

## BRIEF REPORT

### Choice of Difficult Tasks as a Strategy of Compensating for Identity-Relevant Failure

Oliver C. Schultheiss and Joachim C. Brunstein

*University of Erlangen, Erlangen, Germany, and University of Potsdam, Potsdam, Germany*

This study examined the hypothesis that in the realm of self-definitional commitments, individuals try to compensate for failure experiences by increasing their aspirational standards with respect to further self-defining task achievements. Students committed to the self-definition of becoming a physician first received either failure feedback or no feedback on a number-tracking test and were then asked to select from a visual-search test a number of tasks that differed with respect to their difficulty. When the skills involved in these tests were described as relevant to the profession of a physician, students pretreated with failure chose more difficult tasks than their no-feedback counterparts. No such effect emerged when the two tests were described as nonrelevant to students' professional self-definition. Yet, in this nonrelevant condition, task choice was predicted by a measure of self-handicapping.

© 2000 Academic Press

Brunstein (1995, 1999; Brunstein & Gollwitzer, 1996; Brunstein & Olschner, 1996) reported evidence suggesting that self-defining goals play an important role in how individuals cope with failure experiences. A self-definitional commitment can be conceived of as a higher order aspiraton that elicits a persistent striving to acquire a particular identity, such as becoming a competent musician, teacher, or parent (cf. Gollwitzer, 1987). If failure occurs on the way toward realizing such a self-definition (or identity goal), the person will be challenged and motivated to compensate for the experienced self-definitional shortcoming rather than give up and disengage from the identity goal (see Carver & Scheier, 1990, for a similar conception regarding the special role of personally meaningful higher level goals in self-regulatory responses to failure). Failure in a self-defining task will therefore spur extra efforts whenever a person has access to further self-defining tasks.

Address correspondence and reprint requests to Oliver C. Schultheiss, Department of Psychology, University of Michigan, Ann Arbor, MI 48109-1109. Fax: (734) 647-9440. E-mail: [oschult@umich.edu](mailto:oschult@umich.edu).



In contrast, if a person fails to achieve a non-self-defining task, his or her motivation to perform similar tasks, which through the prior failure experience have become aversive, will be reduced thereafter.

In keeping with these hypotheses, Brunstein and Gollwitzer (1996) found that among students committed to professional identity goals, such as becoming a physician or computer scientist, failure feedback in an (initial) identity-relevant task fueled performance increments on a subsequent task described as relevant to the same self-definition. Brunstein and Olschner (1996) replicated this finding among psychology undergraduates who felt committed to the goal of becoming a clinical psychologist and found in addition that students who did not feel committed to this professional goal, but nevertheless received negative feedback as to their performance on a number of psychotherapy tasks, displayed performance decrements on a subsequent diagnosis task described as being relevant to the profession of a clinical psychologist.

A limitation of these studies is that the motivational effects of failure within and outside the realm of self-definitional pursuits were solely assessed through performance measures derived from experimenter-assigned tasks. Thus, the fact that individuals who had experienced an identity-relevant failure subsequently responded with increased effort to a task of fixed difficulty administered by the experimenter tells us nothing about what level of difficulty individuals would select for themselves on the next task if given the choice. It cannot be ruled out a priori that failure at an identity-relevant task may lead to the choice of easy tasks if an alternative identity-relevant opportunity arises. If the individual is more interested in a quick and surefire demonstration of her or his mastery of identity-relevant tasks than in ascertaining that she or he has really got what it takes to realize a self-defining goal, then this would be a suitable strategy.

However, we believe that for individuals who have failed at an identity-relevant task it is not success for the sake of success, regardless of the way it was achieved, that counts but proving to themselves that they actually do possess the competencies necessary to realize the self-defining goal. Therefore, we hypothesized that they should try to compensate a previous identity-relevant failure experience by raising their performance standards with respect to a further identity-relevant task. This postulate complies to Jucknat's (1938) observation that individuals raise, rather than lower, their level of aspiration after failure feedback, provided that the experienced failure in some way impinges on a goal that is closely tied to the person's self-regard. Our notion is also consistent with Locke and Latham's (1990) postulate that difficult tasks, as compared to easy tasks, commonly elicit more effort and thereby stimulate task performance. From this perspective, the choice of difficult tasks can be regarded as a self-regulatory strategy employed by individuals who want to make up for a previous failure in the pursuit of a self-definition they seek to attain. In contrast, if failure is unrelated to a

self-definition, it should be likely to bias an individual's preference in the direction of easy tasks because such tasks may protect the individual against further disappointments. Therefore, in the present research we modified the paradigm used in earlier studies (cf. Brunstein, 1999) and gave participants the opportunity to choose between different difficulty levels on a second task after they had received failure feedback or no feedback on a first task.

We also addressed the possibility that task choices after failure feedback may reflect to a considerable extent interindividual differences in self-handicapping, defined as an individual's desire to proactively protect her or his self-esteem against potential disappointments in evaluative environments (Higgins, Snyder, & Berglas, 1990). One strategy self-handicappers may use to achieve this defensive, ego-protective goal is the choice of very easy or very difficult tasks (Greenberg, 1985; Rhodewalt, 1990). Whereas failure is unlikely to occur in easy tasks, failure in achieving difficult tasks can easily be attributed to the extremity of the choice one has made rather than to a lack of ability or task-related skills. We therefore decided to explore whether interindividual differences in self-handicapping would moderate the postulated effects of failure experiences on individuals' subsequent task preferences within or outside the context of a self-definitional pursuit.

## METHOD

### *Participants*

Participants were 60 students (24 women and 36 men) enrolled at the medical school of the University of Erlangen (Erlangen, Germany). The average age of the sample was 22.6 years ( $SD = 2.54$ ). To obtain a sample of participants truly committed to the goal of becoming a physician, all potential participants were presented the following item: "I cannot imagine a profession more fulfilling than that of a physician." Students who unequivocally endorsed this item were admitted to the study.

### *Design and Procedure*

The study's design was a 2 (task condition: identity-relevant vs nonrelevant)  $\times$  2 (feedback type: failure vs no feedback) factorial. Participants were run individually and randomly assigned to the experimental conditions. Upon arriving, they were first administered a German translation of Rhodewalt's (1990) 25-item Self-Handicapping Scale (SHS). The participants were told that the experiment would require them to work on two types of mental-concentration tests. The experimenter introduced these tests as two different, but complementary measures of a person's ability to master attention-demanding tasks. Using instructions similar to those reported by Brunstein and Gollwitzer (1996, p. 398), the skills involved in these tests were described as being characteristic of highly qualified physicians for half of the participants (identity-relevant task condition). The other half was told that the two tasks had been designed to test the ability to concentrate on performing a task among aircraft pilots (nonrelevant task condition). Data from various samples would be needed to determine the reliability of the two tests.

After that, the procedure was subdivided into a treatment phase and a subsequent test phase. In the treatment phase, the students were administered four test forms adopted from Oswald and Roth's (1987) Number Tracking Test. This test requires the participant to connect as fast

as possible a series of consecutive numbers (1-2-3-4- . . . ) which are arranged together with distractor numbers in a rectangular matrix. In the present experiment, the students were told that they would have 45 s to work on each form. In the identity-relevant task condition, failure feedback was prearranged in the following fashion. Each matrix contained one particular number that was highlighted and printed bold. Failure-feedback students were led to believe that this number would reflect the average performance score medical students had achieved on the respective task in previous studies. In each of the four trials, failure students were stopped about 10 numbers before they could reach the highlighted number. In this way, they were given the impression that their own performance score was consistently worse than the average performance score achieved by medical students. In the nonrelevant task condition, failure feedback was prearranged in the same fashion. However, students exposed to failure were told that the highlighted number would reflect the average performance score achieved by student pilots. In both task conditions, no-feedback participants were stopped at the same numbers as their counterparts in the failure-feedback conditions. Yet their test forms did not contain any information concerning the average performance of medical students or student pilots. Accordingly, no-feedback students received no information about the quality of their task performance.

Based on pilot work, the allegedly allotted 45 s per form as well as the numbers participants were interrupted at during each trial were chosen so as to make it impossible for them notice that they were not actually stopped after 45 s. Consequently, none of the failure- or no-feedback-condition participants showed any suspicion that the experimenter did not in fact stop them after 45 s. To test the effectiveness of the failure manipulation, we administered two bipolar 7-point scales immediately after participants had completed the number tracking forms. The two scales concerned feelings of *satisfaction* (7) versus *dissatisfaction* (1) and *enjoyment* (7) versus *dejection* (1).

In the subsequent test phase, participants were presented a visual-search test that we had designed for the purpose of this experiment. The experimenter introduced this test by presenting to the participants a sheet consisting of a series of 19 rows of geometric symbols on which they were instructed to cross out a target symbol. The rows differed with respect to the number of symbols a participant had to check for the target stimulus within a fixed interval of 10 s and were ordered according to length. The shortest row contained 14 symbols and the longest 32 symbols. Thus, task difficulty was varied as a function of the length of the 19 rows of symbols. In addition, for each row of symbols, we presented the percentage of either medical students (relevant task condition) or student pilots (nonrelevant task condition) who had allegedly succeeded in marking all target symbols within the allotted time. This percentage was varied in steps of 5% between 95% (least difficult task) and 5% (most difficult task). The students were asked to choose from the rows of symbols presented to them 15 tasks that could, but did not need to, differ with respect to their level of difficulty (i.e., the participants could select more than one task from each level of task difficulty). The experimenter continued to say that the greater a task's difficulty (i.e., the length of a row of symbols), the higher the score would be one could earn on the respective trial (ranging from 1 point = *least difficult type of task* to 19 points = *most difficult type of task*). Yet, if a participant would commit only one error while working on a given task (i.e., omit a target or cross out a nontarget) or fail to complete a selected row within the prespecified interval, he or she would receive no point for the respective task. After students had selected 15 tasks, the experimenter interrupted the experiment and fully debriefed the participants about the purpose of the study.

To obtain an overall measure reflecting students' task choice, we summed each difficulty level (i.e., levels ranging from 1 to 19) times the number of tasks a participant had selected at this difficulty level and divided the result by 15. The resulting variable (*mean level of difficulty*) could range from 1 (all rows were chosen from the least difficult type of task) to 19 (all rows were chosen from the most difficult type of task). In the present study participants'

task choices had the form of a normal distribution, with the modal average number of rows selected at the 50% success level ( $M = 2.08$ ) and decreasing numbers of selected rows in the direction both of easier and more difficult row types. Both the easiest (95% chance of success) and the most difficult (5% chance of success) were chosen only very infrequently ( $M_s = 0.65$  and  $0.17$ , respectively). Hence, by tending to choose moderate-difficulty levels, participants tried to strike a balance between optimizing their anticipated score and keeping the risk of failing under control.

### *Pilot Work*

Medical students ( $N = 14$ ) participating in a pilot study in the no-feedback condition indicated on 7-point scales that they had a slightly lower expectancy of doing well on the visual-search test ( $M = 3.36$ ,  $SD = 1.55$ ) than on the number-tracking test ( $M = 3.79$ ,  $SD = 1.42$ ),  $p > .10$ . These findings show that the visual-search test was judged to be at least as difficult as the preceding number-tracking task.

## RESULTS

### *Analysis of the SHS*

We first examined the reliability of the German translation of Rhodewalt's (1990) SHS. As assessed by Cronbach's  $\alpha$ , the internal consistency of the 25-item scale was .56. To obtain a more homogenous scale, we abridged the full SHS to form a 14-item scale, developed and validated by Zuckerman and Kieffer (1998). In our data, the internal reliability of this 14-item SHS was .64 (see Zuckerman & Kieffer, 1998, p. 1628, for the items included in this scale). Although this reliability coefficient was still lower than the reliability coefficients Zuckerman and Kieffer reported for two independent samples of college students ( $\alpha = .74$  and  $.76$ , respectively), it is quite comparable to the reliabilities reported by other researchers for related short-forms of the SHS (for instance, Deppe & Harackiewicz, 1996, reported the internal reliability of a 10-item version of the SHS to be .62). Thus, all subsequent analyses are based on Zuckerman and Kieffer's 14-item version of the SHS. (Noteworthy, however, is that analyses using Rhodewalt's full SHS yielded similar results.)

### *Manipulation Check*

As compared to no-feedback control students, failure-feedback students felt less satisfied ( $M = 3.23$  vs  $4.03$ ),  $t(58) = 3.09$ , and joyful ( $M = 3.96$  vs  $4.66$ ),  $t(58) = 2.99$ , after they had completed the number-tracking tasks ( $ps < .01$ ). No other variable had a reliable impact on these mood ratings.

### *Predicting Task Choice*

To analyze students' task choice, we used the following regression approach: Mean level of task difficulty was used as the dependent variable. As to the predicting variables, the SHS was treated as a continuous variable, whereas dummy codes were used to examine the effects of the experimental

conditions (task condition, 1 = *identity-relevant condition*, 0 = *nonrelevant condition*; feedback type, 1 = *failure feedback*, 0 = *no feedback*). First-order terms of task condition, feedback type, and self-handicapping were entered first into the regression equation, followed by the three two-way interaction terms computed for these variables. Finally, the three-way interaction was tested for significance. To facilitate the interpretation of regression coefficients obtained from hierarchical regression analyses, we followed Aiken and West's (1991) recommendation and converted values of the SHS to  $z$  scores before we computed the interactions.

Together, the predictors accounted for a significant 25.3% of the variance in the mean level of task difficulty,  $F(7, 52) = 9.22, p < .01$ . Regression slopes indicated that none of the first-order terms were significant ( $ps > .10$ ). As to the two-way interaction terms, the interactive effect of task condition and feedback type turned out to be a significant predictor of task difficulty, uniquely accounting for 6.8% of the variance in the dependent measure,  $b = .52, t(53) = 2.03, p < .05$ . Neither the Task Condition  $\times$  SHS interaction nor the Feedback Type  $\times$  SHS interaction was significant ( $ps > .10$ ). Yet the three-way interaction between task condition, feedback type, and self-handicapping significantly added to the predictive power of the equation, uniquely accounting for 13.2% of the variance in students' task choice,  $b = .79, t(52) = 3.03, p < .01$ .

Because the significant Task Condition  $\times$  Feedback Type interaction was directly relevant to the testing of our major hypothesis, we first analyzed the form of this two-way interaction before we examined the nature of the significant Task Condition  $\times$  Feedback Type  $\times$  SHS interaction. Pairwise comparisons between means of the four experimental groups yielded the following results (cf. Table 1): Students exposed to failure in the identity-relevant task condition selected more difficult test tasks than both no-feed-

TABLE 1  
Difficulty of Chosen Tasks as a Function of Task Condition  
and Feedback Type

Feedback type	Task condition			
	Identity-relevant		Identity-nonrelevant	
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
Failure	8.85	1.72	7.08	2.25
No-feedback	7.69	1.66	7.98	1.80

*Note.*  $N = 15$  in each cell; higher scores reflect choice of more difficult tasks.

back control students assigned to the same (relevant) task condition,  $t(28) = 1.88$ ,  $p = .07$ , and failure-feedback students assigned to the nonrelevant task condition,  $t(28) = 2.42$ ,  $p < .05$ . In the nonrelevant task condition, failure-feedback students chose slightly easier test tasks than their no-feedback counterparts, but this difference did not reach statistical significance ( $p > .10$ ).

As to the significant three-way interaction, subsidiary regression analyses revealed the following results: For students assigned to the identity-relevant task condition, neither the SHS first-order term nor the Feedback Type  $\times$  SHS interaction term had a reliable effect on students' task choice ( $ps > .10$ ). In this (identity-relevant) condition, the correlation between the SHS and the mean level of task difficulty was  $r = .06$ . In contrast, for students assigned to the nonrelevant task condition, the Feedback Type  $\times$  SHS interaction was significant,  $t(26) = -2.51$ ,  $p < .05$ . In this (nonrelevant) condition, dispositional self-handicapping was negatively related to task difficulty among failure-feedback students,  $r = -.53$ ,  $p < .05$ , but statistically independent of task difficulty among no-feedback control students,  $r = .28$ ,  $p > .30$ .

## DISCUSSION

The results of the present study provide strong support for our main hypothesis: In the identity-relevant task condition, individuals pretreated with failure as to their performance on the number tracking test chose the most difficult tasks of all participants from the subsequent visual search test. The compensatory nature of their task choice stands out most clearly if one considers that (a) control participants who had also received identity-relevant instructions but no failure feedback picked less difficult test tasks and (b) participants who had received failure feedback that was unrelated to their professional ambition chose the easiest test tasks of all participants. These results complement and extend earlier findings reported by Brunstein (1999) which demonstrate that only within the domain of self-definitional pursuits will individuals who have failed at a previous task show superior performance on a new task that provides an opportunity to compensate the identity-relevant setback.

Moreover, we found that interindividual differences in self-handicapping had no significant impact on the task choice of students exposed to identity-relevant failure feedback. In contrast, self-handicapping clearly predicted participants' task choice in the nonrelevant failure condition inasmuch as students high in self-handicapping chose easier test tasks than their low self-handicapping counterparts. From this finding, it is tempting to conclude that in the realm of self-definitional pursuits (as compared to non-self-definitional pursuits), the desire to compensate for an experienced self-definitional short-

coming overrides the tendency to use defensive strategies in order to protect oneself from further disappointments. Further research is needed to cross-validate this finding.

Finally, although at first glance the choice of difficult tasks observed in failure-feedback participants in the identity-relevant task condition may seem reminiscent of a symbolic self-completion effect as described in Wicklund and Gollwitzer's (1982) work, we believe that this account falls short in explaining the present results. In a nutshell, symbolic self-completion theory predicts that individuals who feel that they are deficient in a self-defining area (e.g., through the experience of failure on an identity-relevant task) seek to reinstate a sense of completeness vis-à-vis their social environment by publicly ascribing to themselves characteristics that symbolize the aspired-to self-definition. Thus, one could argue that failure-feedback participants used the choice of difficult identity-relevant tasks as a means to appear competent within the domain of their self-definition. However, such a strategy would seem short-sighted and could easily backfire as soon as the alleged competence is put to the test, and in the present research, participants fully expected to work on the tasks they had chosen for themselves. Therefore, we suggest that the choice of difficult tasks did not serve a mere symbolic function but represented a self-regulatory attempt at actual effortful compensation of a previous identity-relevant failure. In a similar vein, Wurf and Markus (1991) have pointed out that symbolic validation of the self (e.g., choice of difficult goals as a symbol of one's competence) may be less effective with regard to realizing a self-defining goal than actual achievements (e.g., choosing difficult goals as a means to achieve superior performance). However, it might be fruitful to determine in future research whether individuals who have recently experienced failure within a self-defining context may resort to symbolic self-completion if the way to effortful compensation is blocked.

Taken together, the present findings suggest that in the realm of self-definitional pursuits, individuals who are falling short do not lower their sights, but adopt even more challenging goals to bolster their far-reaching aspirations. Although the strategic choice of more difficult tasks may evoke new risks of failing, it clearly reflects a person's desire to compensate for the prior failure experience. Moreover, as noted above, a major finding in goal-setting research is that high goal difficulty fosters task performance as long as a person remains committed to the difficult goal (Locke & Latham, 1990). In keeping with this proposition, Gendolla (1998) reported that in the domain of identity commitments, difficult tasks, as compared to easy tasks, elicit much more effort as reflected in increases in heart rate and systolic blood pressure. Interestingly, there was no such relationship for non-self-defining tasks. One might therefore speculate that setting more difficult goals after an identity-relevant failure will be likely to trigger an increase in the expendi-

ture of effort that in turn translates into better task performance. Further research is needed to scrutinize this mediational model, though.

## REFERENCES

- Brunstein, J. C. (1995). *Motivation nach Misserfolg* [Motivation following failure]. Göttingen, Germany: Hogrefe.
- Brunstein, J. C. (in press). Motivation and performance following failure: The effortful pursuit of self-defining goals. *Applied Psychology: An International Review*.
- Brunstein, J. C., & Gollwitzer, P. M. (1996). Effects of failure on subsequent performance: The importance of self-defining goals. *Journal of Personality and Social Psychology*, **70**, 395–407.
- Brunstein, J. C., & Olschner, S. (1996). Goal commitment, performance feedback, and motivation to engage in substitute action. In T. Gjesme & R. Nygard (Eds.), *Advances in motivation* (pp. 23–34). Oslo, Norway: Scandinavian Univ. Press.
- Carver, C. S., & Scheier, M. F. (1990). Principles of self-regulation: Action and emotion. In E. T. Higgins & R. M. Sorrentino (Eds.), *Handbook of motivation and cognition: Foundations of social behavior* (Vol. 2, pp. 3–52). New York: Guilford.
- Deppe, R. K., & Harackiewicz, J. M. (1996). Self-handicapping and intrinsic motivation: Buffering intrinsic motivation from the threat of failure. *Journal of Personality and Social Psychology*, **70**, 868–876.
- Gendolla, G. H. E. (1998). Effort as assessed by motivational arousal in identity-relevant tasks. *Basic and Applied Social Psychology*, **20**, 111–121.
- Gollwitzer, P. M. (1987). The implementation of identity intentions: A motivational-volitional perspective on symbolic self-completion. In F. Halisch & J. Kuhl (Eds.), *Motivation, intention, and volition* (pp. 349–369). Berlin, Germany: Springer-Verlag.
- Greenberg, J. (1985). Unattainable goal choice as a self-handicapping strategy. *Journal of Applied Social Psychology*, **15**, 140–152.
- Higgins, R. L., Snyder, C. R., & Berglas, S. (Eds.) (1990). *Self-handicapping: The paradox that isn't*. New York: Plenum.
- Jucknat, M. (1938). Leistung, Anspruchsniveau und Selbstbewusstsein [Performance, aspiration level, and self-confidence]. *Psychologische Forschung*, **22**, 89–179.
- Locke, E. A., & Latham, G. P. (1990). *A theory of goal setting & task performance*. Englewood Cliffs, NJ: Prentice Hall.
- Oswald, W. D., & Roth, E. (1987). *Der Zahlen-Verbindungs-Test (ZVT)* [The number-tracking test] (2nd ed.). Göttingen, Germany: Hogrefe.
- Rhodewalt, F. (1990). Self-handicappers: Individual differences in the preference for anticipatory, self-protective acts. In R. L. Higgins, C. R. Snyder, & S. Berglas (Eds.), *Self-handicapping: The paradox that isn't* (pp. 69–106). New York: Plenum.
- Wicklund, R. A., & Gollwitzer, P. M. (1982). *Symbolic self-completion*. Hillsdale, NJ: Erlbaum.
- Wurf, E., & Markus, H. (1991). Possible selves and the psychology of personal growth. In R. Hogan (Series Ed.) & D. Ozer, J. M. Healy, & A. Stewart (Vol. Eds.), *Perspectives in personality: Vol. 3A. Self and emotion* (pp. 39–62). London: Kingsley.
- Zuckerman, M., & Kieffer, S. C. (1998). Consequences of self-handicapping: Effects on coping, academic performance, and adjustment. *Journal of Personality and Social Psychology*, **74**, 1619–1628.