Report information (authors, country, language, year of study, source of study), (b) setting

information (sampling scope and population type), (c) participant information (*Ns*, social status, age, education, gender, minority representation), (d) treatment information (number of writing sessions, length of each session, spacing of writing sessions, presence of a manipulation check, trauma past/current/mixed), (e) methodological information (attrition, outcomes), and (f) effect size information (statistic type, value, significance, direction). Health outcomes included those measured at least 1 month postwriting, and short term effects were measured as pre- to postwriting task differences. All coding was initially performed by me, following a precise codebook. This codebook was used by two additional raters to recode all studies, yielding a range of agreement from 82 to 100% across all variables, with a mean agreement rate of 93%. Any observed differences between the raters was discussed by all three raters until consensus was reached and that code was used.

Extrapolating Effect Sizes

As the studies included in this review reported a variety of inferential statistics, all results were transformed into Cohen's *d* as the measure of effect size. Cohen's *d* is a standardized mean difference estimate (Hedges, 1981). The majority of transformations and analyses were performed using the software DSTAT (Johnson, 1990). Transformations and analyses not performed by this software were performed following procedures described in Cooper and Hedges (1994). In cases where an effect was noted as nonsignificant, but no other inferential or descriptive information was provided, the effect size was assumed to be zero (Rosenthal, 1984). Cohen's *d*s were computed in two manners: (a) an overall effect size for each study and (b) one effect size for each specific outcome type examined (discussed below), averaged across all outcomes within outcome type and within study. Additionally, one *d* for all short-term distress measures was generated. Although allowing more than one effect size per study can result in nonindependence, it should be noted that the primary analysis used a single effect size from each study. Furthermore, analyses by content group were run independently, and one study never contributed more than one *d* to any one analysis (although studies with a wider range of outcome types did contribute a single *d* to a greater number of analyses). The corresponding correlation coefficient (*r*) for each *d* was also computed.

Evaluation of Effect Sizes

The magnitude and significance of the overall mean weighted effect size was computed for all outcomes (averaged within study) and all studies. This procedure was repeated again for each of the five outcome types, providing an estimate of the mean weighted effect size within each outcome type across all studies. All measures were scored so that when the experimental group was superior to the control group the effect size was in the positive direction, regardless of whether high or low scores on the measure were desirable. Each effect size was weighted inversely to its conditional variance (Shadish & Haddock, 1994). The homogeneity of the effect sizes was examined to determine if the *d*s varied more than would be expected by sampling error. If the homogeneity test is significant, it suggests that there is significant variance among effect sizes and moderator variables should be examined (Hedges & Olkin, 1985). Noncontinuous moderator variables were tested by dividing effect sizes into groups on the basis of study qualities and comparing the mean effect size between groups. This test results in Q_b, the between-group goodness of fit, with an approximate chi-square distribution with $p - 1$ degrees of freedom, where p is the number of groups (Hedges & Olkin, 1985). A brief discussion of each potential moderator variable included in analyses follows.

Moderating Variables

Participant characteristics. Many studies used college students as participants, a population that may not allow generalization of the observed effects. To evaluate the generalizability of the findings, studies using student participants are compared to studies with community or other nonstudent samples. Few studies have reported testing for gender and age differences, although the process or meaning of written disclosure (e.g., willingness to disclose) may differ across gender, age, or both. Consequently, both mean age and gender ratio are tested as moderating variables.

Dose. The studies using a written emotional expression task vary somewhat in the length, number, and duration of the writing sessions, ranging from a single 20-min session to one such session per week for 4 weeks. If the writing intervention functions on a dose–response curve, longer writing sessions or more sessions should increase the effect. The duration of the intervention (i.e., the time from the first to the last writing session) may also moderate the effect, although it is not clear a priori if extending the duration (also increasing the time between sessions) would intensify or dilute any effect.

Essay content characteristics. The relation of experimental participants to the trauma they are to write about has also been varied. Studies have called for participants to write about the most traumatic event in their life (e.g., Greenberg & Stone, 1992), ongoing traumatic events (e.g., Pennebaker, Colder, & Sharp, 1990), or either past or ongoing traumas (e.g., Francis & Pennebaker, 1992). If the effects of writing are based on disinhibition through the modulation of a physiological system (see Pennebaker, 1989), the effect may be influenced by the recency of the trauma (that is, older traumas having a more pronounced effect due to longer inhibition). Recency may also interact with the outcome type if certain systems are more reliant on the physiological drain produced by inhibition. For example, systems possibly more influenced by disinhibition (e.g., immune function) would show greater change if the trauma was in the past than if it was ongoing. Conversely, systems possibly less reliant on disinhibition and more dependent on cognitive factors (e.g., affect) may prove equally influenced by the writing intervention regardless of the trauma being past or ongoing. Recent work by Lutgendorf and colleagues (Lutgendorf, Antoni, Kumar, & Schneiderman, 1994) found that a verbal disclosure induction produced greater benefit for participants writing about older, more troublesome events. Although it is not clear if it is the length of time the trauma has existed or the severity of the trauma expressed that is related to positive outcomes (cf. Greenberg & Stone, 1992), there is accumulating evidence that the nature of the traumas written about may be an important moderating variable. The writing instruction given to participants (i.e., write about past trauma, current trauma, or either past or current trauma) is examined as a moderating variable.

Outcome type. The effect of the written emotional expression task may vary across the type of outcome. Outcome types are groups of outcomes that are conceptually similar, that is, are attempts to measure the same construct (see the Results section for outcome type information). The outcome types used are reported health, psychological well-being, physiological functioning, general functioning, and health behaviors. Certain types of outcomes may be more readily influenced if they are conceptually more closely related to the mechanism of action. For example, if written expression influences health by evoking a cognitive shift, one might expect greater change in cognitive skills than in the function of the immune system. Pennebaker and Francis (1996), for example, found that positive emotion word use in writing predicted health changes but not grade improvements, suggesting that health may be more closely tied to mood than is academic performance. Outcome types are examined to see if any act as moderating factors for the overall effect of the writing task. Additionally, as experimental participants typically experience distress immediately following the writing task (e.g.,

Table 1
Effect Size, Correlation, Significance, Sample Size, and Outcome Types for Each Study

Study	*d*	*r*	*p*	*n*	Outcome types
Esterling, Antoni, Fletcher, Marguiles, & Schneiderman (1994)	2.058	.7241	.0000	40	c, d
Francis & Pennebaker (1992)	.3442	.1729	.2737	41	b, c, d
Greenberg & Stone (1992)	.2628	.1323	.3546	50	a, b, e
Greenberg, Wortman, & Stone (1996)	.2154	.1084	.3865	65	a, b
Hughes (1993)	.8243	.3833	.0000	111	a, b, e
O'Heeron (1992)	.2169	.1093	.4182	56	b, d
Pennebaker & Beall (1986)	.4924	.2430	.0998	46	a, e
Pennebaker, Colder, & Sharp (1990)	.3002	.1493	.0887	130	a, b, d, e
Pennebaker & Francis (1996)	.5362	.2616	.0254	72	a, b, d
Pennebaker, Kiecolt-Glaser, & Glaser (1988)	.4213	.2093	.1406	50	a, b, c, e
Petrie, Booth, & Pennebaker (1995)	.6077	.2961	.0602	40	c
Richards, Pennebaker, & Beall (1995)	.4969	.2439	.0502	64	a, b
Spera, Buhrfeind, & Pennebaker (1994)	.2628	.1328	.4018	41	a, d, e
Overall	.4722	.2298	.0000	806	

Note. Outcome types are indicated as a = reported health, b = psychological well-being, c = physiological functioning, d = general functioning, and e = health behaviors.

increases in symptom reporting, negative mood, heart-rate, etc.), the effect of the writing intervention on short-term (pre- to postwriting) measures is examined separately. (The calculation of the overall effect size is confined to health outcomes measured at least 1 month after the final writing session.)

Publication type. One of the primary strengths of meta-analysis over traditional narrative reviews is the formal inclusion of unpublished work. This reduces potential bias in the review sample that can result from editorial preferences to particular, typically significant, findings. The inclusion of theses, dissertations, and unpublished manuscripts should allow adequate representation of studies that report no effect. Conversely, publication status may reflect underlying differences in methodological quality that produce differences in observed results. Effect sizes of the writing manipulation are compared between published and unpublished work. The correlation between proxies of study quality (randomization, attrition, manipulation checks, etc.) and publication type is also examined.

Results

Effect sizes for each study are shown in Table 1. The mean weighted effect size across all studies and outcomes was d = .47 (r = .23) and was significant at the $p < .0001$ level. The mean weighted effect size excluding the largest outlier was d = .41 (r = .20) and was still significant ($p < .0001$). The difference between these overall effect sizes was not significant, $Q_b = 0.30$, *ns*, so analyses included all effect sizes. As this analysis included only randomized experiments, one can infer the causal relationship that the written emotional expression task leads to positive long-term outcomes. Despite the inclusion of all available unpublished studies, there exists the possibility that unpublished studies with null findings may have been missed, positively biasing the overall effect size (the "file-drawer" problem). It is possible to calculate the number of such null finding studies that would have to exist for the observed effect size to become nonsignificant, which is referred to as the "fail-safe N" (see Cooper & Hedges, 1994). The fail-safe N for this analysis, including all studies used, is 199. As this number may itself be biased by the inclusion of outliers, the fail-safe N was computed excluding the largest outlier, yielding a fail-safe N of 117. Overall, it seems unlikely that unpublished studies (that were not included) would compromise the results.

There existed concern that the calculation of the overall effect size may have been biased due to the fact that Pennebaker was involved in 8 of the 13 studies used in this research synthesis. This may have resulted in artificially high effect sizes for those studies due to "experimenter effects" or low within-group variance (due to increased homogeneity in studies conducted by Pennebaker and his colleagues). Furthermore, if significant effect sizes were limited to work performed by Pennebaker and his colleagues, it would suggest that this research group can reliably elicit effects but would not allow generalizability of these results. Accordingly, a group contrast was performed between those studies in which Pennebaker was an author (n = 8) and those in which Pennebaker was not an author (n = 5). Studies in which Pennebaker was not listed as an author had slightly higher mean effect sizes than those studies in which Pennebaker was listed as an author (mean ds = .57 and .42, respectively), although this difference was not significant (Q_b = 0.90, p = .34). This result indicates that although Pennebaker was involved in the majority of studies included in this research synthesis, effects generated by other research groups are both reliable and not significantly different in magnitude.

Although the overall effect is compelling, it must be examined in light of the considerable variability in effect sizes across studies. The test for homogeneity of effect sizes was significant (Q_w = 22.75, $p < .03$), indicating significant within-group variance and suggesting that moderating variables should be examined. Variability of effect sizes was first corrected for sampling error according to Hunter, Schmidt, and Jackson (1982,

Table 2
Outcome Types

Outcome type formed	Specific outcomes	No. of studies assessing this outcome
Reported health	Health center visits	8
	Self-reported symptoms	3
	Upper respiratory illness	1
Psychological well-being	Positive affect	3
	Negative affect	3
	Happiness	2
	Anxiety	1
	Sadness	1
	Intrusions	1
	Adjustment	1
	General temperament	1
	Adjustment to college	2
	Adjustment to high school	1
Physiological functioning	Phytohemagglutinin	1
	Concanavalin A	1
	T-helper lymphocytes	1
	T-cytotoxic/suppressor lymphocytes	1
	Natural killer cells	1
	Epstein-Barr antibodies	1
	Hepatitis B antibodies	1
	Blood pressure	1
	Heart rate	1
	Triglycerides	1
	Cholesterol	1
	High density lipids	1
	Low density lipids	1
	Uric acid	1
	Albumin	1
	Globulin	1
	Liver function (SGOT, SGPT)	1
General functioning	Reemployment	1
	Grade point average	3
	Absenteeism	2
	Cognitive functioning (thought generation, reaction time)	1
	School behavior	1
Health behaviors[a]	Alcohol use	
	Drug use (including cigarettes, caffeine)	
	Exercise	
	Sleeping habits	
	Eating habits	

[a] The precise health behaviors used in each study were often not provided, so frequencies assessed are not available.

pp. 100–108), with $\sigma_\delta = .313$, supporting an examination of moderators. First, outcome types were examined as moderating variables. Outcomes were grouped by type according to conceptual similarity and groupings used in the articles reviewed (i.e., domains of dependent variables used in the study). Five types were formed: reported health, psychological well-being, physiological functioning, general functioning, and health behaviors. The specific outcomes used to form each outcome type can be seen in Table 2.

Mean weighted effect sizes, correlation coefficients, significance, and effect size variability corrected for sampling error across all outcomes and for each outcome type can be seen in Table 3. The mean effect size associated with each of the outcome types differed significantly from zero, with the exception of health behaviors. The written emotional expression task thus lead to improved reported health, psychological well-being, physiological functioning, and general functioning. Contrasts among group means were performed according to Hedges and Olkin (1985) and the resulting Q_bs were evaluated against a chi-square distribution. The mean effect size for health behaviors was significantly lower than the overall mean effect size, but no other group d was significantly different from the overall d ($Q_b = 8.8$, $p < .005$; all other Q_bs < 2.5, ns). The mean effect size for health behaviors was also significantly lower than each of the other outcome types, all Q_bs > 4.7, ps $< .05$. The effect sizes for psychological well-being and physiological functioning outcomes did not differ from one another, $Q_b = 0.57$, ns, but were significantly higher than the general functioning outcome type, $Q_b = 4.74$, ps $< .01$. As significant within-group effect size variation existed in the psychological well-being and

physiological functioning outcome types (see Table 3), they were examined along with the overall effect size when other moderating variables were tested.

As most studies using written emotional expression reported experimental participants experiencing greater distress during writing than control participants, a mean d was calculated for short-term distress. The written emotional expression task produced a significant rise in experimental participants' pre- to postwriting distress ($d = .84$, $r = .39$, $p < .0001$). This short-term distress d was significantly higher than all health outcome ds, all Q_bs > 3.84, ps $< .05$. Short-term distress d was not related to any of the health outcome ds (all ps $> .40$).

The participant characteristics of student versus nonstudent, age, and gender ratio were tested as moderator variables for overall, well-being, and physiological ds. Students had slightly higher mean ds than nonstudents overall (.49 vs. .39), $Q_b = 0.32$, ns, and within the physiological functioning outcome type (.78 vs. .37), $Q_b = 1.26$, ns. Students had significantly higher d than nonstudents within the psychological well-being outcome type (.76 vs. .34), $Q_b = 3.92$, $p < .05$. Age was not associated with d in any of the three outcome types (all ps $> .10$). Gender ratio (scored as % male) was significantly related to overall d ($\beta = .80$, $p < .05$), but unrelated to psychological well-being or physiological functioning effect sizes.

Three measures of "dose" were examined: number of writing sessions (ranging from 1 to 5), length of each writing session (from 15 to 30 min), and the time period over which the writing sessions were spaced (from 1 to 28 days). Number of writing sessions and length of sessions were unrelated to all ds (all ps $> .10$). The time period of writing was associated with overall d ($\beta = .76$, $p < .02$) such that studies with writing sessions spaced out over a longer period of time had higher mean overall ds, but spacing was not related to psychological well-being or physiological functioning effect sizes.

Whether participants were instructed to write about past, current, or either past or current traumas was not related to the overall effect size (.39 vs. .50 vs. .54, all contrast ps $> .70$). Participants instructed to write about current traumas had significantly higher mean psychological well-being ds than participants instructed to write about either a past or current trauma (.99 vs. .18), $\chi^2_{\text{contrast}} = 14.28$, $p < .001$, whereas participants writing only about past traumas fell between ($M = .56$, contrast ps $> .50$). Although no studies assessing physiological functioning outcomes instructed participants to write about only current traumas, participants asked to write about either past or current traumas had significantly higher mean ds than participants asked to write about only past traumas (1.04 vs. 0.41), $\chi^2_{\text{contrast}} = 3.86$, $p < .05$.

The publication status of studies (published vs. unpublished) was unrelated to overall d (0.48 vs. 0.47), $Q_b = 0.002$, ns, although unpublished studies were associated with higher psychological well-being ds (1.04 vs. 0.25), $Q_b = 16.91$, $p < .0001$. All studies assessing physiological functioning outcomes were published, so no comparisons could be made within this outcome type. Three proxies of study quality (randomization, attrition, and treatment manipulation checks) were not correlated with publication status (all ps $> .10$). Proxies of study quality also did not change over time, suggesting that there was not a general improvement of research methodology over time.

There exists the possibility that a third variable may explain these moderator relationships. The correlation matrix between the moderating variables, as well as their correlation with outcome variables, was examined. In summary, there appears to be little relationship among the moderating variables. Among all of the moderating variables (gender, student status, age, number of sessions, length of sessions, length of time over which the writing sessions were spaced, publication status, and writing content instructions), there were only two significant correlations. Publication status was related to the number of writing sessions ($r = .3$, $p < .01$; studies using more writing sessions were more likely to have been published [to date]), and student status was inversely related to age ($r = -.78$, $p < .0001$). In both cases, the second variable was unrelated to the outcome variable moderated by the first, and therefore cannot explain the moderating relationship. That is, neither number of writing sessions nor age were related to well-being effect sizes (both ps $> .22$).

In summary, there were six moderating variables across the three outcome types found to explain significant within-group variance in effect size. Overall effect sizes were moderated by two variables: Higher percentages of males in a study were related to higher mean effect sizes, as was longer periods over which writing sessions were spaced. Psychological well-being effect sizes were moderated by three variables, each increasing mean effect size: the use of student participants, instructions to write about current traumas (as opposed to past or current trauma), and unpublished studies. Lastly, physiological functioning effect sizes were higher in studies that instructed participants to write about past or current traumas (as opposed to past trauma only).

Discussion

The first goal of this research synthesis was to establish an overall effect size and significance level for the writing task. Results demonstrate that written emotional expression produces significant health benefits in healthy participants. The binomial effect size display (BESD) is a method of showing the practical importance of an effect size (Rosenthal & Rubin, 1982), and is presented as the difference in outcome rates between experimental and control groups. The effect size of $d = .47$ represents a 23% improvement in the experimental group over the control group. For example, illness rates decreasing from 61% in the control group to 38% in the experimental group. This effect size is similar to or larger than those produced by other psychological, behavioral, or educational treatments (Barnes, 1986; Lipsey & Wilson, 1993; Meyer & Mark, 1995; Smith & Glass, 1977; Wells-Parker, Bangert-Drowns, McMillen, & Williams, 1995). Although it is not possible to strictly compare effect sizes between studies when the outcome measures are dissimilar, these findings suggest that the effect of the writing task is similar to that found in other quantitative analyses of psychological interventions.

The question thus becomes how does writing about traumas produce these improvements? Traumatic stress research has

Table 3
Summary of Effect Sizes

Outcome type	k	d	Mean r	Homogeneity within (Q_w)	σ_δ
Overall	13	0.472***	.230	22.751*	.313
Reported health	9	0.421***	.206	3.296	.000
Psychological well-being	9	0.661***	.314	41.803***	.642
Physiological functioning	4	0.681***	.322	15.600**	.649
General functioning	5	0.331***	.163	3.101	.000
Health behaviors	6	0.029	.014	0.226	.000

Note. k = number of studies; d = mean weighted effect size; r = correlation corresponding to mean weighted effect size; σ_δ = effect size variability corrected for sampling error.
* $p < .05$. ** $p < .001$. *** $p < .0001$.

noted the distinction between memories for ordinary and traumatic events; traumatic memories seem immutable, and are more emotional and perceptual in nature (Terr, 1993; van der Kolk, Blitz, Burr, & Hartmann, 1984). Traumatic memories are encoded differently, and are not integrated into a personal narrative (Christianson, 1992; van der Kolk, 1994). Thus, the memory is stored as sensory perceptions, obsessional ruminations, or behavioral reenactments (Janet, 1909; van der Kolk & van der Hart, 1989, 1991). It is the persistence of intrusive and distressing symptoms, avoidance, and hyperarousal that results in observed psychological and biological dysfunction (Creamer, Burgess, & Pattison, 1992; McFarlane, 1988, 1992). One goal in treating traumatic memories is thus to facilitate the processing of traumatic memory (Foa, Steketee, & Rothbaum, 1989; Foa, Rothbaum, & Molnar, 1995). Foa and Riggs (1993) noted that traumatic memories are particularly disorganized, and treatments aimed at organizing memory should thus be more effective. This is supported by work in both clinical and healthy populations. DiSavino and colleagues (DiSavino et al., 1993) analyzed victims' trauma-related narratives during exposure and found that evidence of decreasing disorganization over time was associated with improvement. Similarly, Pennebaker (Pennebaker, 1993; Pennebaker, Mayne, & Francis, 1997) found that a narrative becoming more focused and coherent over writing sessions was associated with increased improvement. Writing about the traumatic event may force the transduction of the memories from sensory–affective components into an organized, linguistic format (Pennebaker, Mayne, & Francis, 1997), facilitating processes central to the treatment of traumatic memory. Namely, the deconditioning of traumatic memories and affectual–physiological responses, and the restructuring of dissociated traumatic memories into a personal, integrated narrative (Foa & Kozak, 1986; Herman, 1992; van der Hart, Steele, Boon, & Brown, 1993).

The writing task produced superior health outcomes in several outcome types, each measured at least 1 month postwriting: reported health, psychological well-being, physiological functioning, and general functioning. As each outcome type was improved by the writing task, the overall effect is not solely dependent on any one outcome type. That is, if one subgroup of outcomes (e.g., health center visits) was responsible for the entire effect, other outcome types would not have shown improvement. Effect sizes did, however, differ across outcome types. Psychological well-being and physiological functioning outcomes had higher effect sizes than reported health or general functioning outcomes, and the effect size for reported health outcomes was higher than general functioning outcomes.

Changes in psychological well-being may result from cognitive shifts about the trauma following writing. Tests of cognitive shifts using traditional information processing models (e.g., assessments of reaction time), however, have found no support for this hypothesis (Pennebaker & Francis, 1996). Alternatively, the increased usage of insight words (e.g., *understand, realize*) is associated with more improvement (Pennebaker, 1993). The strong effect on physiological measures provides support for the biological impact of writing. Written expression may free physiological resources previously used for inhibition (Pennebaker, 1989, 1993). Alternatively, memories for trauma-related subjects can result in alterations in a variety of psychophysiological systems, including: autonomic, neurohormonal, neuroanatomical, and immunogical changes (see van der Kolk, 1994). Assimilation of the traumatic memory may lead to reductions in the intrusions and hyperreactivity associated with traumatic memories, ultimately attenuating physiological responses. Unfortunately, there is currently no information on the relation between ongoing intrusive thoughts, assimilation, and physiological functioning.

The impact of writing on reported health outcomes may be lower than on physiological functioning because overall health is only partially mediated by physiological competence. For example, decrements in immune function may negatively impact health to a smaller degree as other factors (diet, exercise, etc.) also influence resistance to disease. Similarly, the impact of writing on general functioning may be lower yet because it is in turn mediated by changes in well-being, reported health, and physiological function. For example, re-employment may be more likely for individuals with improved well-being (who may be more pleasant), whereas grade point average will be higher for those who are healthier (who don't miss classes). Pennebaker, Mayne, and Francis (1997) found that outcomes included in the general functioning category (grade point average and re-employment) were associated with physical health outcomes. It should also be noted that the relationship among outcome types is likely quite dynamic.

The effect size for health behaviors was negligible and nonsignificant. Although consistent with earlier research (Pennebaker, 1993), this is surprising in light of changes in other domains. Successful change of health behaviors requires a complex set of conditions, including both the intent to change and the behavioral execution of these intentions (Miller, Shoda, & Hurley, 1996). Health behaviors would thus seem to be more influenced by commitment than by emotional factors, although the importance of Person × Situation interactions is noted (Miller et al., 1996). In fact, the negative affect produced by writing ("hot" representations; Miller et al., 1996) may specifically undermine successful health behavior change.

Short-term distress was also increased by the writing task. Short-term distress has been thought to be related to long-term improvement (Pennebaker, 1993). Average short-term distress was unrelated to all long-term outcomes examined. Thus, although all studies report mean increases in distress, experiencing relatively more short-term distress does not appear to lead to greater benefit. This supports the view that the trauma-relevant fear network must be activated for improvement to be made (e.g., Foa & Kozak, 1986; Foa, Riggs, Massie, & Yarczower, 1993). It suggests, however, that this may be a boolean process—short-term distress may be required for cognitive change, but the amount of short-term distress is not related to improvement.

Considering moderator variables, student participants were found to have significantly higher effects for psychological well-being outcomes than nonstudents (although overall effect size did not differ). Most student participants were first year or transfer students who were writing about the stress of college. It is possible that much of this stress involved issues of insecurity and self-esteem common to college students. The writing task may more directly impact students' well-being because their writing topics are ongoing hassles that have immediate affectual consequences. Examination of students' essays, however, reveals they were not trivial; topics included isolation and loneliness (54% of participants), loss of family (51%), and even thoughts of suicide (11%; from Pennebaker, Colder, & Sharp, 1990). As nonstudent participants were older on average (48.5 vs. 18.8 years), it is plausible they had more rigidly defined views of the self, making it more difficult for writing to produce change (cf. Epstein, 1991; Harber & Pennebaker, 1992; Horowitz, 1986). That age was unrelated to well-being outcomes, however, lessens the plausibility of this explanation.

The proportion of male participants was positively related to the overall effect size ($\beta = .80$), suggesting that writing may be more effective for males. As traditional sex roles make it less likely for men to disclose a trauma or express emotion than women (e.g., Ptacek, Smith, & Zanas, 1992), they may experience greater benefit due to lower prewriting levels of emotional expression. Males also tend to use more problem-focused coping (Ptacek et al., 1992), and may focus more on the trauma when writing—a difference that may facilitate the beneficial effects of expression (Pennebaker, 1993; Solomon, Avitzur, & Mikulincer, 1990).

One measure of dose, the amount of time over which the writing intervention was spaced, was positively related to the overall effect size. This unexpected result implies that lengthening the time course of the writing task would increase its effect. Number and length of writing sessions were unrelated to improvement. The salutory process (e.g., the integration of negative information) may progress over a period of time, increasing the benefit to the writer (Horowitz, 1986; Suls & Fletcher, 1985). Similarly, prolonged exposure strategies are thought to provide greater opportunity for improvement (Foa & Riggs, 1993).

Unpublished studies were associated with higher well-being effect sizes. As available proxies of study quality were not correlated with publication status, this finding is not likely a result of obvious methodological differences. Because the typical assumption of publication bias assumes that published studies will have higher effect sizes (Smith, Glass, & Miller, 1980), this finding is unexpected, although it should be noted that effect sizes were higher for unpublished studies only within one specific outcome type.

The instructional set given to participants regarding the trauma they were supposed to write about (past trauma, current trauma, or either) was also related to effect size. Instructional set was unrelated to overall effect size, but participants writing about only current traumas had well-being outcomes superior to those of participants instructed to write about any trauma (either past or current). Addressing ongoing traumas more intimately linked to daily life may produce greater well-being change than addressing past traumas that may be less salient to daily experience (similar to students writing about ongoing trauma). Participants assigned to write about any trauma (past or current) had physiological outcomes superior to those of participants assigned to write about only past traumas. This seems contrary to inhibition theory, where past traumas should have the greatest physiological load and produce more benefit when disinhibited (cf. Lutgendorf et al., 1994).

Although it is possible to conduct a research synthesis with 2 studies, the fewer the number of studies, the less stable the results (Rosenthal, 1995). It seems valid to perform this analysis with 13 studies, but analyses using subgroups of studies should be viewed more cautiously. Although the writing task produces health benefits in healthy participants, there may be negative interactions with other treatments or a subset of participants. Unfortunately, research synthesis is limited to the data collected within the studies, and examination of this issue is not currently possible. Use of the writing task may be limited by hesitation to foster distress in the absence of support resources (e.g., a therapist). Exposure therapy, though acknowledged as effective, can have negative effects on its participants (e.g., Pitman et al., 1991), and writing produces exposure without a therapist present to control the degree of exposure. Conversely, in all of the studies examined in this analysis, few participants reported difficulty in dealing with the negative emotions evoked by writing.

The hypothesis that the writing task facilitates cognitive processing and assimilation of traumatic memories must be tested by assessing ongoing cognitive, affective, and physiological changes. This will allow evaluation of whether cognitive changes (e.g., assimilation) drive affectual and physiological changes. Evaluation of short-term distress should occur over time, as well as include measures of participants' reluctance to disclose

distress. The effects of moderating variables revealed by this synthesis should also be examined. For example, manipulating the spacing of writing sessions may suggest an optimal spacing, perhaps related to trauma severity (cf. van der Hart et al., 1993). More care should also be taken in assessing and evaluating participant characteristics; that students react differently than nonstudents suggests that processes underlying change may differ between various populations. Finally, evaluation of the writing task as a potential intervention needs to examine a number of issues: whether specific types of trauma are related to outcomes, the role of writing parameters (e.g., the use of insight words), and the effect of the writing task in clinical samples of both psychological (e.g., posttraumatic stress disorder) and somatic (e.g., chronic disease) nature.

References

References marked with an asterisk indicate studies included in the meta-analysis.

Barnes, D. (1986). Promising results halt trial of anti-AIDS drug. *Science, 234,* 15–16.

Breuer, J., & Freud, S. (1966). *Studies on hysteria* New York: Avon. (Original work published 1895)

Christianson, S. (1992). *Handbook of emotion and memory.* Hillsdale, NJ: Erlbaum.

Cooper, H., & Hedges, L. (1994). *The handbook of research synthesis.* New York: Sage.

Creamer, M., Burgess, P., & Pattison, P. (1992). Reaction to trauma: A cognitive processing model. *Journal of Abnormal Psychology, 101,* 452–459.

DiSavino, P., Turk, E., Massie, E., Riggs, D., Penkower, D., Molnar, C., & Foa, E. (1993, November). *The content of traumatic memories: Evaluating treatment efficacy by analysis of verbatim descriptions of the rape scene.* Paper presented at the 27th Annual Meeting of the Association for the Advancement of Behavior Therapy, Atlanta, GA.

Donnelly, D. A., & Murray, E. J. (1991). Cognitive and emotional changes in written essays and therapy interviews. *Journal of Social and Clinical Psychology, 10,* 334–350.

Epstein, S. (1991). The self-concept, the traumatic neurosis, and the structure of personality. In D. Ozer, J. M. Healy, Jr., & J. Stewart (Eds.), *Perspectives in personality* (Vol. 3, pp. 63–98). London: Jessica Kingsley.

*Esterling, B., Antoni, M., Fletcher, M., Marguiles, S., & Schneiderman, N. (1994). Emotional disclosure through writing or speaking modulates Epstein-Barr virus antibody titers. *Journal of Consulting and Clinical Psychology, 10,* 334–350.

Esterling, B. A., Antoni, M. H., Kumar, M., & Schneiderman, N. (1990). Emotional repression, stress disclosure responses, and Epstein-Barr viral capsid antigen titers. *Psychosomatic Medicine, 52,* 397–410.

Fawzy, F. I., Fawzy, N. W., Hyun, C. S., Elashoff, R., Guthrie, D., Fahey, J. L., & Morton, D. L. (1993). Malignant melanoma: Effects of an early structured psychiatric intervention, coping, and affective state on recurrence and survival 6 years later. *Archives of General Psychiatry, 50,* 681–689.

Florin, I., Freudenberg, G., & Hollander, J. (1985). Facial expressions of emotion and physiologic reactions in children with bronchial asthma. *Psychosomatic Medicine, 47,* 382–393.

Foa, E., & Kozak, M. (1986). Emotional processing of fear: Exposure to corrective information. *Psychological Bulletin, 99,* 20–35.

Foa, E., & Riggs, D. (1993). Post-traumatic stress disorder in rape victims. In J. Oldham, M. B. Riba, & A. Tasman (Eds.), *American Psychiatric Press review of psychiatry* (Vol. 12, pp. 273–303). Washington, DC: American Psychiatric Press.

Foa, E., Riggs, D., Massie, E., & Yarczower, M. (1993). Post-traumatic stress disorder in rape victims: Exposure to corrective information. *Psychological Bulletin, 99,* 20–35.

Foa, E., Rothbaum, B., & Molnar, C. (1995). Cognitive–behavioral treatment of post-traumatic stress disorder. In M. Friedman, D. S. Charney, & A. Y. Deutch (Eds.), *Neurobiological and clinical consequences of stress: From normal adaptation to post-traumatic stress disorder* (pp. 483–494). New York: Raven Press.

Foa, E., Steketee, G., & Rothbaum, B. (1989). Behavioral/cognitive conceptualizations of post-traumatic stress disorder. *Behavior Therapy, 20,* 155–176.

*Francis, M. E., & Pennebaker, J. W. (1992). Putting stress into words: The impact of writing on physiological, absentee, and self-reported emotional well-being measures. *American Journal of Health Promotion, 6,* 280–287.

Goldstein, H., Edelberg, R., Meier, C., & Davis, L. (1988). Relationship of blood pressure and heart rate to experienced anger and expressed anger. *Psychosomatic Medicine, 50,* 321–329.

Glass, G. V. (1976). Primary, secondary, and meta-analysis. *Educational Researcher, 5,* 3–8.

*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). Health and psychological effects of emotional disclosure: A test of the inhibition–confrontation approach. *Journal of Personality and Social Psychology, 71,* 588–602.

Grinker, R. R., & Spiegel, J. P. (1945). *Men under stress.* Philadelphia: Blakiston.

Harber, K., & Pennebaker, J. (1992). Overcoming traumatic memories. In S. A. Christianson (Ed.), *The handbook of emotion and memory: Research and theory* (pp. 359–386). Hillsdale, NJ: Erlbaum.

Hedges, L. V. (1981). Distribution theory for Glass's estimator of effect size and related estimators. *Journal of Educational Statistics, 6,* 107–128.

Hedges, L. V., & Olkin, I. (1985). *Statistical methods for meta-analysis.* Orlando, FL: Academic Press.

Herman, J. (1992). *Trauma and recovery.* New York: Basic Books.

Horowitz, M. (1986). *Stress response syndromes* (2nd ed.). New York: Jason Aronson.

*Hughes, C. F. (1993). *Effects of expressing negative and positive emotions and insight on health and adjustment to college.* Unpublished doctoral dissertation, Southern Methodist University.

Hunter, J. E., Schmidt, F. L., & Jackson, G. B. (1982). *Meta-analysis: Cumulating research findings across studies.* Beverly Hills, CA: Sage.

Jamner, L., Schwartz, G., & Leigh, H. (1988). The relationship between repressive and defensive coping styles and monocyte, eosinophile, and serum glucose levels: Support for the opioid peptide hypothesis of repression. *Psychosomatic Medicine, 50,* 567–575.

Janet, P. (1909). *Les névroses.* Paris: Flammarion.

Jensen, M. (1987). Psychobiological factors predicting the course of breast cancer. *Journal of Personality, 55,* 317–342.

Johnson, B. T. (1990). *DSTAT: Software for the meta-analytic review of research literatures.* Hillsdale, NJ: Erlbaum.

Krantz, A., & Pennebaker, J. W. (1995). *Bodily versus written expression of traumatic experience.* Manuscript submitted for publication.

L'Abate, L. (1992). *Programmed writing: A self-administered approach for interventions with individuals, couples, and families.* Pacific Grove, CA: Brooks/Cole.

L'Abate, L., Boyce, J., Fraizer, L., & Russ, D. (1992). Programmed

writing: Research in progress. *Comprehensive Mental Health Care, 2,* 45–62.

L'Abate, L., Boyce, J., Russ, D., & Bird, G. (in press). Programmed writing: Two follow-ups and one application with clinical out-patients. *Contemporary Psychodynamics.*

Lange, A. (1994). Writing assignments in the treatment of grief and traumas from the past. In J. Zeig (Ed.), *Ericksonian methods: The essence of the story* (pp. 377–392). New York: Brunner/Mazel.

Larson, D. (1990). Self-concealment: Conceptualization, measurement, and health implications. *Journal of Social and Clinical Psychology, 9,* 439–455.

Lepore, S. J., Silver, R. C., Wortman, C. B., & Wayment, H. A. (1996). Social constraints, intrusive thoughts, and depressive symptoms among bereaved mothers. *Journal of Personality and Social Psychology, 70,* 271–282.

Lipsey, M., & Wilson, D. (1993). The efficacy of psychological, educational, and behavioral treatment. Confirmation from meta-analysis. *American Psychologist, 48,* 1181–1209.

Lutgendorf, S., Antoni, M., Kumar, M., & Schneiderman, N. (1994). Changes in cognitive coping strategies predict EBV-antibody titre change following a stressor disclosure induction. *Journal of Psychosomatic Research, 38,* 63–78.

McFarlane, A. (1988). The longitudinal course of posttraumatic morbidity: The range of outcomes and their predictors. *Journal of Nervous and Mental Disease, 176,* 30–39.

McFarlane, A. (1992). Avoidance and intrusion in posttraumatic stress disorder. *Journal of Nervous and Mental Disease, 180,* 439–445.

Meyer, T., & Mark, M. (1995). Effects of psychosocial interventions with adult cancer patients: A meta-analysis of randomized experiments. *Health Psychology, 14,* 101–108.

Miller, S., Shoda, Y., & Hurley, K. (1996). Applying cognitive–social theory to health-protective behavior: Breast self-examination in cancer screening. *Psychological Bulletin, 119,* 70–94.

Mumford, E., Schlesinger, H. J., & Glass, G. V. (1983). Reducing medical costs through mental health treatment: Research problems and recommendations. In A. Broskowski, E. Marks, & S. K. Budman (Eds.), *Linking health and mental health* (pp. 257–273). Beverly Hills, CA: Sage.

Murray, E. J., Lamnin, A. D., & Carver, C. S. (1989). Emotional expression in written essays and psychotherapy. *Journal of Social and Clinical Psychology, 8,* 414–429.

Murray, E. J., & Segal, D. L. (1994). Emotional processing in vocal and written expression of feelings about traumatic experiences. *Journal of Traumatic Stress, 7,* 391–405.

*O'Heeron, R. C. (1992). *Confronting life events: Inner-city adolescents write about personal traumas.* Unpublished doctoral dissertation, California School of Professional Psychology.

Pennebaker, J. (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. (1990). *Opening up: The healing power of confiding in others.* New York: Morrow.

Pennebaker, J. (1991, Jan/Feb). "Writing your wrongs." *American Health, 10,* 64–67.

Pennebaker, J. (1993). Putting stress into words: Health, linguistic, and therapeutic implications. *Behavioral Research 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., & Chew, C. H. (1985). Deception, electrodermal activity, and inhibition of behavior. *Journal of Personality and Social Psychology, 49,* 1427–1433.

*Pennebaker, J. W., Colder, M., & Sharp, L. K. (1990). Accelerating the coping process. *Journal of Personality and Social Psychology, 58,* 528–537.

*Pennebaker, J., & Francis, M. (1996). Cognitive, emotional, and language processes in disclosure. *Cognition and Emotion, 10,* 601–626.

Pennebaker, J. W., & Harber, K. (1993). A social stage model of collective coping: The Loma Prieta earthquake and the Persian Gulf war. *Journal of Social Issues, 49*(4), 125–146.

Pennebaker, J. W., Hughes, C. F., & O'Heeron, R. C. (1987). The psychophysiology of confession: Linking inhibitory and psychosomatic processes. *Journal of Personality and Social Psychology, 52,* 781–793.

*Pennebaker, J. W., Kiecolt-Glaser, J., & Glaser, R. (1988). Disclosure of traumas and immune function: Health implications for psychotherapy. *Journal of Consulting and Clinical Psychology, 56,* 239–245.

Pennebaker, J. W., Mayne, T., & Francis, M. (1997). Linguistic predictors of adaptive bereavement. *Journal of Personality and Social Psychology, 72,* 863–871.

Pennebaker, J. W., & O'Heeron, R. C. (1984). Confiding in others and illness rate among spouses of suicide and accidental death victims. *Journal of Abnormal Psychology, 93,* 473–476.

Pennebaker, J. W., & Susman, J. R. (1988). Disclosure of traumas and psychosomatic processes. *Social Science and Medicine, 26,* 327–332.

*Petrie, K., Booth, R., & Pennebaker, J. (1995). Disclosure of trauma and immune response to a hepatitis B vaccination program. *Journal of Consulting and Clinical Psychology, 63,* 787–792.

Pitman, R., Altman, B., Greenwald, E., Longpre, R., Macklin, M., Poire, R., & Steketee, G. (1991). Psychiatric complications during flooding therapy for posttraumatic stress disorder. *Journal of Clinical Psychiatry, 52,* 17–20.

Ptacek, J. T., Smith, R. E., & Zanas, J. (1992). Gender, appraisal, and coping: A longitudinal analysis. *Journal of Personality, 60,* 747–770.

Rachman, S. J. (1980). Emotional processing. *Behavior Research and Therapy, 18,* 51–60.

*Richards, J., Pennebaker, J., & Beall, W. (1995). *The effects of criminal offense and disclosure of trauma on anxiety and illness behavior in prison inmates.* Unpublished manuscript.

Rosenthal, R. (1984). *Meta-analytic procedures for social research.* Beverly Hills, CA: Sage.

Rosenthal, R. (1995). Writing meta-analytic reviews. *Psychological Bulletin, 118,* 183–191.

Rosenthal, R., & Rubin, D. (1982). A simple, general pupose display of magnitude of experimental effect. *Journal of Educational Psychology, 74,* 166–199.

Scheff, T. J. (1979). *Catharsis in healing, ritual, and drama.* Berkeley: University of California Press.

Shadish, W., & Haddock, C. (1994). Combining estimates of effect size. In H. Cooper & L. Hedges (Eds.), *The handbook of research synthesis.* New York: Sage.

Shadish, W. R., & Ragsdale, K. (1996). Random versus nonrandom assignment in controlled experiments: Do you get the same answer? *Journal of Consulting and Clinical Psychology, 64,* 1290–1305.

Smith, M., & Glass, G. (1977). Meta-analysis of psychotherapy outcome studies. *American Psychologist, 32,* 752–760.

Smith, M., Glass, G., & Miller, T. (1980). *The benefits of psychotherapy.* Baltimore: Johns Hopkins University Press.

Solomon, Z., Avitzur, E., & Mikulincer, M. (1990). Coping styles and post-war psychopathology among Israeli soldiers. *Personality and Individual Differences, 11,* 451–456.

*Spera, S., Buhrfeind, E., & Pennebaker, J. (1994). Expressive writing and coping with job loss. *Academy of Management Journal, 3,* 722–733.

Spiegel, D., Bloom, J. R., Kraemer, H. C., & Gottheil, E. (1989). Effects

of psychosocial treatment of patients with metastatic breast cancer. *Lancet, ii,* 888–891.

Stock, W. (1994). Systematic coding for research synthesis. In H. Cooper & L. Hedges (Eds.), *The handbook of research synthesis.* New York: Sage.

Suls, J., & Fletcher, B. (1985). The relative efficacy of avoidant and non-avoidant coping strategies: A meta-analysis. *Health Psychology, 4,* 249–288.

Terr, L. (1993). *Unchained memories.* New York: Basic Books.

van der Hart, O., Steele, K., Boon, S., & Brown, P. (1993). The treatment of traumtic memories: Synthesis, realization, and integration. *Dissociation, 6,* 162–180.

van der Kolk, B. (1994). The body keeps the score: Memory and the evolving psychobiology of posttraumatic stress. *Harvard Review of Psychiatry, 1,* 253–265.

van der Kolk, B., Blitz, R., Burr, W., & Hartmann, E. (1984). Nightmares and trauma: Lifelong and traumatic memories in veterans. *American Journal of Psychiatry, 141,* 187–190.

van der Kolk, B., & van der Hart, O. (1989). Pierre Janet and the breakdown of adaptation in psychological trauma. *American Journal of Psychiatry, 146,* 1530–1540.

van der Kolk, B., & van der Hart, O. (1991). The intrusive past: The flexibility of memory and the engraving of trauma. *American Imago, 48,* 425–454.

Wells-Parker, E., Bangert-Drowns, R., McMillen, R., & Williams, M. (1995). Final results from a meta-analysis of remedial interventions with drink/drive offenders. *Addiction, 90,* 907–926.

Willensky, D. (1993, June). "Writing off the unemployment blues." *American Health, 12,* 35.

Wortman, P., Smyth, J., Langenbrunner, J., & Yeaton, W. (in press). Clinical consensus panels and research synthesis: A comparison of methods. *International Journal of Technology Assessment in Health Care.*

Yeaton, W., Langenbrunner, J., Smyth, J., & Wortman, P. (1995). Exploratory research synthesis: Methodological considerations for addressing limitations in data quality. *Evaluation and the Health Professions, 18,* 283–303.

Received April 9, 1996
Revision received November 20, 1996
Accepted June 12, 1997 ■