Emotion
2004, Vol. 4, No. 2, 207–214

1528-3542/04/$12.00 DOI: 10.1037/1528-3542.4.2.207

Inducing Affective States With Success–Failure Manipulations: A Meta-Analysis

Lauri Nummenmaa
University of Turku

Pekka Niemi
University of Turku and Stavanger University College

Meta-analytic techniques were used to assess whether successes and failures can be used experimentally to induce affective states. Data from 32 studies, with a total of 2,468 participants, were reviewed. Methods for producing success–failure experiences, as well as the resulting affective reactions, were analyzed. Effect sizes as a result of various methods of induction were calculated. The success–failure manipulation turned out to be capable of reliably inducing both positive and negative affective reactions. A framework for using success–failure manipulations in affect induction is presented.

Using an affective state as an independent variable has been one of the standard paradigms in emotion research. Although there are some reliable observations of natural situations causing affective reactions (see, e.g., Parrott & Sabini, 1990), on many occasions, the use of naturally occurring emotions might be impractical, unreliable, or even impossible. Thus, there have been numerous attempts to develop experimental procedures that would induce affective states in participants (for reviews, see Gerrards-Hesse & Spies, 1994; Martin, 1990; Westermann, Spies, Stahl, & Hesse, 1996). Essential presuppositions are that the resulting affective reactions be similar across individuals, and the different affective reactions resulting from different manipulations be clearly recognized and differentiated.

It is obvious that induction paradigms vary in terms of advantages and disadvantages, not least with respect to validity. Their theoretical rationales can be questionable (Clore, 1994, p. 183); it is sometimes unclear whether the procedure induces mood or emotion (see Gerrads-Hesse & Spies, 1994). Onset and offset of affective reactions are also often unclear, and sometimes the affective reactions might be too mild (Philippot, 1993). Perhaps the greatest problem is related to the ecological validity of the induction procedure. Watching films or pictures, listening to music, or imagining oneself in various situations certainly bears an analogy to affects in life outside the laboratory. However, in such situations, people are not *involved* in the emotion-eliciting situation. Therefore, there is a clear need to develop standardized emotion-eliciting procedures that resemble the real-life situations in which people experience emotions.

Success–Failure Manipulations as a Possible Alternative

Successes and failures are experienced in many everyday situations, and they usually result in affective reactions. Such experiences can be produced in the laboratory with relatively simple arrangements. What is more, doing so makes it possible to avoid the problems with theoretical rationale and ecological validity. The affective reactions to successes and failures depend on people's motivation and how they estimate their ability to succeed (Atkinson, 1957). Additionally, the resulting affects can be seen as caused by attribution of the causal locus, the stability and controllability of the success–failure situation (Weiner, 1986). Moreover, the resulting affect can be seen as caused by appraisal of the success–failure experience (Scherer, 1999).

In success–failure manipulation (SFM), the participant is presented with a need-related task, and the

Lauri Nummenmaa, Department of Psychology, University of Turku, Turku, Finland; Pekka Niemi, Department of Psychology, University of Turku, and Center for Reading Research, Stavanger University College, Stavanger, Norway.

Correspondence concerning this article should be addressed to Lauri Nummenmaa, Department of Psychology, University of Turku, FIN-20014, Turku, Finland. E-mail: lauri.nummenmaa@utu.fi

actual or perceived outcome of the task is manipulated so that the participant experiences either success or failure. In the simplest form of SFM, the participant is asked to compare his or her performance with that of some criterion group. However, the task itself is so ambiguously constructed that it is virtually impossible for the participant to estimate his or her own success before the feedback is given. Alternatively, the experiences of success and failure can be produced through manipulation of the task difficulty. In SFM, task feedback or actual performance on the task is thought to affect the person's appraisals of the situation that, in turn, modulate the experienced emotion.

SFMs have other significant advantages. Sometimes it is crucial that the participant does not know that his or her emotional state is being manipulated (Parrott & Hertel, 1999). When using success and failure in affect induction, it is fairly simple to disguise the true intention of the task (i.e., affect induction). Another advantage is that with SFM, it is possible to pinpoint the onset of the affective reaction, which is important in physiological measurements of emotion, such as electrodermal activity, in which the responses are modulated on a varying baseline. Probably the greatest advantage of the SFM is its strong ecological validity, as the participants are actually acting in and experiencing the affect-eliciting situations. If it can be empirically shown that experimentally produced successes and failures systematically result in different affective reactions, and a sufficient standardization of the experimental manipulation can be attained, success and failure experiences might offer an ecologically valid means of manipulating the participant's affective state. The aim of this meta-analysis was to evaluate the merits of experimental SFMs in this respect.

The Meta-Analysis

Literature Search

The meta-analysis includes studies published through the end of August 2002. Several search methods were used. The PsycINFO and ERIC databases were searched to retrieve documents containing the terms *emotion, mood, affective, success,* and *failure* either in article title, abstract, or keywords; a similar search was conducted on the EBSCOhost, Academic Press IDEAL, and Elsevier Science Direct full-text electronic journal databases; articles referred to in articles found by the two preceding methods were examined. Studies were accepted for the meta-analysis if they met the following criteria: (a) The study had investigated either experimentally or quasi-experimentally the relation between success–failure experience and affective reaction labeled either *mood, emotion,* or *affect;* (b) the participants were involved in actual situations resulting in successes or failures. For example, studies using imagination of success and failure were omitted; (c) participants' affective state was measured with a self-report. Altogether, the meta-analysis included data from 2,468 participants in 32 studies.

The following questions were addressed: (a) What procedures can be used for experimental production of success and failure experiences? (b) What affective reactions do successes and failures result in? (c) What is the relative strength of the different affects resulting from successes and failures? (d) What factors in the design of the study determine the strength of the affective reaction?

The meta-analysis proceeded as follows: Each study was coded with respect to sample size, affect measure, resulting affects, employed SFM, and experimental design; whenever possible, the association between success–failure experience and experienced affect was transformed into a common measure of effect size (r); the effect sizes from different studies were combined and their magnitude was examined; and the relationships between effect sizes and experimental design and the SFM used were examined.

Tasks Used in SFMs

Use of SFM offers a wide array of possible tasks in which the participant can succeed or fail. For the purposes of the meta-analysis, we classified them as follows: (a) Bogus[1] intelligence tests ($n = 8$), (b) other cognitive tasks ($n = 13$), (c) exam feedback ($n = 3$), and (d) tests of social perception skills ($n = 4$). Miscellaneous studies on affective reactions after sports competition, interpersonal rejection, and different motor tasks were treated separately ($n = 4$).

Categorizing the Affective Reactions

The terminology describing affective reactions is diverse, with no consensus on how to describe affective states in words. Success has been reported to result in, for example, *positive affect, joviality, gladness, happiness, pride, self-assurance, attentiveness, curiosity,* and *gratefulness*. Failure has been reported

[1] By *bogus* we refer to a test that is either a real test that has been manipulated in some way or a test that does not actually tap intelligence at all.

to result in, for example, *negative affect, depressed mood, depression, anxiety, hostility,* and *shame*. Because of the muddled terminological clarity, a classification of measured affects was necessary to obtain interpretable results from the meta-analysis. We classified the affects into three categories according to their valence (positive-negative). Terms such as *depression, negative affect, fear, sadness, hostility, guilt, shame,* and *anxiety* were classified as negative affect (Category 1). *Positive affect, joviality, self-assurance, attentiveness, serenity, pride,* and *happiness* were classified as positive affect (Category 2). When the affective state was measured on a bipolar scale (positive-negative), we labeled the affect *bipolar valence* (Category 3). We calculated total affectivity per study as the mean of all effect sizes from the different measured affects in each study. If the affect term did not fall into Categories 1–3, we did not analyze it separately but included it in total affectivity. It is clear that analysis based only on the valence dimension of affects is rather crude. However, very fine-grained categorization would have made it impossible to draw any major conclusions because many affects were measured only in a few studies. The categorization based on valence made it possible to combine the results of the reviewed studies effectively.

Calculating and Combining Effect Sizes

Although statisticians have suggested reporting the effect size in addition to p values for decades (Olejnik & Algina, 2000), none of the reviewed 32 studies did so. We calculated effect sizes (r) for affective reactions in success–failure situations to estimate the strength of affective reactions resulting from success versus failure. Because very few studies used a control group, we decided to compare the postmanipulation affectivity between success and failure situations. If a pretest of affect was administered, we used emotion change scores when calculating effect sizes because using only postmanipulation scores in all instances would have led to less exact estimates of effect size. If effects were reported to be nonsignificant but sufficient information was given, we calculated r nonetheless.

Calculating effect sizes was based on means and variances and the number of participants in each treatment group, or alternatively, F-test values and degrees of freedom (see Rosenthal, 1984; Rosenthal & DiMatteo, 2001). If only sample size and p value were given, we calculated a conservative estimate of effect size by converting the p values to the corresponding standard normal deviate equivalent using the table of Z values and dividing the Z score by the square root of the sample size. If groups were referred to as *balanced,* we assumed equal group sizes in all computations. If the information was insufficient for all the above procedures, or the design of the study or its reporting did not permit analysis of main effects of success and failure, we assumed r to be zero, which gives a conservative estimate of effect. It should be noted, however, that such statistical methods for calculating effect sizes rely on procedures that assume normality in the original population. Because tests of normality assumption were not reported in any of the targeted studies, our analysis is based on the assumption that tested distributions stem from normally distributed populations. Although not a serious threat, violation of the assumption of normality in the original data may affect our results.

In multifactor designs, we collapsed the data over groups. Consequently, these analyses have a maximum of one fixed factor (success–failure). We were able to compute effect sizes from 32 studies, resulting in 42 effect sizes. These effect sizes cannot be considered to be independent of each other because the analysis included more than one study from a number of laboratories and because multiple affects were measured in many studies (we calculated effect sizes for each of the affects measured).

The reviewed studies applied a wide variety of methods for assessing the participants' affective state (see the Appendix). Although there is evidence for an affect measure being associated with the obtained effect sizes (Westermann et al., 1996), analyzing effect sizes measure by measure would have led to uninformative and scattered results, as the number of manipulation checks was large. Therefore, we have pooled all the results by affect regardless of the applied measure.

To obtain weighted mean effect sizes (r_m) for the affect categories presented above, we transformed individual effect sizes by using Fisher's Z-r transformation and calculated weighted Z_r means. We transformed mean Z_r:s back to r_m:s. We calculated a 95% confidence interval for each mean effect size. Average effect sizes, with confidence intervals for negative and positive affect and bipolar valence, are presented in Table 1. Mean effect sizes and their confidence intervals were similarly calculated for effects of manipulation type and task type on average affectivity (see Table 1). The confidence intervals can be used to test whether two correlation coefficients are statistically different from each other. If there is no overlap, the correlations can be considered unequal.

Table 1
Mean Effect Sizes (M_r) and 95% Confidence Intervals (CI) for Manipulation Type and its Effect on Average Affectivity and Tasks Used In Success–Failure Manipulations (SFM) and Their Effect on Average Affectivity

Variable	M_r	*CI*	n_{part}	*n*	File drawer estimate
Measured affect					
1. Negative affect	0.32	0.30, 0.35	1,312	17	2,655
2. Positive affect	0.34	0.30, 0.36	971	11	2,028
3. Bipolar valence	0.39	0.36, 0.41	791	14	4,018
4. Average affectivity	0.37	0.35, 0.39	2,540	32	6,355
Manipulation type					
1. Feedback	0.37	0.35, 0.39	1,563	20	3,967
2. Task difficulty	0.38	0.35, 0.41	763	9	1,994
3. Quasi-experimental	0.30	0.24, 0.36	214	3	394
Task used in SFM					
1. Bogus intelligence tests$_{2,4,5}$	0.33	0.30, 0.37	602	8	1,284
2. Other cognitive tasks$_{1,3,4,5}$	0.42	0.39, 0.44	996	13	2,993
3. Exam feedback$_{2,4,5}$	0.30	0.25, 0.35	317	3	569
4. Test of social perception skills$_{1,2,3,5}$	0.60	0.56, 0.64	244	4	1,300
5. Other tasks$_{1,2,3,4}$	0.17	0.12, 0.22	381	4	210

Note. Subscripts indicate nonoverlapping confidence intervals in relation to other tasks in subtable. n_{part} = number of participants.

Because nonsignificant results stand little chance of being published, the present sample of studies might not comprise a random sample of all SFM studies conducted. Therefore, we conducted a file drawer analysis to estimate whether the present results were robust against the file drawer problem. Our method was based on Rosenthal's (1995) suggestions; that is, we estimated the number of participants in studies with null effects required to lower the r_m:s below .1, which is traditionally considered the lower limit of a small effect. This is a conservative method for a file drawer analysis, as null effects are defined as an effect size of zero.

Results

If one or several confidence intervals for effect sizes include zero, the usefulness of the associated affect manipulation would be seriously undermined. The data compiled for Table 1 pass this test. In fact, all the effect sizes for affective reactions actually exceed the traditional limit of moderate effect ($r > .24$). Hence, we can conclude that experimentally or quasi-experimentally induced successes and failures reliably result in at least moderate affective reactions. Manipulating task difficulty, giving false feedback, or using quasi-experimental SFMs proved to be equally effective (see Table 1). Reliable effects resulted after all manipulations. Hence, SFMs can be used in both experimental and quasi-experimental designs. It is not, however, irrelevant which task is used. Success or failure in different tasks results in affective reactions of different magnitudes (see Table 1). Tests of social perception skills clearly elicit the strongest affects. Bogus intelligence tests, cognitive tasks, and exam feedback are less powerful, with the mixed category, other tasks, being the weakest method. Moreover, the file drawer analysis suggests that approximately three times more data with null results should exist to invalidate the effects discovered in this meta-analysis. Therefore, we conclude that the file drawer problem is not a serious threat to our results.

Discussion

The meta-analysis shows that successes and failures do result in different affective reactions. This can safely be concluded despite the fact that some studies included in the analysis suffered from methodological flaws (e.g., pretest of affective state was not measured in some studies). The confidence intervals of mean effect sizes show that, first, both positive and negative affective reactions differ significantly from zero, and second, they are of approximately equal magnitude (see Table 1). Consequently, the reviewed studies support the hypothesis that successes and failures reliably result in affects that are clearly differentiated on the positive-negative axis. Gerrads-Hesse and Spies (1994) suggested that it is easier to induce negative than positive affect with SFMs. This notion is not supported by our meta-analysis, as confidence intervals of effect sizes for positive and negative affect do not overlap. Failure increases the negatively valenced affect, whereas success increases the positively valenced affect. Beyond this, the analysis of resulting affective reactions is, however, controversial because of the classification of affects used in the analysis.

Affective Reactions to Successes and Failures

Why, then, do different studies report different affective reactions resulting from quite similar cognitive tasks? The first possibility is that measurement scales vary among studies. Some scales (e.g., bipolar scale negative-positive) are not extensive enough with regard to possible affective reactions, whereas affect terminology in others is ambiguous. It is possible that successes and failures result in simultaneous experiences of multiple affective reactions, but the measurement scales unnecessarily limit the participants' responses. It is possible that different emotions occur partly because of the construction of the success–failure task itself, but some variance is clearly the result of the different measures used. Such confusion could clearly be avoided using scales drawn from well-established emotion terminology such as basic emotions or scales like the Positive and Negative Affect Schedule (Watson, Clark, & Tellegen, 1988) or Self-Assessment Manikin ratings (see, e.g., Lang, Bradley, & Cuthbert, 2001) that have been tested for reliability and validity.

An alternative explanation is that the actual affective reaction is partly determined by the appraisals made in the situation. The emotional reactions to successes and failures depend on how the events are attributed (Weiner, 1986). However, analysis of the emotion-eliciting situations and the possible appraisals were rendered impossible because explicit information about task instructions, feedback, and task setting were not routinely given.

How to Manipulate Successes and Failures

Our results showed that both random and task-congruent feedback result in affective reactions congruent with the feedback, as mean effect sizes for task difficulty manipulation and feedback manipulation do not differ (see Table 1). Although false feedback results in affective reactions that are congruent with the feedback, it may make participants suspicious of the true meaning of the task, given that the task is not constructed in such a way that the participant is not able to estimate his or her own performance. Therefore, extreme feedback should be avoided (Hammen & Krantz, 1976). For example, university students will probably not believe that they score in the 10th percentile of overall norms on an intelligence test. It should also be noted that the timing of the presentation of standards for comparison might influence the affective reaction. These are more pronounced when the standards are presented after, as opposed to before, the performance (Schul & Schiff, 1995).

From the point of view of appraisal theories, the task should be somehow ego-involving in order to result in an affective reaction. This means that the task should be planned so that the success versus failure is significant to the participant. Neither success nor failure produces an emotional response in all situations. Therefore, the choice of a certain task should always be justified. SFM can also be used quasi-experimentally in natural situations in which people experience success or failure. Such procedures result in clearly differentiated affective reactions (see Table 1). The major disadvantage with using naturally occurring emotional reactions is that participants cannot be randomly assigned to different conditions (Parrott & Hertel, 1999). However, a quasi-experimental method has certain advantages if used carefully. First, in many natural situations, the success or failure is meaningful to the individual, and second, the resulting affective reactions can be strong if the situation is meaningful to the individual.

What Tasks Should Be Used in SFMs?

Our meta-analysis suggests that the most powerful affective reactions have resulted from social perception tasks. Cognitive tasks, bogus intelligence tests, motor performance tests, and exam feedback elicited affective reactions of lesser but mutually equal strength (see Table 1). Very few authors, however, have justified their choice of task. Mere success or failure in any task is not a sufficient elicitor of affect. Therefore, the participant must be involved in the task, considering success or failure important. Without this, he or she cannot make appraisals that lead to an experience of affect. Therefore, the ecological validity of the task is improved if the task is connected to meaningful situations in which people typically experience successes or failures. From this point of view, different social skills tests and bogus intelligence or ability tests are suitable for students or adults. Only one study involved children (Ward, Friedlander, & Silverman, 1987), so we cannot draw conclusions about suitable tasks for them, but it is likely that any involving and competitive play/game task would be relevant (see Salonen, Lepola, & Niemi, 1998).

Conclusion

Experimental studies have mainly focused on emotion induction procedures that are easy to control and standardized at the possible expense of ecological validity. Successes and failures frequently experienced in everyday life ensure affective reactions. In the laboratory, it is fairly easy to create situations in which participants can succeed or fail, and these situations can be made very realistic. Our meta-analysis suggests that resulting positive and negative affective reactions are of clearly measurable magnitude, and the SFM can also be easily controlled and standardized. However, on the basis of this meta-analysis, it was impossible to conclude how success–failure tasks should be designed to induce differentiated emotional states. Despite the above reservations, the SFM seems to offer an ecologically valid and effective method for inducing affective states. On the basis of the results of this meta-analysis, we suggest the following when inducing affects with SFM:

1. The need-related situation should be involving enough for the participants; that is, the participants should have the need and will to achieve. From this point of view, tests of social skills and bogus intelligence tests are well-suited tasks for adults.

2. Manipulation of task difficulty should be preferred over giving false feedback.

3. The SFM can most reliably be used in studies in which the valence of emotional state is to be induced—the current evidence is not sufficient to conclude whether such manipulations are suitable for the reliable induction of discrete emotional states.

References

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

Atkinson, J. W. (1957). Motivational determinants of risk-taking behavior. *Psychological Review, 64,* 359–372.

*Baucom, D. H., & Aiken, P. A. (1981). Effects of depressed mood on eating among obese and nonobese dieting and nondieting persons. *Journal of Personality and Social Psychology, 41,* 577–585.

*Brown, J. D., & Dutton, K. A. (1995). The thrill of victory, the complexity of defeat: Self-esteem and people's emotional reactions to success and failure. *Journal of Personality and Social Psychology, 68,* 712–722.

*Brown, J. D., & Marshall, M. A. (2001). Self-esteem and emotion: Some thoughts about feelings. *Personality and Social Psychology Bulletin, 27,* 575–584.

*Chartier, G. M., & Ranieri, D. J. (1989). Comparison of two mood induction procedures. *Cognitive Therapy and Research, 13,* 275–282.

Clore, G. L. (1994). Why emotions require cognition. In P. Ekman & R. J. Davidson (Eds.), *The nature of emotion: Fundamental questions* (pp. 181–191). New York: Oxford University Press.

*Cohen, L. H., Pane, N., & Smith, H. S. (1997). Complexity of the interpersonal self and affective reactions to interpersonal stressors in life and in the laboratory. *Cognitive Therapy and Research, 21,* 387–407.

*Egloff, B. (1998). The independence of positive and negative affect depends on the affect measure. *Personality and Individual Differences, 25,* 1101–1109.

*Egloff, B., & Krohne, H. W. (1996). Repressive emotional discreteness after failure. *Journal of Personality and Social Psychology, 70,* 1318–1326.

*Feick, D. L., & Rhodewalt, F. (1997). The double-edged sword of self-handicapping: Discounting, augmentation, and the protection and enhancement of self-esteem. *Motivation and Emotion, 21,* 147–163.

*Forgas, J. P., Bower, G. H., & Moylan, S. J. (1990). Praise or blame? Affective influences on attributions of achievement. *Journal of Personality and Social Psychology, 59,* 809–819.

*Gendolla, G. H. E. (1997). Surprise in the context of achievement: The role of outcome valence and importance. *Motivation and Emotion, 21,* 165–193.

Gerrads-Hesse, A., & Spies, K. (1994). Experimental inductions of emotional states and their effectiveness: A review. *British Journal of Psychology, 84,* 55–79.

Hammen, C. L., & Krantz, S. (1976). Effect of success and failure on depressive cognitions. *Journal of Abnormal Psychology, 85,* 577–586.

*Idson, L. C., Liberman, N., & Higgins, E. T. (2000). Distinguishing gains from nonlosses and losses from nongains: A regulatory focus perspective of hedonic intensity. *Journal of Experimental Social Psychology, 36,* 252–274.

*Ingram, R. E. (1984). Information processing and feedback: Effects of mood and information favorability on the cognitive processing of personally relevant information. *Cognitive Therapy and Research, 8,* 371–386.

*Ingram, R. E., Johnson, B. R., Bernet, C. Z., Dombeck, M., & Rowe, M. (1992). Vulnerability to distress: Cognitive and emotional reactivity in chronically self-focused individuals. *Cognitive Therapy and Research, 16,* 451–472.

*Ingram, R. E., Smith, T. W., & Brehm, S. S. (1983). Depression and information processing: Self-schemata and the encoding of self-referent information. *Journal of Personality and Social Psychology, 45,* 412–420.

*Krohne, H. W., Pieper, M., Knoll, N., & Breimer, N. (2002). The cognitive regulation of emotions: The role of success versus failure experience and coping dispositions. *Cognition & Emotion, 16,* 217–243.

Lang, P. J., Bradley, M. M., & Cuthbert, B. N. (2001). *International Affective Picture System (IAPS): Instruction manual and affective ratings* (Tech. Rep. No. A-5). Gainesville: University of Florida, The Center for Research in Psychophysiology.

*Larsen, R. J., & Ketelaar, T. (1989). Extraversion, neuroticism and susceptibility to positive and negative mood induction procedures. *Personality and Individual Differences, 12,* 1221–1228.

*Linville, P. W. (1985). Self-complexity and affective extremity: Don't put all of your eggs in one cognitive basket. *Social Cognition, 3,* 94–120.

Martin, M. (1990). On the induction of mood. *Clinical Psychology Review, 10,* 669–697.

*McFarland, C., & Buehler, R. (1997). Negative affective states and the motivated retrieval of positive life events: The role of affect acknowledgment. *Journal of Personality and Social Psychology, 73,* 200–214.

*McFarland, C., & Ross, M. (1982). Impact of causal attributions on affective reactions to success and failure. *Journal of Personality and Social Psychology, 43,* 937–946.

*Mehlman, R. C., & Snyder, C. R. (1985). Excuse theory: A test of the self-protective role of attributions. *Journal of Personality and Social Psychology, 49,* 994–1001.

*Neumann, R., Seibt, B., & Strack, F. (2001). The influence of mood on the intensity of emotional responses: Disentangling feeling and knowing. *Cognition & Emotion, 15,* 725–747.

Olejnik, S., & Algina, J. M. (2000). Measures of effect size for comparative studies: Applications, interpretations, and limitations. *Contemporary Educational Psychology, 25,* 241–286.

Parrott, W. G., & Hertel, P. (1999). Research methods. In T. Daleigsh & T. Power (Eds.), *Handbook of cognition and emotion* (pp. 3–19). West Sussex, England: Wiley.

Parrott, W. G., & Sabini, J. (1990). Mood and memory under natural conditions: Evidence for mood incongruent recall. *Journal of Personality and Social Psychology, 59,* 321–336.

Philippot, P. (1993). Inducing and assessing differentiated emotion-feeling states in the laboratory. *Cognition & Emotion, 7,* 171–193.

*Rhodewalt, F., & Morf, C. C. (1998). On self-aggrandizement and anger: A temporal analysis of narcissism and affective reactions to success and failure. *Journal of Personality and Social Psychology, 74,* 672–685.

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

Rosenthal, R. (1995). Writing meta-analytic reviews. *Psychological Bulletin, 118,* 183–192.

Rosenthal, R., & DiMatteo, M. R. (2001). Meta-analysis: Recent developments in quantitative methods for literature reviews. *Annual Review of Psychology, 52,* 59–82.

Salonen, P., Lepola, J., & Niemi, P. (1998). The development of first graders' reading skill as a function of preschool motivational orientation and phonemic awareness. *European Journal of Psychology of Education, 13,* 155–174.

*Sanna, L. J., Meier, S., & Wegner, E. A. (2001). Counterfactuals and motivation: Mood as input to affective enjoyment preparation. *British Journal of Social Psychology, 40,* 235–256.

Scherer, K. (1999). Appraisal theory. In T. Dalgleish & T. Power (Eds.), *Handbook of cognition and emotion* (pp. 637–659). West Sussex, England: Wiley.

Schul, Y., & Schiff, M. (1995). On the costs and benefits of ignorance: How performance satisfaction is affected by knowing the standard prior to performance. *Personality and Social Psychology Bulletin, 21,* 491–501.

*Stiensmeier-Pelster, J. (1989). Attributional style and depressive mood reactions. *Journal of Personality, 57,* 581–599.

Ward, L. G., Friedlander, M. L., & Silverman, W. K. (1987). Children's depressive symptoms, negative self-statements, and causal attributions for success and failure. *Cognitive Therapy and Research, 11,* 215–227.

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.

Weiner, B. (1986). An attributional theory of achievement motivation and emotion. *Psychological Review, 92,* 548–573.

Westermann, R., Spies, K., Stahl, G., & Hesse, F. (1996). Relative effectiveness and validity of mood induction procedures: A meta-analysis. *European Journal of Social Psychology, 26,* 557–580.

*Whitley, B. E. (1986). The relationship of informational attributions to affective response to success and failure. *The Journal of Social Psychology, 120,* 453–457.

*Wierzbicki, M., & Westerholm, P. (1994). A comparison of cognitive and behavioral inductions of negative mood. *Journal of Psychology, 128,* 651–657.

*Wilson, G. V., & Kerr, J. H. (1999). Affective responses to success and failure: A study of winning and losing in competitive rugby. *Personality and Individual Differences, 27,* 85–99.

(Appendix follows)

Appendix

Reviewed Studies Included in the Meta-Analysis

Study	*n*	Task	Affect measure	Manipulation	Design	$R_{negative}$	$R_{positive}$	$R_{bipolar}$	$R_{average}$
Baucom & Aiken (1981)	56	Cognitive	Other public	Feedback	Experimental	.34			.34
Brown & Dutton (1995; Exp. 1)	172	Cognitive	PANAS	Difficulty	Experimental			.48	.48
Brown & Marshall (2001)	72	Cognitive	Own	Difficulty	Experimental			.24	.24
Chartier & Ranieri (1989)	60	Cognitive	Other public	Feedback	Experimental	.33			.33
Cohen, Pane, & Smith (1997; Study 2)	157	Other	MAACL	Feedback	Experimental	.00	.16		.08
Egloff (1998)	80	Cognitive	PANAS	Difficulty	Experimental	.46	.50		.48
Egloff & Krohne (1996)	100	Cognitive	PANAS	Difficulty	Experimental	.15	.12		.13
Feick & Rhodewalt (1997)	121	Return of exam	Other public	Feedback	Experimental			.30	.30
Forgas, Bower, & Moylan (1990; Study 1)	64	Cognitive	Own	Feedback	Experimental	.62	.47		.55
Gendolla (1997; Study 5)	60	Cognitive	Own	Difficulty	Experimental				.37
Idosn, Liberman, & Higgins (2000; Study 3)	45	Cognitive	Own	Feedback	Experimental			.74	.74
Ingram (1984)	32	Bogus IT	MAACL	Feedback	Experimental	.69			.69
Ingram et al. (1992; Study 1)	58	Bogus IT	MAACL	Feedback	Experimental	.26			.26
Ingram, Smith, & Brehm (1983)	47	Social perception	MAACL	Feedback	Experimental	.36			.36
Krohne, Pieper, Knoll, & Breimer (2002)	82	Bogus IT	PANAS	Difficulty	Experimental	.35	.47		.41
Larsen & Ketelaar (1989)	67	Bogus IT	Own	Feedback	Experimental			.18	.18
Linville (1985; Study 1)	59	Bogus IT	Other public	Feedback	Experimental			.00	.00
McFarland & Buehler (1997; Study 1)	58	Social perception	Own	Feedback	Experimental			.59	.59
McFarland & Ross (1982)	48	Social perception	Own	Feedback	Experimental		.50		.50
Mehlman & Snyder (1985)	96	Cognitive	MAACL	Feedback	Experimental	.45			.45
Neumann, Seibt, & Strack (2001; Study 1)	40	Other	Own	Feedback	Experimental	.00	.14		.07
Parrott & Sabini (1990; Study 1)	124	Return of exam	Own	Quasi-exp.	Quasi-exp.	.24	.28	.18	.23
Rhodewalt & Morf (1998; Study 1)	128	Bogus IT	Other public	Feedback	Experimental	.30	.30		.30
Rhodewalt & Morf (1998; Study 2)	130	Bogus IT	Other public	Feedback	Experimental	.46	.46		.46
Sanna, Meier, & Wegner (2001; Study 2)	72	Return of exam	Own	Quasi-exp.	Quasi-exp.			.41	.41
Schul & Schiff (1995; Study 1)	91	Social perception	Own	Feedback	Experimental			.74	.74
Schul & Schiff (1995; Study 2)	166	Other	Own	Feedback	Experimental			.26	.26
Stiensmeier-Pelster (1989; Study 2)	46	Bogus IT	Other public	Difficulty	Experimental			.28	.28
Ward, Friedlander, & Silverman (1987)	91	Cognitive	Own	Difficulty	Experimental			.25	.25
Whitley (1986)	60	Cognitive	Own	Difficulty	Experimental			.64	.64
Wierzbicki & Westerholm (1994)	40	Cognitive	MAACL	Feedback	Experimental	.39			.39
Wilson & Kerr (1999)	18	Other	Other public	Quasi-exp.	Quasi-exp.	.46	.24		.35

Note. $R_{negative}$, $R_{positive}$, $R_{bipolar}$, and $R_{average}$ = effect sizes for negative affect, positive affect, bipolar affectivity, and average affectivity, respectively; exp. = experiment; PANAS = Positive and Negative Affect Schedule; MAACL = Multiple Affect Adjective Checklist; IT = intelligence.

Received July 2, 2003
Revision received October 28, 2003
Accepted October 28, 2003 ■