Effectiveness of Distal and Proximal Goals as Transfer-of-Training Interventions: A Field Experiment

Travor C. Brown

The effectiveness of distal outcome goals, proximal plus distal outcome goals, and being urged to “do your best” interventions on self-efficacy and transfer was investigated in a field experiment involving government employees (N = 72). Six weeks after the training session took place, both participants who were urged to do their best and those who set proximal plus distal goals had increased transfer (that is, generalization and maintenance) relative to those who set outcome goals. There was no significant difference in the transfer level of participants urged to do their best and those who set proximal plus distal goals.

According to a recent survey, the average employee receives twenty-eight hours of training a year, at a cost of more than $800 per employee (ASTD, 2003). Yet despite the billions of dollars and employee hours that organizations dedicate to education and training programs, there is limited evidence that the skills learned during these programs are transferred to the workplace (Baldwin and Ford, 1988). One estimate suggests that employees transfer less than 10 percent of training and development expenditures back to their workplace (Georgenson, 1982). Given such estimates of the limited amount of training transferred back to the workplace, transfer of training—the application and

Note: The author thanks Gary P. Latham, Kate Dupre, the anonymous reviewers, and the editorial board for comments on earlier versions of this article. The support of the steering committee members (Susan Arscott, Roma Bridger, Bill Morrissey, and Karen Taylor) of the Public Sector Leadership Development Program is greatly appreciated. This study was funded through an internal research grant from the Memorial University Social Sciences and Humanities Research Council of Canada. A preliminary version of this article was presented June 2003 at the annual meeting of the Administrative Sciences Association of Canada.
usage of skills learned in the classroom posttraining—has become an important issue for organizations (Belcourt, Wright, and Saks, 2000; Noe, 2002). In a seminal article, Baldwin and Ford (1988) discussed two key transfer-of-training factors: (1) maintenance of learned material over time (that is, maintenance), and (2) generalization of learned material from the classroom to the workplace context (generalization).

In the area of posttraining transfer strategies, or interventions that occur after training (Salas and Cannon-Bowers, 2001), several studies used interventions grounded in social cognitive theory (Bandura, 1986). The use of theory in training research is important because the training literature has historically been criticized for lack of theoretical foundation as well as lack of rigorous and systematic analysis (Campbell, 1971; Latham, 1988; Tannenbaum and Yukl, 1992). In essence, social cognitive theory is grounded in an agency perspective, where people are seen as being contributors, through self-regulation, to the circumstances of their lives (Bandura, in press). As such, an important variable in this theory is self-efficacy, or one's confidence that one can perform a given task (Bandura, 1986, 1997). This task-specific confidence is seen as being a key determinant of performance; the higher one's self-efficacy, the higher one's level of task performance (Bandura, 1997). This is because self-efficacy affects effort, persistence, and task strategies (Lee and Bobko, 1994; Locke, Frederick, Lee, and Bobko, 1984; Locke and Latham, 1990). Thus, self-efficacy represents an important variable for successful training interventions for at least two reasons. First, it represents a trainee's self-assessment concerning the extent to which he or she can master the training content; thus, it is a measure of one's readiness to learn (Noe, 2002). Second, empirical research has shown that trainees' self-efficacy to perform the training task correlates positively with their subsequent posttraining performance on that task (Brown and Latham, 2000; Gist, 1989; see also review in Salas and Cannon-Bowers, 2001). For these reasons, scholars have argued that the effectiveness of training programs is largely dependent on the extent to which these programs strengthen trainees' self-efficacy to perform the training task (Belcourt, Wright, and Saks, 2000).

In terms of transfer interventions grounded in social cognitive theory, research has shown that relapse prevention, where participants are trained to anticipate and overcome obstacles to using the training content at work (Tziner and Haccoun, 1991), can improve transfer. Similarly, verbal self-guidance, where trainees learn to convert negative self-statements to positive statements that guide behavior, has been shown to improve transfer in terms of self-efficacy (Brown and Morrissey, 2004).

Goal-setting theory (Locke and Latham, 1990) is complementary to social cognitive theory in that it too identifies self-efficacy as a means of improving performance. Self-efficacy is also a key variable in this theory; people with high self-efficacy set difficult goals and are highly committed to them. A high level of goal commitment is desired because the positive effect of goals on
performance is strongest when individuals are committed to the goals they set (Locke and Latham, 2002). Goal-setting theory is founded on a number of core findings (see Locke and Latham, 1990). First, people who set specific, difficult goals outperform those who do not set goals or those who set vague goals (“I will do my best”). Second, there is a positive, linear relationship between goal setting and performance; the higher the goal, the higher the performance. Third, commitment is needed for goals to have a positive effect on performance. More than five hundred field and laboratory studies support these core findings (Locke and Latham, 2002). Given this wealth of empirical evidence, goal setting is considered to be one of the most valid and practical managerial theories for improving individual performance in the workplace (Pinder, 1997). This study was grounded in goal-setting theory and examined the impact of the type of goal on both self-efficacy and performance, in this case transfer of training.

Several studies have examined the effectiveness of goal-setting transfer-of-training interventions (Werner, O’Leary-Kelly, Baldwin, and Wexley, 1994; Wexley and Baldwin, 1986). Limitations of these studies include that (1) self-efficacy, a key variable for training effectiveness, was not measured; (2) the samples consisted of university students; (3) manipulation checks (goal commitment, specificity and difficulty) to assess the effectiveness of the goal-setting interventions were lacking; and (4) the usual do-your-best (DYB), or control treatment, was not provided to participants who did not set goals.

Gist and her colleagues have also conducted several transfer-of-training studies involving goal setting. Specifically, they found that a self-management training intervention consisting of writing a self-contract that contained goals and self-monitoring of one’s progress relative to the contract resulted in superior transfer to goal setting alone (Gist, Bavetta, and Stevens, 1990; Gist, Stevens, and Bavetta, 1991). In subsequent studies, Gist and her colleague (Gist and Stevens, 1998; Stevens and Gist, 1997) found that trainees who set an outcome goal (which they defined as a performance orientation where participants were urged to achieve superior outcomes) had a lower level of transfer relative to trainees who set a learning goal (defined as a mastery orientation where participants were urged to improve their skills).

The research of Gist and her colleagues creates a solid foundation for transfer-of-training research, but there are a number of limitations of this stream of research. First, these studies used simulated negotiation tasks rather than true workplace tasks. Hence, one cannot examine how far these results can generalize to workplace contexts and tasks. Second, this research involved student samples; thus generalizing these findings to older adults in workplace settings is unknown. Third, none of these studies used a control group (a DYB). Thus, the degree to which outcome goals would have resulted in superior performance relative to being urged to DYB was not tested. Fourth, these studies lacked goal-setting manipulation checks of goal specificity, difficulty and commitment; hence one cannot assess the relative effectiveness of
the goal-setting interventions in terms of goal commitment, difficulty, and specificity. In summary, the overarching limitation of these studies, from a human resources development (HRD) perspective, is the context. Each study was conducted in a simulated setting and was removed from the workplace setting. As a result, how well the findings from these studies generalize to the workplace is unknown.

However, the findings of Gist and her colleagues are consistent with recent goal-setting research. As argued by Locke and Latham (2002), the developers of goal-setting theory, outcome goals are less effective when people have not fully mastered a task. This is because such goals can cause people to focus on the consequences of failure; they fail to focus on developing the skills needed to master the task at hand. This argument is supported by research studies where participants who were trained to set specific, difficult performance goals performed worse than those who were urged to DYB (Kanfer and Ackerman, 1989; Earley, Connolly, and Ekegren, 1989).

Although these recent findings appear to contradict a core finding of goal-setting theory (that people who set specific, difficult goals outperform those urged to DYB), a potential cause of these negative goal-setting findings may be the type of goal set—namely, outcome goals to achieve a certain quality or quantity level. A potential antidote to this negative goal-setting effect involves breaking down a long-term, outcome goal (a distal goal) into several shorter-term, benchmark goals (proximal goals). In a laboratory experiment involving young adults, Latham and Seijts (1999) found that proximal plus distal outcome goals increased performance relative to being urged to DYB. These goals also resulted in increased self-efficacy. A potential explanation for these findings is that when the environment is dynamic, people need more feedback concerning their performance; thus, these proximal goals serve as benchmarks because they provide more regular feedback concerning performance than do distal goals (Locke and Latham, 2002). Limitations of Latham and Seijts’ study are that the context was a laboratory setting involving young students and that a simulated task (assembly of toys) was used. Again, a key limitation from an HRD perspective is that generalization of these findings to contexts involving adults in a workplace setting remains untested.

In terms of transfer of training, Latham and Seijts (1997) have argued that augmenting a distal goal with proximal goals should facilitate transfer. This is because trainees who set proximal goals should experience “small wins” that can motivate them to continue striving to attain the distal goal. However, in the absence of these proximal goals, trainees would not experience small wins and the associated motivational impact. But to my knowledge, no study to date has tested the effectiveness of proximal plus distal outcome goals as a transfer mechanism.

Given this backdrop, the purpose of this study was to assess the degree to which transfer interventions containing proximal plus distal goals, distal goals, and no goals (that is, being urged to do your best) could increase
trainee self-efficacy and transfer, in terms of maintenance and generalization. On the basis of the previously reviewed goal-setting studies, I hypothesized that:

**Hypothesis 1:** Participants urged to do their best have higher self-efficacy (1A) and transfer (1B) than those who set a distal outcome goal.

**Hypothesis 2:** Participants who set proximal plus distal goals have higher self-efficacy (2A) and transfer (2B) than those who set a distal outcome goal.

**Hypothesis 3:** Participants who set proximal plus distal goals have higher self-efficacy (3A) and transfer (3B) than those who are urged to do their best.

**Methods**

The sample consisted of seventy-two federal and provincial government employees working in the same Canadian province. These participants were enrolled in a university-based professional training program that was designed to develop competencies that the governments, through their internal needs analyses, identified as critical for managerial effectiveness. The participants of this training program were current managers at all levels (supervisors, managers, directors) as well as people being groomed for leadership positions. As such, they were representative of the target population of public sector managers. This particular sample of employees was selected for the training program by the organizations in question on the basis of the fit between the training content and the trainees’ developmental needs. On average, these participants were 42.00 (standard deviation = 5.63) years old, female (59 percent), and employed in a managerial position (78 percent). A true experimental design was used as the trainer randomly assigned participants to one of three study conditions. Participant sex ($\chi^2 = 1.11, p > .05$), age ($F = .61, p > .05$), and managerial status (manager or nonmanager; $\chi^2 = 3.13, p > .05$) did not differ significantly by study condition.

An advantage of this sample was that it consisted of full-time employees performing organizationally relevant tasks. As such, the sample overcomes a previously discussed limitation of past research concerning goal-setting transfer of training interventions, as well as the past research concerning proximal and distal goals—namely, the use of students performing simulated tasks. Thus, the sample enabled the first field experiment to compare distal and proximal plus distal goals in an HRD context.

**Procedure**

The key elements of the procedure consisted of the delivery of core training, study invitation, transfer condition assignment and training, time 2 data collection, and calculation of time 2 variables.
Core Training. All participants in all study conditions attended identical one-day self-awareness training programs. In total, four one-day sessions were conducted in two locations, given the geographic diversity of the organizations in question. Because the governments wished to ensure equity in terms of training delivery and content, identical training content (slides, surveys, handouts, exercises, and so on) was used in each session and the same trainer (the experimenter) conducted all four sessions. The use of multiple, but identical, training sessions is consistent with past training research (Gist, Bavetta, and Stevens, 1990).

In the core training program, trainees completed several self-awareness surveys. These specific surveys were the Fundamental Interpersonal Relationship Orientation Behaviors (FIROB; Schutz, 1958), which examines one's need for affection, inclusion, and control; Learning Styles Inventory (LSI; Kolb, 1999), which examines how one solves problems and learns; Locus of Control (Rotter, 1971), which examines the extent to which a person has an external locus of control (the person believes that she has little ability to affect her fate) as opposed to an internal locus of control (the person believes that he can influence his fate); and Tolerance of Ambiguity (Budner, 1982), or one's comfort with complex and novel situations.

In addition to completing these surveys, participants took part in discussions and exercises that examined how the differing leadership styles represented by these surveys can affect workplace effectiveness (for example, the extent of a manager's need for control minimizing effective delegation, how a reflective problem-solving style can be misinterpreted as “non-” participation in workplace decisions, and so on). These particular instruments were chosen by the program steering committee (made up of training representatives of the governments and the university's management training center, as well as the trainer) because the employees in question worked in provincial and federal government organizations undergoing substantial reorganization and realignment. Among the challenges faced by these organizations were: (1) an aging labor force, where many employees would retire in the next five years; (2) changes in the mandates and services provided by these departments; (3) a geographically diverse client and employee base; and (4) the continuous introduction of new technology and work processes. As such, this training program was specifically created to assist in developing new and emerging leaders. Senior management, as well as the governmental training staff who assisted in the initial needs analysis for the training program, felt that people managers, and future people managers, needed better understanding of themselves to further develop their managerial effectiveness. Consequently, they assisted in designing a one-day session that would help current and emerging leaders understand their own leadership style. This view is consistent with managerial education textbooks that emphasize the importance of self-awareness as a critical element of effective managerial skills (de Janasz, Dowd, and Schneider, 2002; Whetten and Cameron, 2002).
Study Invitation. At the end of the core training program, the trainer and the trainees discussed the importance of transferring the knowledge and skills learned in the core training program back to their workplace. Participants were then invited to take part in the research study. At that time, they were informed that their participation would consist of taking part in an in-class activity, completing a short survey in class, and answering a short survey six weeks after training. They were also informed that (1) their participation in the study was voluntary, (2) their surveys would not be shown to their employer, and (3) the final report would not present individual data.

Transfer Condition Assignment and Training. At the end of the core training, the trainer randomly assigned participants to one of three conditions: distal outcome goal, proximal plus distal goal, and DYB. To minimize contamination, participants of each condition met in different rooms for approximately thirty minutes and were given written instructions appropriate for their study condition. In each condition, the experimenter explained the importance of minimizing contamination. Participants were also asked not to share the specifics of their transfer condition with colleagues.

In the distal outcome goal condition, participants \( n = 22 \) were asked to individually set a specific, difficult goal for the number of times they would use the skills and knowledge gained from the training session back in their workplace over the coming six weeks. In table teams of two to four people, these participants shared their goal with their goal-setting peers. They were then informed that they could change their goals on the basis of feedback from their table group peers. This discussion was included in the intervention because having participants publicly state their goals can increase goal commitment and self-efficacy (Locke and Latham, 1990). At the end of the thirty-minute session, participants completed surveys assessing their goal, goal commitment, and self-efficacy.

The proximal plus distal outcome goal condition largely replicated the procedure of the distal outcome goal condition. The only exception was that participants \( n = 26 \) were asked to break their six-week goal into goals for two, four, and six weeks. They too met in small table teams to discuss their goals as well as completed surveys assessing their goal, goal commitment, and self-efficacy.

In the DYB condition, participants \( n = 24 \) were urged to do their best to use the knowledge and skills developed in the core training over the coming six weeks. Similar to the other conditions, they then met in small table teams to discuss why they needed to do their best to transfer the training content back to their workplace. They also completed self-efficacy questionnaires. However, they did not complete goal-level and goal-commitment surveys because they did not set a goal. Thus, the condition is consistent with the comparison (DYB) conditions in other goal-setting studies (see reviews in Locke and Latham, 1990, 2002).

Time 2 Data Collection. Six weeks after the training session, a posttraining survey was faxed as well as e-mailed to each participant at his or her workplace.
This survey assessed goal specificity, goal difficulty, training maintenance, training generalization, and self-efficacy. The six-week time frame was chosen for several reasons. First, examination of past research found that transfer was often measured four to eight weeks following a transfer-of-training intervention (Gist, Stevens, and Bavetta, 1991; Stevens and Gist, 1997; Werner, O'Leary-Kelly, Baldwin, and Wexley, 1994; Wexley and Baldwin, 1986). Second, as argued by Baldwin and Ford (1988), maintenance and generalization decay over time. As a result, I wished to pick a time frame that represented a reasonable but realistic time period for transfer. From my consultations with the steering committee, I felt that six weeks was an appropriate time frame.

Calculation of Time 2 Scores. A research assistant (RA), a graduate industrial relations and human resources student, received a copy of the posttraining questionnaire. Using the training material from the program (slides, handouts, exercises, and surveys), she independently created a structured score guide for each of the questions that assessed generalization. She also assessed content validity of this measure by verifying that all information assessed by the questions used to create the measure was relevant to the material covered in the training program. Depending on the question, her score guide allocated one or two points per question (some questions had two parts). This RA then evaluated each respondent’s survey using this structured score guide. A second RA, a graduate business student, reviewed the score guide for completeness and concurred with the initial score guide developed by the first RA. He then independently evaluated each survey using the same structured score guide. This second student was blind to the evaluation scores of the first student. Both students were familiar with the instruments being used because both had taken a management skills course where similar training content was presented. In addition, both RAs were blind to study hypotheses and study conditions, as well as which participants were assigned to which study conditions.

Measures

The key measures in this study were the manipulation checks (that is, goal specificity, difficulty, and commitment) and the dependent variables (that is, self-efficacy and transfer).

Manipulation Checks. According to goal-setting theory, goal setting is most effective when individuals set a specific, difficult goal and are committed to it (Locke and Latham, 2002). Further to the recommendations of Locke and Latham (1990), the goal-setting manipulation checks included in this study were (1) the actual goal set by participants, which is actually an objective measure of goal difficulty; (2) goal commitment; (3) goal specificity; and (4) goal difficulty. These manipulation checks were included for two reasons: to examine the effectiveness of the goal-setting interventions (the degree to which participants set specific, difficult goals and were committed to them), and to
examine whether participant scores on these measures differed significantly by goal-setting condition. This is important because, if, for example, participants who set proximal plus distal goals were found to have a higher level of these goal manipulation checks; a rival explanation for any significant differences in the dependent variable scores (self-efficacy and transfer) would be that these differences were not caused by the type of goal set per se but rather by the fact that these proximal plus distal goal participants set more specific, difficult goals and were more committed to their goals than the participants who set distal goals. Note that only participants in the distal and the proximal plus distal goal conditions were asked the goal-setting manipulation check questions; participants in the DYB condition did not set goals. This is consistent with past goal-setting studies (Brown and Latham, 2002; Winters and Latham, 1996).

Consistent with past studies (Brown and Latham, 2002; Seijts and Latham, 2001; Winters and Latham, 1996), the goal-level question asked participants to state their goals. Goal commitment was assessed using the five-item scale of Klein and others (2001). This scale was chosen because it is “highly stable and robust across a variety of settings and contexts” (Klein and others, 2001, p. 51). In fact, this five-item scale, when analyzed using a meta-analytical technique and factor analysis involving a sample of seventeen studies and more than twenty-nine hundred subjects, was found to be reliable (α = .74) with construct validity as demonstrated by the high factor loadings of the items (all above .65) on a single factor and a Kaiser criterion demonstrating that 49 percent of the variance in the sample was accounted for by the single factor of goal commitment. Goal specificity and difficulty were assessed using one question each adapted from Brown and Latham (2000) and Latham, Mitchell, and Dossett (1978; “My goal was specific”; “I believe that my goal was difficult”). The latter three measures were assessed using a five-point Likert scale (1 = strongly disagree and 5 = strongly agree). Consistent with Kervin (1992), all three measures were deemed to have theoretical validity in that on their face they appeared to reflect the variables they were designed to assess.

The dependent variables in this study were self-efficacy and training transfer, following Baldwin and Ford’s factors of maintenance and generalization (1988). Each measure was assessed using questionnaires sent to participants six weeks after training. Self-efficacy was also assessed immediately after goal setting.

**Self-Efficacy.** Self-efficacy is task-specific; there is no standardized scale that can be used for all studies (Latham and Pinder, 2005). Hence, a self-efficacy scale was created for this study following the recommendations of Lee and Bobko (1994). Specifically, the six-item scale asked participants if they believed that they could use the knowledge and skills gained through the self-awareness training three, five, ten, fifteen, and twenty times over the next six weeks (self-efficacy magnitude: yes or no) and how confident they were that they could do so (self-efficacy strength: 1 = no confidence to 10 = totally
confident). A composite score was created by summing all self-efficacy strength levels (as measured by the ten-point scale) when the person's magnitude was yes ("I believe I can do this"). This composite was chosen because past research has found it to be valid and reliable. Specifically, Lee and Bobko (1994), through two studies involving university students, found that the composite scale resulted in high convergent validity as demonstrated by a significant correlation between the composite self-efficacy score and four other methods of computing self-efficacy (in both studies the correlation coefficients ranged between 0.28 and 0.94; all were significant at \( p < .01 \)). A subsequent workplace training study (Brown and Latham, 2000) found a composite self-efficacy scale to be reliable (\( \alpha = .81 \) and .78 at times 1 and 2, respectively). The use of a frequency measure was chosen because (1) it is consistent with other goal-setting studies (Latham and Seijts, 1999; Seijts and Latham, 2001) involving proximal goals; and (2) the participants set goals concerning the frequency with which they could use the training content back at work; thus self-efficacy was also measured in terms of usage of training material back at work. The particular frequencies (three, five, ten, fifteen, twenty) were chosen on the basis of a pilot study in which participants, none of whom participated in the present study, were asked how frequently they felt they could use this material back at work; responses ranged from zero to twenty. Self-efficacy was assessed immediately after goal setting and in the mail-out survey six weeks after training.

**Maintenance.** Because maintenance refers to the process of continuing to use the skills from training back at the workplace (Noe, 2002), I measured maintenance in terms of self-report of usage of training material. Specifically, six questions were used. Five asked participants to self-assess the degree to which they had used various training material (LSI, Locus of Control, FIROB, Tolerance of Ambiguity, Overall Training Content) back at the workplace. These "usage" questions were assessed on a five-point scale (1 = not at all, 5 = very frequently). A sixth question asked the participants to state how frequently they had used the knowledge gained from this training session over six weeks. This "frequency" question was assessed using a six-point scale (1 = not at all, and 6 = one or more times a day).

**Generalization.** Consistent with Noe (2002), generalization was operationalized as the ability to apply training knowledge and skills to workplace problems and situations that were similar to the content of the training program. In this case, trainees were given eight questions or scenarios that assessed how well they could apply training content to workplace scenarios. None of the scenarios were covered in the training program. These were multiple-choice, open-ended, fill-in-the blank, and true-false questions. Two samples are, “What do you believe Jane, who is high on external locus of control, would say if she did not get a promotion when her friend Joan did?” and “Name two jobs which would be well-suited for a person high on active experimentation.”
Results

The analysis focused on an examination of respondents versus nonrespondents, manipulation checks, and the dependent measures of self-efficacy and transfer (that is, generalization and maintenance).

Respondents Versus Nonrespondents. Of the seventy-two people who took part in the core training and completed time 1 measures, only fifty-one (70.83 percent) returned the time 2 survey to the researcher. The number of responses for DYB, distal outcome goal, and proximal plus distal outcome goal conditions were fifteen, eighteen, and eighteen, respectively. Although a response rate of more than 70 percent is considered to be high for field research (Kervin, 1992), I conducted an analysis to determine whether time two respondents differed from nonrespondents. One-way ANOVAs failed to detect significant differences between respondents \((n = 51)\) and nonrespondents \((n = 21)\) in terms of goal commitment \((F = 0.00, p > .05)\), time 1 self-efficacy \((F = 0.00, p > .05)\), or age \((F = 0.54, p > .05)\). Nor did Pearson chi-square analyses detect differences between time 2 respondents and nonrespondents by gender \((\chi^2 = 1.64, p > .05)\), condition \((\chi^2 = 2.12, p > .05)\), or role (manager or nonmanager; \(\chi^2 = 0.21, p > .05)\). Taken together, these analyses suggest that there was no evidence of systematic nonresponse bias based on demographic variables, time one measures, or condition.

Manipulation Checks. As previously discussed, the goal-setting manipulation checks consisted of goal level, goal commitment, goal difficulty, and goal specificity. The mean goal set by participants in the distal goal condition was to use the skills and knowledge 12.67 (SD = 9.52) times over the next six weeks. Participants in the proximal plus distal goal condition set a goal of using the skills and knowledge 3.78 (SD = 2.68), 6.53 (SD = 5.72) and 9.27 (SD = 8.80) times in two, four, and six weeks, respectively.

In terms of goal commitment, difficulty, and specificity, goal-setting participants were highly committed to their goals \((M = 20.94, SD = 2.93, \alpha = .77, \text{maximum score} = 25)\), and they perceived their goals as moderately difficult \((M = 2.80, SD = 0.90, \text{maximum score} = 5)\) and specific \((M = 3.51, SD = 0.89, \text{maximum score} = 5)\). Table 1 presents the correlations between these measures. One-way ANOVAs revealed that these three measures did not differ significantly between distal and proximal plus distal goal conditions \((F = 0.42, p > .05; F = 0.05, p > .05; F = 1.10, p > .05)\), respectively.

Self-Efficacy. Composite scores for time 1 and time 2 self-efficacy were calculated consistently with the recommendations of Lee and Bobko (1994). The grand mean for time 1 and time 2 measures were 31.10 (SD = 12.94) and 26.36 (SD = 13.56), respectively, indicating that participants had moderate self-efficacy. The means and standard deviations of both self-efficacy measures, in addition to the transfer measures, by condition are presented in Table 2. Box plots of time 1 and time 2 measures failed to detect any outliers. Both time 1 and time 2 measures were highly reliable \((\alpha = .89 \text{ and } .90, \text{respectively})\) and
demonstrated test-retest reliability ($r = 0.44, p < .01$). Consistent with goal-setting theory, self-efficacy at time 1 and time 2 correlated with goal commitment ($r = 0.42, p < .01; r = 0.29, p < .10$, respectively). ANOVA failed to detect significant differences between the self-efficacy scores of study conditions at time 1 ($F = 0.23, p > .05$, eta$^2 = 0.01$) or time 2 ($F = 0.57, p > .05$, eta$^2 = 0.02$). Hence, the hypotheses concerning self-efficacy (1A, 2A, 3A) were not supported.

**Maintenance** Overall, the mean responses to each of the five usage questions ranged from a low of 2.92 (SD = 0.94) to a high of 3.26 (SD = 0.92) on a five-point scale. The mean response to the frequency question was 4.25 (SD = 1.04) on a six-point scale. Taken together, the measures suggest a moderate level of maintenance of the training skills six weeks after training. Because the five usage questions were assessed using a five-point scale and one frequency question was assessed using a six-point scale, z scores were calculated

---

### Table 1. Correlation Matrix

<table>
<thead>
<tr>
<th></th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
</tr>
</thead>
<tbody>
<tr>
<td>Time 1 self-efficacy</td>
<td></td>
<td>1.00</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(M = 31.10, SD = 12.94, $\alpha = .89$)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Time 2 self-efficacy</td>
<td>0.44*</td>
<td></td>
<td>0.29***</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(M = 26.36, SD = 13.56, $\alpha = .90$)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Goal commitment</td>
<td>0.42*</td>
<td>0.29***</td>
<td></td>
<td>0.19</td>
<td>0.23</td>
<td></td>
</tr>
<tr>
<td>(M = 20.94, SD = 2.93, $\alpha = .77$)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Goal difficulty</td>
<td>0.09</td>
<td>-0.17</td>
<td>-0.02</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(M = 2.80, SD = 0.90)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Goal specificity</td>
<td>0.27</td>
<td>0.13</td>
<td>0.19</td>
<td>0.23</td>
<td></td>
<td></td>
</tr>
<tr>
<td>(M = 3.51, SD = 0.89)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Maintenance</td>
<td>0.29**</td>
<td>0.42*</td>
<td>0.19</td>
<td>0.07</td>
<td>0.32***</td>
<td></td>
</tr>
<tr>
<td>(M = 0.00, SD = 4.09, $\alpha = .78$)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Generalization</td>
<td>0.02</td>
<td>-0.08</td>
<td>0.11</td>
<td>0.11</td>
<td>-0.14</td>
<td>0.11</td>
</tr>
<tr>
<td>(M = 7.99, SD = 1.46)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*Significant at $p = .01$, **$p = .05$, ***$p = .10$.

---

### Table 2. Key Measures, Means, and Standard Deviations by Condition

<table>
<thead>
<tr>
<th></th>
<th>Do Your Best</th>
<th>Distal Outcome</th>
<th>Proximal Plus Distal Outcome</th>
</tr>
</thead>
<tbody>
<tr>
<td>Self-efficacy (time 1)</td>
<td>31.26 (13.33)</td>
<td>32.41 (12.69)</td>
<td>29.85 (13.20)</td>
</tr>
<tr>
<td>Self-efficacy (time 2)</td>
<td>29.20 (12.00)</td>
<td>26.24 (13.66)</td>
<td>24.11 (14.95)</td>
</tr>
<tr>
<td>Maintenance (z score)</td>
<td>0.46 (2.57)</td>
<td>-1.28 (3.22)</td>
<td>1.52 (4.22)</td>
</tr>
<tr>
<td>Generalization</td>
<td>8.61 (0.68)</td>
<td>7.53 (1.71)</td>
<td>8.19 (1.20)</td>
</tr>
</tbody>
</table>

Note: Means, standard deviations, and Cronbach’s alpha coefficients are given in parentheses. Correlation coefficients exclude outliers.
before creating the maintenance variable. The resulting six-item scale was found to be reliable ($\alpha = .78$). Box plots of this measure revealed three outliers, one in each study condition; they were removed before subsequent analysis. Consistent with past research concerning self-efficacy and posttraining performance, two-tailed Pearson correlations found a significant relationship between participants’ use of training skills (maintenance) and both time 1 and time 2 self-efficacy ($r = 0.29, p = .05; r = 0.42, p < .01$, respectively). Consistent with past research concerning self-efficacy and posttraining performance, ANOVA also revealed a main effect for study condition on maintenance ($F = 2.84, p < .07, \text{eta}^2 = 0.11$). The hypotheses were directional; one-tailed $t$ tests were used consistent with the recommendations of Weinberg and Abramowitz (2002). In support of hypotheses 1B and 2B, participants who were urged to do their best and those who set proximal plus distal goals had higher maintenance scores than those who set distal outcome goals ($t = 1.68, p = .05; t = 2.18, p < .025$, respectively). However, contrary to hypothesis 3B, there was no significant difference in the maintenance scores of DYB and proximal plus distal goal participants ($t = 0.82, p > .05$). The means and standard deviations by condition are presented in Table 2.

**Generalization.** Two graduate students independently assessed generalization. Thus, consistent with past research (Brown, 2003; Smith-Jentsch, Salas, and Baker, 1996), two-tailed correlations were run between the scores of the raters in order to assess the level of interrater reliability. The results revealed that the correlations between the scores that the two raters provided for each of the eight questions were all greater than 0.75 ($p < .001$). Hence, the total generalization score was calculated by summing the mean score of both raters on all eight items. The resulting generalization measure had a mean of 7.99 (SD = 1.46; maximum score =10). This suggests a high level of generalization of training content to workplace scenarios not covered in training. A box plot revealed one extreme outlier in the DYB condition; it was removed before the subsequent analysis. ANOVA revealed a main effect for study condition on generalization ($F = 2.81, p = .07, \text{eta}^2 = 0.11$). In support of hypotheses 1B and 2B, one-tailed planned $t$ tests revealed that participants in the DYB and proximal plus distal goal conditions had a higher level of generalization ($t = 2.22, p < .025; t = 1.35, p < .10$, respectively) than those who set a distal outcome goal. Contrary to hypothesis 3B, the generalization scores of DYB participants did not differ significantly from those of proximal plus distal goal participants ($t = 1.15, p > .05$).

**Discussion**

As previously discussed, training research has often been criticized for its lack of theoretical grounding. Yet the advantages of theory-based training are that theory constitutes a framework for developing effective training design as well as assessing why training interventions work or fail (Latham and Crandall,
In this study, which was grounded in goal-setting theory, I sought to examine the extent to which transfer interventions containing proximal plus distal goals, distal goals, and being urged to do your best could improve self-efficacy and transfer of training (maintenance and generalization). Given the importance of transfer for both practitioners and researchers, this study makes significant contributions to the HRD field.

First, contrary to my expectations, self-efficacy did not differ between the experimental conditions. Yet past research suggests that proximal plus distal goals (Seijts and Latham, 1999) and DYB (Kanfer and Ackerman, 1989) interventions can increase self-efficacy. One possible explanation for the null effect is that I measured self-efficacy six weeks after goal setting, when these previous studies measured post-self-efficacy the same day as the intervention. In addition, the past studies examined self-efficacy in the context of simulated versus workplace settings. However, consistent with goal-setting (Locke and Latham, 1990) and social-cognitive (Bandura, 1986) theories, self-efficacy correlated positively with both goal commitment and subsequent performance of the skills learned in training (maintenance). Thus, the study reaffirms the importance of increasing trainees’ self-efficacy for training to be effective.

Second, these results offer additional evidence that distal outcome goals are not effective transfer-of-training interventions. These goals resulted in lower transfer (maintenance and generalization) than both being urged to DYB or setting proximal plus distal goals. These results are consistent with past goal-setting research, where studies have shown that when participants are learning new tasks, being urged to DYB results in superior performance relative to setting distal outcome goals (Kanfer and Ackerman, 1989; Latham and Seijts, 1999). Thus, for HRD practitioners and researchers alike there is now considerable evidence to support the conclusion that distal outcome goals are not effective in bringing about an increase in transfer when participants are learning new skills.

This conclusion concerning distal outcome goals should not be interpreted as meaning that all outcome goals are ineffective when used as transfer interventions. In fact, research suggests that two types of outcome goal may be very effective for transfer. First, research by Wexley and Baldwin (1986) and Werner, O’Leary-Kelly, Baldwin, and Wexley (1994) suggests that a behavioral outcome goal, where participants used checklists to set a goal of demonstrating skills learned in training, can bring about positive transfer. This finding is consistent with that of Brown and Latham (2002), where trainees who set behavioral outcome goals based on behavioral observation scales (Latham and Wexley, 1977) had superior performance relative to trainees urged to DYB. Thus it would appear that behavioral outcome goals can be effective transfer interventions.

Proximal plus distal goals represent the second type of outcome goal that can be an effective transfer intervention. The results of this study show that such proximal goals increase maintenance and generalization relative to
distal outcome goals. The results concerning transfer (maintenance and generalization) are consistent with the recent research concerning proximal plus distal goals where these goals have been found to be superior to being urged to DYB (Latham and Seijts, 1999). As such, these results support the argument of Latham and Seijts (1997) concerning the effectiveness of proximal plus distal goals as transfer mechanisms.

A third contribution of this study, for both HRD and goal-setting research, is that it is the first to assess the relative effectiveness of distal outcome versus proximal plus distal outcome goal-setting techniques as transfer-of-training interventions. Moreover, this study addresses several of the limitations of the previously reviewed transfer studies (Gist, Bavetta, and Stevens, 1990; Gist, Stevens, and Bavetta, 1991; Gist and Stevens, 1998; Stevens and Gist, 1997; Wexley and Baldwin, 1986; Werner, O’Leary-Kelly, Baldwin, and Wexley, 1994)—namely, the use of nonworkplace settings and largely simulated tasks, the lack of goal-setting manipulation checks, the use of student samples, and the lack of DYB comparison groups. As such, this study adds to the goal-setting literature by extending the validity of proximal plus distal goals to a population (adult employees) and context (maintenance and generalization of training skills) not previously examined.

Implications for Practice. The issue of transfer is of key importance to HRD practitioners. Given the billions of dollars spent each year on training and the limited evidence that skills learned in training transfer to the workplace, practitioners need transfer interventions that have been systematically investigated. Thus the results of the present study have three significant implications from an applied perspective. First, it shows how a very short proximal plus distal goal-setting intervention (thirty minutes) can have a positive effect on transfer. Moreover, this intervention did not require use of expensive consultants or copyrighted materials. Rather, it simply required that trainees set a goal (with benchmarks), discuss the goal with cotrainees, and then record it on paper. This simple but effective exercise could easily be added to many organizational training sessions. HRD professionals should therefore consider augmenting their current training programs with similar proximal plus distal goal interventions.

Second, the evidence of this study, combined with the literature reviewed earlier, makes a strong case for the argument that distal goals should not be set when people are learning new skills. A clear take-away from this study is that these distal goals must be augmented with short-term or benchmark goals, for trainees to effectively transfer training material back to the workplace.

Third, in addition to those reported in the literature review, this study shows that an increase in trainee self-efficacy concerning the skills being taught is associated with an increase in posttraining performance of these skills. This suggests that HRD practitioners should continue to develop and use training interventions that are complementary to, or grounded in, social cognitive theory.
Limitations and Directions for Future Research. Five potential limitations of this study are identified. First, the generalization of these findings to other settings and populations needs to be examined. As previously stated, fifty-one government employees across three conditions responded to the time 2 survey. The degree to which these findings hold true for other environments (nonprofit, private sector, and so on) as well as other populations (such as non-managerial employees) now needs to be investigated.

Second, the participants of this study were employed in organizations that are supportive of training. The organizations took part in an in-depth needs assessment, played an active role in designing the training program through their participation on the steering committee, and were supportive of the current experimental study because they, too, were interested in assessing ways to improve transfer. Past research has shown that supportive work environments enhance transfer (Burke and Baldwin, 1999; Richman-Hirsch, 2001). However, I did not include a measure of environmental supportiveness in this study. Future research should be conducted that examines the extent to which environmental supportiveness has an impact on the effectiveness of the present interventions.

Third, although graduate students assessed (that is, marked) the generalization questions, usage of the training material (maintenance) was assessed using self-report measures. Such measures have been used in previous works (see review in Salas and Cannon-Bowers, 2001), and there is some evidence concerning the reliability and validity of self-report measures for performance assessments (see review in Latham and Wexley, 1994), but Gist and her colleagues Stevens and Bavetta (1990, 1991) were able to benefit from simulations and so did not rely solely on self-reported measures for usage of training content. Unfortunately, this was not possible in this study. That being said, it is important to note that the findings concerning the effectiveness of being urged to DYB and setting a proximal goal plus distal goals were found for both the self-reported maintenance measure as well as the graduate student assessment of generalization, suggesting that the self-scores were accurate. Nevertheless, future HRD research should endeavor to use incremental sources of data to examine the usage of skills back at the workplace. In particular, research has shown that peer ratings of workplace performance are superior to self- and managerial assessments (see Brown and Latham, 2000; see also review in Latham and Wexley, 1994). Thus, in future studies, HRD practitioners and researchers should consider the use of peer assessments of how well trainees use the skills developed in training back on the job.

Fourth, the sample size was relatively small, with fifty-two respondents across the three study conditions. This sample size may have resulted in lack of statistical power to detect relationships that did indeed exist. In particular, it may explain the null effects found for self-efficacy. Hence, a type II error may be present in this study. Therefore, it should be replicated using a larger sample to see if these findings hold true.
A fifth limitation of this study is that I did not include a behavioral outcome condition when these goals have been shown to increase transfer. To my knowledge, no study has examined the extent to which behavioral outcome goals bring about an increase in self-efficacy or performance relative to proximal plus distal outcome goals. Thus, there is a need for a future HRD study that compares the relative effectiveness of these two goal types.

References


Travor C. Brown is associate professor of human resources and labor relations at Memorial University, St. John's, Newfoundland.