Eyes on the Prize or Nose to the Grindstone? The Effects of Level of Goal Evaluation on Mood and Motivation

Linda Houser-Marko
University of Illinois–Chicago

Kennon M. Sheldon
University of Missouri

These studies tested the hypothesis that evaluating goal feedback in terms of a primary, longer term goal can be risky for future motivation. Study 1 was a 2 x 2 experiment in which framing level (primary goals/subgoal) and feedback valence (success/failure) were manipulated for participants during a verbal skills task. In the primary goal failure condition, there was increased negative mood and decreased positive mood and expectancy for subsequent trials, even while controlling for goal difficulty and importance. Study 2 was an 8-week study throughout which participants were asked to evaluate their progress regarding a primary goal (class grade goal) or subgoal (weekly study hours goal), and success or failure varied naturally. When progress was lacking, participants in the primary goal condition experienced the largest decreases in mood and expectancy. These results suggest that it is optimal to evaluate goal progress at the lower, subgoal level, particularly after failure feedback.

Keywords: long-term personal goals; short-term personal goals

Folk psychology is full of contradictory adages, each adage sounding plausible by itself, but both of which cannot be true. One such contradiction can be seen in the area of goal striving. Should we “keep our eyes on the prize” or “keep our noses to the grindstone”? The first saying advocates always retaining in mind the ultimate goal, the thing we are really trying to get, as a way of revitalizing and recharging our energies when things go poorly (King, 2001). The second saying advocates stoically plugging away at more concrete or proximal goals, avoiding the temptation to fantasize about the prize prematurely (Oettingen, 1996). What are the costs and benefits of each type of approach? These studies will shed light further on this question with an emphasis on the process of personal goal pursuit.

Levels of Goals

Personal goals often exist in hierarchical relationship (Carver & Scheier, 1990; Kruglanski et al., 2002), in which one goal serves another. Primary goals are ends in themselves and represent the true object of striving. Primary goals are dependent on the successful completion of related subgoals. Subgoals are not ends but rather exist only because of the primary goals, as means to the end (Kruglanski et al., 2002). Thus, subgoals are dependent on primary goals for their meaning and impetus. Primary goals and subgoals also differ in their degree of specificity (abstract vs. concrete), time frame (long term vs. short term), and self-relevance (direct vs. indirect; Bandura, 1988).

The Evaluation Process

In addition to being hierarchically and structurally related to each other, primary goals and subgoals may...
be thought of as functionally linked in feedback loop systems. Feedback loop (or control theory) models are useful in discussing the process of goal pursuit and evaluation and are well established in the motivation literature (Carver & Scheier, 1990, 1998; Miller, Galanter, & Pribram, 1960; Powers, 1973; Scheier & Carver, 1988). A feedback loop system begins with the current state that is compared to a desired state or goal, such as to read a chapter in a history textbook. When a discrepancy is detected, a behavior is enacted (the student reads the chapter), and then the result feeds back to the beginning of the system. The new state (that he or she read the history chapter) is compared to the desired state again to determine if there is a discrepancy and to decide on further action. The lack of a perceived discrepancy tells the student he or she can move on to the next assignment or subgoal within the overall plan. The current article asks, What is the optimal level of identification (Vallacher & Wegner, 1987) to use when the discrepancy process is initiated by failure feedback to maintain one’s future motivation?

The Costs and Benefits of Primary Goal Construal

Being ends in themselves, primary goals tend to represent what the individual really cares about. Primary goals tend to be more relevant and meaningful for identity and the self-concept (Brunstein, 2000) because they occupy a higher level of the action system that is hierarchically “closer” to the global self (Carver & Scheier, 1998). Certainly, giving attention to one’s higher level goals can have benefits. For example, counseling often tries to help people discover meaningful goals to pursue (Michalak, Klappheck, & Kosfelder, 2004). Encouraging people to write about their “best possible future self” has been shown in experimental research to convey physical health benefits, enhance people’s sense of meaning and purpose in life, and increase their subjective well-being (King, 2001). Thus, it seems important to aspire to future “prizes” in life.

At the same time, desiring such prizes can make people vulnerable. Because high-level goals are so important and self-defining, feeling that one has failed at such goals may be extremely dispiriting (Wrosch, Scheier, Miller, Schulz, & Carver, 2003). This problem may be especially pernicious because high-level goals are long term and require many steps to complete. Thus, there are many occasions along the way to construe momentary difficulties or setbacks in terms of their implications for the longer term goal. When such implications become salient, a person’s emotional reaction to momentary failure might be amplified, causing his or her expectancies regarding the primary goal to decline. This might lead to a reduction in commitment toward that primary goal, and also a reduction in commitment to the linked subgoals, and then to further performance difficulties or failures. In other words, evaluating at the primary goal level too early or after a subgoal failure might affect pursuit of the goals via discouragement regarding the primary goal. A downward spiral might ensue in which both the initial primary goal and the subgoals that serve it are deenergized and finally abandoned.

For example, consider a person with the goal of running a marathon 6 months hence (the primary goal). Suppose that his or her subgoal during a particular week is to go running three times, for 10 miles each time. Further suppose that a family emergency and a chaotic workweek prevent the person from running at all during the week. We suggest that it would be more constructive for this person to restrict his or her interpretation of this situation, construing it as a failure this week and one that does not severely impact the primary goal. In this case, the person’s motivation to pursue the next weekly goal will be unflagging. However, if he or she instead thinks, “I’m never going to reach my marathon goal,” then his or her motivation for the subsequent week may wane.

The Action Identification and Control Theory Perspectives

Often, life pursuits and personal projects involve a trade-off between doing what is meaningful at the higher level and what is manageable at the lower level (Little, 1989). But what determines our focus on meaning versus managing? Action identification theory (AIT) directly addresses this (Vallacher & Wegner, 1987, 1989) by noting that many actions can be described in either low-level (“I’m threading the needle”) or high-level terms (“I’m sewing the costume”) and further noting that most people prefer more meaningful higher level identifications.

What happens when difficulties arise? In this case, AIT asserts that people will lower their level of attention, focusing more concretely on momentary problems to be solved when the action is personally difficult or when the context demands. For instance, if a person is asked to eat Cheetos with chopsticks, his or her optimal level of identification would be at the lower level of grasping the Cheeto and bringing it to his or her mouth, instead of higher level identifications such as having a snack or enjoying the taste (Vallacher & Wegner, 1987). Thus, the optimal approach from the AIT perspective is to frame one’s goal pursuits at the higher, primary goal level when things are going well and frame one’s goal pursuits at the subgoal level when the situation demands it; to properly trade off the “meaning” and the “managing.”
One study showed that when a person assigned to give a speech was led to believe that the audience would be easy or difficult to persuade and had either a high level of identification (no manipulation) or a low level (an action disruption made them focus more on low-level actions), the person's performance on the speech differed significantly (Vallacher, Wegner, & Somoza, 1989). In essence, the low-level identification was helpful for rated speech performance when participants thought the audience would be hard to persuade, but high-level identification led to better performance when the audience was thought to be easier to persuade.

The current research applies and extends AIT by considering the consequences of receiving feedback or self-evaluating one's progress, framed either in concrete subgoal terms or more abstract primary goal terms. We experimentally tested the hypothesis that keeping one's nose to the grindstone has better motivational and mood consequences, at least in the case of momentary failure or difficulties. Although this hypothesis can be derived from AIT, it has not been specifically tested by prior AIT research that has not manipulated success and failure feedback, has not studied hierarchically linked personal goals, and has not studied the effect of AIT processes on people's motivation for future tasks. For example, in the Vallacher et al. (1989) experiment described previously, the dependent variable was current performance, not future motivation; the manipulated variable was task difficulty versus ease, rather than success versus failure feedback following the task; and the task (speech giving) was not necessarily important to the participant.

Our hypothesis can also be derived from Carver and Scheier's (1990, 1998) control theory of action and self-regulation, which emphasizes the necessity of completing a series of linked subgoals to approach a longer term, primary goal (i.e., completing a plan; Miller et al., 1960). When discrepancies are detected at the subgoal level, it is only logical that people should narrow their attention to that level, providing themselves with the most relevant comparison standard for use in self-regulation and discrepancy reduction. Thus, as does AIT, control theory suggests that one should keep one's nose to the grindstone, until such a time as that subgoal is completed—only then would it be appropriate to raise one's gaze, to identify and begin pursuing the next subgoal. The erstwhile marathon runner is better off thinking about how to meet the weekly goal next week rather than thinking about the longer term goal next year.

Other research also supports the idea that premature high-level construals can be problematic. Emmons (1992) defined concrete (low-level) versus abstract (high-level) goal focus as an individual difference, which he measured by classifying participants’ set of personal strivings as either concrete or abstract, as a whole. He showed that high-level strivers were more prone to experience depression, arguing that the depression resulted because high-level strivers are not as clear on how to achieve their more abstract goals and because they suffer more from any perceived failures due to the greater meaningfulness and self-relevance of their goal-set as a whole.

Attribution theory also supports the “premature high-level construal” idea. Generalization of failure occurs when a person makes global attributions for specific failures (Carver & Scheier, 1998) and has been shown to be associated with depression (Carver, 1998), reduced academic performance (Peterson & Barrett, 1987), and maladaptive coping styles and off-task cognitions (Mikulincer, 1989). Notably, the attribution literature has considered the tendency toward generalization primarily as an individual difference variable, just as Emmons (1992) considered abstract goal striving as an individual difference. However, in the current research we are studying primary versus secondary goal framing as a contextual variable, which can be influenced by the framing of specific information. Notably, AIT researchers have examined action identification both as an individual difference and as a manipulated variable. However, AIT’s individual difference variable focuses on the traitlike construal of common actions rather than focusing on personal goals per se.

The Costs and Benefits of Subgoal Construal

Although thinking of one’s actions and outcomes in terms of concrete subgoals may be less meaningful or interesting to participants, there is reason to believe it is more optimal, at least when the subgoal is difficult or uncertain. For example, Bandura and Schunk (1981) and Bandura (1988) showed that focusing participants’ attention on a proximal or short-term goal rather than a distal or longer term goal is associated with greater performance and associated self-efficacy, and Manderlink and Harackiewicz (1984) showed that such a focus is correlated with higher attainment expectations. Thus, the research comparing performance outcomes as a function of focusing on proximal and distal goals indicates a benefit to keeping one’s nose to the grindstone in the context of the workplace (Latham & Sejts, 1999).

We suggest that negative evaluations regarding short-term goals are likely to have less impact on people’s mood because such goals are less directly meaningful and self-defining and because short-term failures can be more quickly and easily rectified. In this restricted case, negative mood serves properly as a signal for self-correction (Carver & Scheier, 1990) but does not spread to
the domain of the more global self-construal. Also, lower level negative evaluations are likely to have an eventual positive impact on people’s mood because low-level evaluations keep people’s attention focused on the concrete steps that need to be taken, leading to adaptive behavior in the end. Again, because primary goals are longer term structures, logically, progress toward them should only be assessed periodically. In short, keeping your nose to the grindstone may help a person to dodge the pitfalls of goal failure.

The Current Research

The main hypothesis of the two studies was that framing task feedback at the primary goal level amplifies the negative effects of failure, as manifested in decreased expectancy, less positive affect, and more negative affect. Thus, the combination of failure and the primary goal evaluation will be contrasted with all other cells of a 2 × 2 design. Less important, we expected to see a main effect of success/failure feedback but did not expect a main effect of level of evaluation. These predictions were tested in both an experiment and a naturalistic longitudinal study.

Although our hypotheses have parallels within a variety of other literatures, the current research addresses a gap in the literature by (a) focusing on assigned (Study 1) or self-generated (Study 2) high- or low-level personal goals rather than assigned high- or low-level actions, as has past AIT research; (b) randomly assigning success versus failure feedback rather than letting success or failure be determined by differences in participant ability, as has some past high-versus low-level goal research; and (c) examining changes in future motivation and expectancy across multiple trial blocks as outcomes rather than examining current performance in a single block or task as an outcome. Again, we hoped to specifically address and answer the question When the going gets tough, should we remotivate ourselves by keeping our eye on the prize or by turning our nose to the grindstone?

STUDY 1

This experiment manipulated the hierarchical level and valence of performance feedback. Participants played a difficult skill-related game while receiving false feedback regarding either a primary goal or a subgoal. Then they rated their attitudes about the goal and their associated affect during the task. The 2 × 2 design helped us to ascertain the effects of level of goal (primary/subgoal) and valence of feedback (success/failure) on subsequent motivation and mood. A no feedback condition was also included as a hanging control condition.

Method

Participants

In all, 118 students (41 males and 77 females) from the University of Missouri–Columbia participated in this study. The incentive for participating was a $10 gift certificate or partial credit for those in an introductory psychology course. The mean age of participants was 21.5 years old. There were 23 participants in the subgoal/success (SS) condition, 24 participants in the primary goal/success condition (PS), 23 participants in the subgoal/failure (SF) condition, 23 participants in the primary goal/failure (PF) condition, and 25 participants in the no feedback (NO) condition.

Procedure

Participants came into the lab in small groups of up to 5 people and engaged in the task alone. The experimenter told participants they would be playing a computer word game designed to assess verbal ability. Participants were also told they would receive points for correct responses within each block of puzzles, and the points would accumulate to result in a final letter grade. Finally, participants were asked to keep in mind two goals while playing the game: (a) to earn the best grade for the task (primary goal) and (b) to earn the most points for each block (subgoal). The participants then made initial ratings of important constructs regarding both the primary goal and the subgoal (see following).

Next, participants started the computer game. Word puzzles from the Remote Associates Test (Mednick & Mednick, 1967) were presented on the computer for participants to solve. The participants were instructed that they should respond as quickly as possible to the computer prompts, that their responses would be scored by the computer for accuracy and speed, and that each puzzle presentation would time-out if a maximum time (25 seconds) was reached, though they were not told the specific amount of time. All participants received the same puzzles. Puzzles were presented in five blocks that contained five puzzles each. Difficult puzzles were used in all blocks so that the failure feedback would be believable.

The game program was tailored to present one of the five conditions, depending on the participant number that was drawn at random and entered into the computer by the participant. This kept the experimenter blind to the condition.
Success or failure feedback. After each block, the computer presented feedback that was determined by the participant’s condition assignment. The type of feedback was consistent throughout all of the blocks for every participant. The computer reported the following for the experimental conditions: “This completes the block. Press ‘5’ to view your results.” The next screen read, “Your responses are evaluated on accuracy and speed” and then presented participants with one of the following types of messages:

- Overall, you are on target to receive the grade of “A” compared to other college students. (PS)
- Overall, you are on target to receive the grade of “D” compared to other college students. (PF)
- For this block, the points you have earned put you in the 87th percentile compared to other college students. (SS)
- For this block, the points you have earned put you in the 47th percentile compared to other college students. (SF)

The specific percentile or letter grade was varied slightly within a range within each block. Participants in the no feedback condition went on to the next block without receiving the message about viewing results and did not receive any feedback.

Following Blocks 1 and 3, the participants made ratings of their current affect, expectancy, and intended effort on their assigned level of goal. At the end of the experiment participants were debriefed, and the experimenter inquired whether they were suspicious of the computer feedback. Most participants did not question the feedback. Those who were slightly concerned said the puzzles were so difficult that they thought it was plausible that the feedback was true. In other words, it was difficult for them to gauge how well they were doing relative to their peers. No participants were dropped due to suspicion.

Measures

Demographics. The participants reported their gender, age, ethnicity, year in school, current grade point average (GPA), and ACT score, where applicable.

Initial perceived goal difficulty. Participants were asked to rate both the primary goal and the subgoal as to their perceived difficulty of each goal. A not at all (1) to extremely (5) scale was used for these two ratings.

Initial perceived goal importance. Participants were asked to rate both the primary goal and the subgoal as to their perceived importance of each goal. A not at all (1) to extremely (5) scale was used for these two ratings.

State affect during the game. Participants rated six mood words after Blocks 1 and 2. The positive words were pleased, content, and satisfied. The negative words were discouraged, depressed, and agitated. The instructions read, “Select the number on the rating scale that best describes the way you feel. To what degree does this word describe you?” A not at all (1) to very much (5) scale was employed. The reliability for the positive mood scores were alpha = .896 and alpha = .940 for Blocks 1 and 3; the reliability for the negative mood scores were alpha = .836 and alpha = .791 for Blocks 1 and 3. Positive mood and negative mood were moderately correlated ($r = -.564$) and so they were treated separately.

Expectancy for future goal performance. After Blocks 1 and 3, participants rated the question “How well do you expect to do in the future on this goal?” on a scale of not at all well (1) to extremely well (5) for the specific goal level to which they were assigned that was shown at the top of the computer screen.

Results

Preliminary Analyses

To test for unexpected initial condition differences, means were examined for the initial variables of GPA and ACT using ANOVA. There were no condition differences for GPA or ACT (all $p$ values > .10).

To consider unique qualities of the two types of goal, the initial goal ratings for both types were compared within participant using paired $t$ tests. For initial perceived goal difficulty, the primary goal rating ($M = 3.32$) was higher than the subgoal rating ($M = 3.07$), $t(117) = 4.29, p < .01$. For initial perceived goal importance, the primary goal rating ($M = 4.14$) was higher than for the subgoal ($M = 3.91$), $t(117) = 4.11, p < .01$. These findings support our assumptions regarding basic differences between primary and secondary goals.

Hypothesis Tests

Expectancy for future blocks. We focused on the ratings made after Block 3, reasoning that the effects should be cumulative and require a few repetitions of feedback to emerge. To examine change in expectancy resulting from the feedback for the 2 × 2 experiment, expectancy after Block 3 was examined using an ANOVA with Block 1 expectancy as a baseline to control for earlier levels of expectancy. There was a main effect for valence of feedback, $F(1, 88) = 19.81, p < .01$, and a main effect for level of goal, $F(1, 88) = 3.86, p = .05$. There was also a significant interaction, $F(1, 88) = 9.33, p < .01$. These results are shown in Table 1. Expectancy was lowest for the primary/failure condition ($M = 2.04$) compared to subgoal/failure ($M = 2.96$), primary goal/success ($M = 3.92$), and
subgoal/success ($M = 3.74$). These unadjusted means are graphed in Figure 1.

In a planned contrast, the PF condition differed from all other conditions, $F(4, 112) = 7.14$, $p < .01$, including the no feedback condition. Again, expectancy in the primary/failure condition differed from all others. It is possible that this effect is reducible to differences in the initial perceived difficulty or importance of the primary goals, compared to the subgoals. Perhaps the primary/secondary distinction is just a proxy for these known constructs? To consider this we used an ANCOVA with the initial goal ratings of difficulty and importance for both levels of goal as four covariates.

Consistent with the previous model, there was a main effect for valence of feedback, $F(1, 84) = 15.89$, $p < .01$, and a marginal main effect for level of goal, $F(1, 84) = 3.69$, $p = .05$. There was also a significant interaction, $F(1, 84) = 7.90$, $p < .01$. None of the four covariates were significant.

**Negative affect.** To examine change in negative affect resulting from feedback in the $2 \times 2$ design, negative affect after Block 3 was examined using an ANOVA with Block 1 negative affect as a baseline to control for initial levels of negative affect. There was a main effect for valence of feedback, $F(1, 88) = 18.32$, $p < .01$, and no main effect for level of goal. There was a significant interaction, $F(1, 88) = 4.15$, $p < .05$. These results are shown in Table 2. Negative affect was highest for the PF condition ($M = 2.62$) compared to SF ($M = 2.27$) and the success conditions, PS ($M = 1.55$) and SS ($M = 1.94$); these unadjusted means are graphed in Figure 2.

In a planned contrast, the PF condition differed from all other conditions, $F(4, 112) = 6.33$, $p < .01$, including the no feedback condition. Thus, change in negative affect was significantly greater in the PF condition.

To consider the potential influence of the initial rated goal constructs, we used an ANCOVA with Block 1 negative affect as a baseline to control for initial levels of negative affect and initial goal ratings of difficulty and importance to control for goal type differences. There was a main effect for valence of feedback, $F(1, 84) = 15.10$, $p < .01$, and no main effect for level of goal. The interaction effect was slightly reduced but remained marginally significant interaction, $F(1, 84) = 3.49$, $p = .065$. None of the four covariates were significant.

**Positive affect.** Positive affect after Block 3 was examined using an ANOVA with Block 1 positive affect as a baseline to control for initial levels of positive affect.
TABLE 3: Study 1: Analysis of Variance for Positive Affect (PA)

<table>
<thead>
<tr>
<th>Source</th>
<th>df</th>
<th>F</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Baseline PA</td>
<td>1</td>
<td>94.54***</td>
<td>.000</td>
</tr>
<tr>
<td>Success/Failure (SF)</td>
<td>1</td>
<td>11.41***</td>
<td>.001</td>
</tr>
<tr>
<td>Primary/Subgoal (PS)</td>
<td>1</td>
<td>6.60**</td>
<td>.012</td>
</tr>
<tr>
<td>SF × PS</td>
<td>1</td>
<td>3.82**</td>
<td>.054</td>
</tr>
<tr>
<td>Within-group error</td>
<td>88</td>
<td>(0.31)</td>
<td></td>
</tr>
</tbody>
</table>

NOTE: Value in parentheses represents mean square errors.
**p < .05, ***p < .01.

Figure 3 Study 1: Means for Block 3 positive affect (unadjusted).

Brief Discussion

The data from this experiment indicate that there was indeed an effect of valence of feedback (success or failure) in interaction with goal level (primary goal vs. a subgoal). When failure feedback was given for the primary goal, participants had decreases in task expectancy and positive mood as well as increases in negative mood. Perhaps because the primary goal was felt to be more aligned with global self-goals, perceived failure on the primary goal had more of a negative impact on mood and expectancy. Furthermore, the observed effects could not be reduced to differences in the perceived difficulty or importance of the initial primary and subgoals.

The strengths of Study 1 lay in the experimental design. Participants were randomly assigned to the conditions, so any differences between conditions were due to the valence of feedback and goal level. Also, success and failure feedback was given as false feedback, eliminating any confound regarding actual participant abilities. However, there were some limitations to this study due to the experimental laboratory setting. To test goals in the lab, we needed to use somewhat artificial goals that may not have been very meaningful to participants. The laboratory setting only allowed us to look at goals that could be pursued in a very short time frame. In addition, participants were given false feedback that was somewhat artificial and the effects of such false feedback may not generalize to what occurs in our daily lives. We addressed these issues in Study 2 by looking at college students’ goals that were self-evaluated over a longer period of time.

STUDY 2

Study 2 was a longitudinal experimental study that allowed us to look at the effects of evaluating progress at one or the other level of goal over time, with success or failure varying naturally for individuals. This design allowed for the study of important personal goals while still having participants focus on either the distal or primary goal or a more proximal subgoal associated with the primary goal. The study included one experimental factor: assigned level of goal evaluation, which was maintained consistently throughout the 8 weeks of the study. Level of goal evaluation, taken in conjunction with self-reported success or failure on that assigned goal, was expected to affect the participants’ thoughts about the goals, their associated mood, and their perceived performance for the course.

To examine the robustness of the effects, we again controlled for several rated goal variables. In addition...
to controlling for both primary and secondary difficulty and importance, we also used a more refined measure of importance, based on self-determination theory (Deci & Ryan, 1985, 2000). Specifically, participants rated their perceived locus of causality for each goal, allowing for computation of a measure of felt “self-concordance” regarding the two goals (Sheldon, 2002; Sheldon & Houser-Marko, 2001).

Method

Participants

Students from the University of Missouri–Columbia participated in this study for partial course credit for an introductory psychology course. In all, 113 university students participated in the first questionnaire; 105 participants completed at least 7 of the 8 weeks of the study, and this was considered the final sample. Attrition analyses showed there were no differences between the retained participants and the 8 dropouts for any of the initial measures (all p values > .10).

There were 43 males, 69 females, and 1 participant who did not indicate gender. Of the participants, 12 said they were Black, 88 White, 8 Asian, 2 Hispanic, 2 indicated “other” race, and 1 did not indicate race. The mean age of participants was 19.5 years old. Of the final sample, there were 54 participants in the grade goal focus (primary goal) condition and 51 participants in the study hours goal focus (subgoal) condition. Eight participants were not included in the final sample due to an error in the game computer program.

Procedure

Participants were first asked to set two related academic goals: a primary goal of obtaining a letter grade for a course (grade goal) and a subgoal of studying a target number of hours each week for the same course (study hours goal). The participants rated both the grade goal and study hours goal regarding several goal variables in the first session questionnaire (see following). Then in the weekly surveys the type of goal focus was manipulated by asking the participants to evaluate their progress on only one of the two goals.

Participants completed the seven weekly surveys over the Internet. These short surveys inquired about the participants’ current mood and progress on either the grade goal or the study hours goal, depending on condition assignment.

The final questionnaire of the study was completed online at the end of 8 weeks. Participants completed final measures of mood, ratings of perceived performance, and ratings of their feelings about their grade goal and study hours goal.

First Session Materials

Demographics. The participants reported their gender, age, ethnicity, year in school, and high school or previous year grade point average.

Baseline current affect. Participants completed the Positive and Negative Affect Schedule (Watson, Clark, & Tellegen, 1988), which is a list of 20 mood words (10 positive and 10 negative), for a measure of mood. The instructions were to “indicate to what extent you feel this way, in the past month.” A baseline positive and negative affect score was computed for each participant. Reliability of the positive affect scale was alpha = .856 and the negative affect scale was alpha = .859.

Initial perceived goal difficulty. Participants were asked to rate both the grade goal and the study hours goal as to their perceived difficulty of each. A not at all (1) to extremely (5) scale was used for these two ratings.

Initial perceived goal importance. Participants were asked to rate both the grade goal and the study hours goal as to their perceived importance of each goal. A not at all (1) to extremely (5) scale was used for these two ratings.

Initial goal expectancy. Participants were asked to rate both the grade goal and the study hours goal as to the perceived expectation to do well on each goal. A not at all (1) to extremely (5) scale was used for these two ratings.

Self-concordance of goals. Participants were asked to rate both the grade goal and the study hours goal as to their perceived intrinsic, identified, introjected, and extrinsic motivation regarding each goal. A not at all (1) to extremely (5) scale was used for these ratings. A self-concordance score was calculated for each goal in which intrinsic and identified were summed and then the sum of introjected and extrinsic was subtracted (Sheldon, 2002; Sheldon & Houser-Marko, 2001).

Weekly Materials

Goal progress for each week. Again, weekly questionnaires asked the following questions regarding either the grade goal or the study hours goal, depending on the condition to which the participant was assigned. The questionnaires were essentially the same from week to week. Goal progress was determined from four achievement-related questions (i.e. “How would you rate your progress on this goal this week?”) on a Likert-type scale ranging from very poor (1) to exceptional (5).
Reliability of these four items within each week ranged from $\alpha = .823$ to $\alpha = .907$. A mean cumulative progress score was calculated. Overall reliability for the mean level of progress reported in the seven weekly questionnaires was $\alpha = .891$.

**Dichotomous success/failure evaluation.** Each week participants evaluated their goal by making a dichotomous choice of success or failure on the goal. Then they were asked to write a few sentences about their success or failure to make this evaluation more salient.

**Current positive and negative mood.** Participants also rated several mood words each week. Participants were asked to read each word and “indicate to what extent you have felt this way in the past week.” Overall reliability for the averaged weekly questionnaires was $\alpha = .911$ for positive mood and $\alpha = .924$ for negative mood.

**End of Study Materials**

**Final positive and negative mood.** The Week 8 survey asked participants to rate the same six mood words as were used in the weekly surveys for a measure of final mood. Reliability of the final positive and negative mood items was $\alpha = .907$ and $\alpha = .883$, respectively. Positive and negative mood were moderately correlated ($r = -.47$) and so were treated separately.

**Perceived performance.** In Week 8, participants were asked to indicate the grade that they expected to receive for the course at the end of the semester. The letter grades were coded $A = 4$, $B = 3$, $C = 2$, and $D = 1$. The mean expected grade was $M = 3.13$.

**Analyses**

We tested our primary hypothesis in two different ways. First, we took advantage of the longitudinal aspect of the study by looking at the cumulative effects of progress and the assigned level of goal focus on mood and perceived performance. In other words, the repeated framing of evaluation at one versus the other level of goal, in combination with overall rated progress, was expected to impact final mood and perceived performance. This approach considers the participant’s own mean evaluation of progress across the 7 weeks prior to the final 8th week. Cumulative progress could be thought of as a continuous measure of the success/failure concept. These analyses take a between-subjects approach using multiple regression models.

Second, an orthogonal analytic approach was used to consider reports at the week level. As each participant made a dichotomous choice of success or failure for each week, each weekly report could be thought of as a two-factor experiment, similar to Study 1. Notably, with this approach, the number of success/fail observations increased from 105 participants to approximately 630 weekly reports. To account for the person-level dependence of responses, we used multilevel analyses using SAS Proc Mixed. These analyses take a within-subject approach that models a person’s fluctuations around his or her own mean.

**Results**

**Preliminary Analyses**

To test for unexpected initial condition differences, means were examined using ANOVA. Participants’ GPA did not differ by condition ($p > .10$). Also, ratings for goal importance, goal difficulty, and self-concordance did not differ by condition (all $p$ values $> .10$).

The goal ratings were also compared to examine within-person differences between the grade goal and the study hours goal. Consistent with Study 1, we found that the grade goal was rated higher than the study hours goal for both importance, $t(112) = 5.72$, $p < .01$, $M = 4.40$ and $M = 3.93$, respectively, and difficulty, $t(112) = 5.38$, $p < .01$, $M = 4.01$ and $M = 3.55$. The within-person ratings of goal expectancy for either type of goal were not significantly different, $t(112) = 1.83$, $p = .07$, $M = 3.87$ and $M = 3.76$. Thus, the grade goal and study hours goal differed on the constructs of importance and difficulty (as in Study 1), but not expectancy, for participants.

**Negative Affect Outcome**

The model predicting change in negative affect included baseline negative affect; assigned goal level; rated progress on the assigned goal; rated difficulty, importance, expectancy, and self-concordance for both the primary and secondary goals; and the interaction term for assigned goal level and progress. First, in Step 1, baseline negative affect, assigned goal level (coded $\text{subgoal} = -1$, $\text{primary goal} = 1$), mean cumulative progress, and the goal control variables were entered. In Step 2, the interaction term for goal level and mean cumulative progress was entered. There was a main effect of baseline negative affect, $\alpha = .39$, $t(89) = 3.99$, $p < .01$, and a near main effect of cumulative progress, $\alpha = -.19$, $t(89) = -2.01$, $p = .047$. There were no main effects for level of goal focus or for any of the goal rating variables (i.e., subgoal and primary goal difficulty, expectancy, and importance).

Most important, the interaction of level of goal focus and mean cumulative progress marginally accounted for
unique variance, $\alpha = -0.19$, $t(89) = -2.05$, $p = 0.043$. As can be seen in Figure 4, the costs of framing failure for the grade goal are clear as negative affect was the highest in that case. However, negative affect was also the lowest after high progress in the grade goal condition. A test of the simple slope of the line for primary goal focus was significant, $t(99) = 2.675$, $p < 0.01$. The slope for the line for subgoal focus was not different from zero, $t(99) = -0.446$, $p > 0.10$, thus final negative affect did not differ when focusing on the low-level goal and experiencing differing degrees of progress.

**Positive Affect Outcome**

A similar model was tested for positive affect in which baseline positive affect; assigned goal level; cumulative progress; rated goal difficulty, importance, expectancy, and self-concordance; and the interaction term were included. First, in Step 1, baseline positive affect, assigned level of goal focus (coded as $\text{grade goal focus} = 1$ and $\text{study hours goal focus} = -1$), mean cumulative progress, and the goal control variables were entered. In Step 2 the interaction term for condition and mean cumulative progress was entered. In the final model there was a significant main effect of progress, $\alpha = 0.38$, $t(89) = 3.65$, $p < 0.01$. The effect of baseline positive affect was marginally significant, $\alpha = 0.18$, $t(89) = 1.72$, $p = 0.08$. There was not a main effect for level of goal focus. The interaction term was also nonsignificant ($p > 0.10$). Thus, positive mood was not influenced by the level of goal evaluation combined with cumulative progress.

**Perceived Performance Outcome**

Previous GPA was controlled for in a model predicting expected course grade. In addition, assigned level of goal focus (coded $\text{subgoal} = -1$, $\text{primary goal} = 1$), cumulative progress, and primary and secondary goal difficulty, importance, expectancy, and self-concordance were all entered in Step 1. In Step 2 the interaction term for condition and mean cumulative progress was entered.

There was a significant main effect of mean cumulative progress, $\alpha = 0.26$, $t(85) = 2.77$, $p < 0.01$. Previous GPA was also a significant predictor of expected course grade, $\alpha = 0.25$, $t(85) = 2.94$, $p < 0.01$. There was no main effect of level of goal focus, $p > 0.10$.

More important, the interaction of level of goal focus and mean cumulative progress accounted for unique variance, $\alpha = 0.24$, $t(95) = 2.57$, $p < 0.01$. As can be seen in Figure 5, there are greater costs as well as greater benefits for perceived performance when a person focuses on evaluating his or her primary goal. A test of the simple slope of the line for primary goal focus was significant, $t(99) = 3.37$, $p < 0.01$. The slope for the line for subgoal focus was not different from zero, $t(99) = 0.64$, $p > 0.10$, thus expected grade differed when progress reports focused on the primary goal but not when they focused on the subgoal.

**Week-Level Analyses**

Again, the design of this study allowed for a within-subject analysis in which outcomes during each week could be examined for the effects of that participant’s weekly dichotomous choice of success or failure. A multilevel model (using Proc Mixed in SAS) was created in which week was the Level-1 factor and person was the Level-2 factor, to account for the within-person variation or variation around a person’s own mean.

**Negative Affect for the Week**

Assigned level of goal focus (coded as $\text{grade goal focus} = 1$ and $\text{study hours goal focus} = -1$),
success/failure (coded as success = 1 and failure = −1), and the interaction term for level of focus and success/failure were entered into the first mixed model. There was a main effect of success/failure, $B = −0.567$, $t(617) = −16.93$, $p < .01$, as well as a near-significant Goal Focus × Success/Failure interaction predicting state negative mood, $B = −0.063$, $t(617) = −1.86$, $p = .06$ (see Table 4). This interaction had a similar form to the ones plotted earlier, with a somewhat smaller effect size.

**Positive Affect for the Week**

Assigned level of goal focus (coded as grade goal focus = 1 and study hours goal focus = −1), success/failure (coded as success = 1 and failure = −1), and the interaction term for level of focus and success/failure were entered into the second mixed model. There was a main effect of success/failure, $B = 0.644$, $t(617) = 17.35$, $p < .01$, as well as a significant Goal Focus × Success/Failure interaction predicting state positive affect, $B = 0.072$, $t(617) = 1.95$, $p = .05$ (see Table 5). This interaction also had a similar form to the ones plotted earlier, with a smaller effect size.

**Brief Discussion**

Study 2 built on Study 1 by examining the feedback-framing effects in a natural goal pursuit setting in which the goals were more meaningful to the college student participants. This longitudinal experiment showed the cumulative effects of consistently focusing evaluation at one level of goal. Cumulative progress and level of goal focus interacted to differentially affect final negative mood and expected grade for the semester.

**DISCUSSION**

To summarize the findings, Study 1 had an experimental design in which level of goal evaluation was crossed with valence of feedback. In a controlled setting, we found that when a person received failure feedback framed for a primary goal, he or she experienced the highest levels of negative affect and the lowest levels of positive affect and expectancy for future tasks. This

**TABLE 4:** Study 2: Multilevel Mixed Model Predicting Weekly Negative Affect

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Estimate</th>
<th>SE</th>
<th>df</th>
<th>t</th>
<th>Z</th>
<th>Significance</th>
</tr>
</thead>
<tbody>
<tr>
<td>Estimates of fixed effects</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Intercept</td>
<td>2.371</td>
<td>0.0627</td>
<td>103</td>
<td>37.80</td>
<td></td>
<td>&lt; .01</td>
</tr>
<tr>
<td>Primary/Subgoal (PS)</td>
<td>0.032</td>
<td>0.0627</td>
<td>103</td>
<td>0.51</td>
<td></td>
<td>&lt; .60</td>
</tr>
<tr>
<td>Success/Failure (SF)</td>
<td>−0.567</td>
<td>0.0335</td>
<td>617</td>
<td>−16.93</td>
<td></td>
<td>&lt; .01***</td>
</tr>
<tr>
<td>SF × PS</td>
<td>−0.063</td>
<td>0.0333</td>
<td>617</td>
<td>−1.88</td>
<td></td>
<td>&lt; .06*</td>
</tr>
<tr>
<td>Estimates of covariance parameters</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Unstructured covariance (1, 1)</td>
<td>0.318</td>
<td>0.0537</td>
<td></td>
<td>5.92</td>
<td></td>
<td>&lt; .01***</td>
</tr>
<tr>
<td>Residual</td>
<td>0.441</td>
<td>0.0251</td>
<td></td>
<td>17.55</td>
<td></td>
<td>&lt; .01***</td>
</tr>
</tbody>
</table>

**TABLE 5:** Study 2: Multilevel Mixed Model Predicting Weekly Positive Affect

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Estimate</th>
<th>SE</th>
<th>df</th>
<th>t</th>
<th>Z</th>
<th>Significance</th>
</tr>
</thead>
<tbody>
<tr>
<td>Estimates of fixed effects</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Intercept</td>
<td>2.710</td>
<td>0.0615</td>
<td>103</td>
<td>44.07</td>
<td></td>
<td>&lt; .01***</td>
</tr>
<tr>
<td>Primary/Subgoal (PS)</td>
<td>−0.028</td>
<td>0.0614</td>
<td>103</td>
<td>−0.46</td>
<td></td>
<td>&lt; .60</td>
</tr>
<tr>
<td>Success/Failure (SF)</td>
<td>0.644</td>
<td>0.0372</td>
<td>617</td>
<td>17.35</td>
<td></td>
<td>&lt; .01***</td>
</tr>
<tr>
<td>SF × PS</td>
<td>0.072</td>
<td>0.0370</td>
<td>617</td>
<td>1.95</td>
<td></td>
<td>&lt; .051**</td>
</tr>
<tr>
<td>Estimates of covariance parameters</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Unstructured covariance (1, 1)</td>
<td>0.278</td>
<td>0.0504</td>
<td></td>
<td>5.52</td>
<td></td>
<td>&lt; .01***</td>
</tr>
<tr>
<td>Residual</td>
<td>0.556</td>
<td>0.0316</td>
<td></td>
<td>17.56</td>
<td></td>
<td>&lt; .01***</td>
</tr>
</tbody>
</table>
primary goal/failure combined effect differed from all other cells in planned contrasts for the dependent measures of negative and positive affect and future expectancy.

For Study 2, all participants set both a grade goal and a study hours goal and evaluated their progress during the semester on their assigned level of goal only. We found that when a person self-reported lack of progress framed in terms of his or her primary goal, he or she experienced the highest levels of negative affect and the lowest levels of perceived performance at the end of the 8-week study. In addition, we were able to examine the responses for each week considering the two dichotomous factors of self-reported success versus failure and primary versus subgoal focus, similar to the design of Study 1. Here, the interaction of weekly success/failure and level of focus predicted relatively stronger negative and positive mood for that week. Thus, taking two orthogonal approaches to the data, both cumulative and within varying slices of goal pursuit, Study 2 showed that self-reported failure regarding a primary goal was associated with greater negative affect, less positive affect, and lower perceived performance.

Taken together, these two studies form a persuasive picture regarding the problematic effects of framing failure feedback in terms of the primary goal. Consistent with AIT, high-level identification negatively impacted mood and future motivation in the case of failure feedback. The picture is a little unclear regarding success feedback, however. For Study 1, the effects of success were equivalent when framed in either primary or subgoal terms. For Study 2, however, when progress was high (success), the participants who were focused on the higher level goal had higher perceived performance and lower negative affect. Thus, in Study 1 there was no benefit when success was framed in primary terms, but in Study 2, there was such a benefit. Because there are several differences between the two studies (i.e., experimental vs. longitudinal, trivial vs. more meaningful goals, success randomly assigned vs. naturally varying), it is impossible to tell from these data whether eyes on the prize is a more optimal affective regulation strategy when one is experiencing success. When things are going well with our current concerns, should we gratify ourselves by envisioning the attainment of our lifelong ambitions? Or do we risk getting too far ahead of ourselves, even in this case? Future research could further explore the possible benefits and drawbacks of high-level success evaluations.

Some readers may be reflecting on underlying similarities between this research and past research. It is worth considering these to further illuminate the possible contributions of our studies. First, the issue of “framing short-term failure in primary goal terms” has been approached in numerous guises in the past. In the individual difference literature, the issue has surfaced as trait pessimism versus optimism (Seligman, 1991), global versus specific attributional styles (Carver, 1998), entity versus incremental self-theories (Dweck, 1999), and concrete versus abstract goal focus (Emmons, 1992). In the clinical literature, the issue has surfaced in terms of “abstinence violation” effects (Marlatt & Gordon, 1985) and “what the hell” effects (Cochran & Tesser, 1996), in which small failures initiate psychological dynamics that lead to global failure. We suggest that all of these phenomena are underlain by the same common mistake—of framing momentary failure feedback in primary or long-range terms. This is understandable as the only reason we have the subgoal is to approach the primary goal, and it is natural to check to see if it is working. However, one has to be careful not to check too soon, especially early on in the process. “Am I there yet?” is not the best way to begin a long journey! Again, we suggest that we have identified an important common process underlying these overlapping concepts within the literature.

Another contribution of our studies, supporting AIT, is to show that people’s level of identification regarding a personal goal can be influenced by contextual forces (i.e., our experimental manipulations) and do not just represent stable individual differences. Thus, the “common mistake” mentioned earlier, of evaluating long-term progress too soon, might perhaps be avoided if mentors are careful about the language they use. It appears that motivating authorities (teachers, coaches, managers) should try to reinforce the importance of the subgoal, especially when a setback has just occurred. “Yes, we lost the game—now let’s get back to the practice field.” In this case, the level of identification is optimized, allowing attention to return adaptively to the plan or sequence leading toward longer term goals (Carver & Scheier, 1990).

Limitations

Both studies utilized self-report of mood and feelings about the goals. This brings up the possibility of demand characteristics. Participants rated their mood after receiving task feedback in Study 1 and after evaluating their own goal progress in Study 2. Thus, they may have felt some demand to report feeling better after success and feeling worse after failure. Although these demand effects might account for the main effect of success/failure, it is unlikely that the demand for such responses could account for the interactive influence of level of goal focus.

Another study limitation was that Study 2 relied on participants’ self-reports and self-evaluation of progress.
versus failure instead of randomly assigning them to progress versus failure. However, this limitation is somewhat mitigated by the experimental results of Study 1 and by the fact that Study 2 examined the processes in a naturalistic context that is very meaningful to students.

Future Directions

These studies both focused on academically relevant goals in that the Study 1 participants played a game for a letter grade and the Study 2 participants specifically pursued goals for a current class in the semester. However, we believe that the hierarchical framing of goal feedback would also apply to other domains of motivated behavior, such as physical activity and training, diet and eating habits, or career goals. We expect that no matter what the domain of motivated activity, as long as that activity matters to the person and requires multiple steps to achieve, the same effects would occur. However, this conjecture awaits further research.

Although these studies have robustly shown the effects of level of goal identification by controlling for difficulty and importance, future research could explore additional qualities of the levels of goals, such as time frame, stability, and clarity, to ensure that level of framing is not reducible to these quantities. Future research might also investigate level of framing as a chronic individual difference variable, as per AIT (Vallacher & Wegner, 1989), to evaluate whether such differences moderate the effects.

Our studies attempted to confine participants’ attention to just one level of goal, to test the effects of sustained focus at either one or the other level. However, people likely shift naturally between different levels of goal, perhaps frequently or infrequently. This might occur due to demands of the situation, for example when the task becomes difficult versus remaining routine. There also may be individual differences in the tendency to shift between levels or in the skill of shifting appropriately and fluidly. Some may get “stuck” and fail to shift focus when necessary (i.e., “state-oriented” individuals who are unable to get their mind off of failures; Kuhl & Baumann, 2000). These processes deserve further attention.

Another research direction concerns taking into account the association strengths between low- and high-level goals. As the subgoal and primary goal are associated to some degree within a hierarchy, they share cognitive activation and accessibility tendencies (Kruglanski et al., 2002). The strength of association between the subgoal and the primary goal will vary for individuals and for differing goals, and this would affect the likelihood of coactivation and thus premature evaluation. For example, it is likely that a primary goal would become accessible and salient when a subgoal was primed and the link between the goals was strong (Kruglanski et al., 2002; Shah, Kruglanski, & Friedman, 2003). Future research could consider the strength of cognitive links between personal goals and the effects of success and failure feedback as well as the effects of priming either level of goal on the outcomes of mood and motivation. Priming thoughts of a lower level goal might help a person to more easily make the shift to that level of focus and evaluation when necessary.

Conclusion

These studies provide a fresh look at what Little (1989) referred to as the trade-off between doing what is meaningful (high-level focus) and what is manageable (low-level focus). There certainly are benefits to thinking of goals at a higher level; these higher level goals are more self-relevant and holistic and give us a sense of direction in our lives. However, when focusing solely on the higher level goal, a person is more vulnerable to the detrimental and amotivating effects of momentary failure. This does not mean we should never look up from what we are doing: There are times when the prize must be referenced, reexamined, or even rejected. However, as a general strategy, it seems that especially during times of difficulty, it is more beneficial to keep your nose to the grindstone.

REFERENCES


Received December 20, 2007
Revision accepted May 30, 2008