Explaining Performance Variability: Contributions of Goal Setting, Task Characteristics, and Evaluative Contexts

Susan E. Jackson
University of Maryland

Sheldon Zedeck
University of California, Berkeley

A laboratory experiment was conducted in order to test the effects on performance and satisfaction of goal-setting, task-characteristic, and evaluative contexts. Two hundred and sixty-three students participated in a 4 goal conditions (no goal, do your best, easy goal, and difficult goal) X 3 evaluative contexts (control, peer evaluation, and compliance) X 2 task characteristics (low and high variety) X 2 (order of task presentation) factorial design; all subjects worked on two tasks (manual and cognitive). Univariate analyses of multivariate analyses of variance results revealed: (a) Performance on the cognitive task was significantly affected by type of goal, task variety, and evaluative context, and (b) performance on the manual task was affected by task variety and evaluative context but not by type of goal. For both tasks, satisfaction was adversely affected by the presence of goals but was unaffected by evaluative contexts. For the cognitive task only, satisfaction was significantly higher in the low-variety condition. Research examining the effects of several potentially important task characteristics is suggested in order to develop a better understanding of goal-setting effects.

In a recent review of the goal-setting literature, Locke, Shaw, Saari, and Latham (1981) concluded that the relationship between goal setting and performance is reliable, persistent, and strong. Specific, difficult goals led to higher performance than did nonspecific, "do-your-best" goals in 90% of the studies they reviewed in which the goals could be assumed to have been accepted by the subjects. However, the strength of this relationship varies considerably from study to study. Important characteristics of goals that may limit or enhance the goal-setting effect have been extensively discussed (see Locke et al., for a review of these studies). Furthermore, performance quantity and quality have both been examined. Despite such careful consideration of both conceptual and empirical issues, a large amount of performance variability is unexplained by goal condition. Three potentially important contributors to performance variability are situational factors, task characteristics, and individual differences. The present study examines the relative contributions to performance of situational factors, task characteristics, and goals; the effects of individual differences (e.g., ability) are controlled for but, not explicitly examined.

Situational Factors

Most goal-setting research has ignored the potential impact on performance of the setting in which behavior occurs, although it is known that performance levels can be reliably improved simply by changing the test situation from one in which the person is performing in isolation from others to one in which others are present. Triplett's (1897) early work in this area demonstrated that competition against others improves performance more than does "pacing" (racing to beat an established time goal). Triplett theorized that competition led to better performance than pacing because it released latent energy (increased goal commitment) that was otherwise unavailable to the bicyclist. Zajonc (1965) has made a similar argument to ex-
plain social facilitation effects. Cottrell (1972) has qualified this argument somewhat, proposing that others are a source of drive or arousal only when they are perceived as important sources of evaluation.

In organizations, a worker may perceive co-workers and/or supervisors as potential evaluators. The importance of co-workers’ evaluations in determining the impact of goals on performance in work settings has been recognized in discussions of the concepts of peer competition (Latham & Baldes, 1975; Steers & Porter, 1974; Terborg, Note 1) and evaluation apprehension (White, Mitchell, & Bell, 1977). Employees working in groups are likely to compete with each other if the performance of each member will be known by others. When performance levels are public, behavior may be affected by the individual’s apprehension about how co-workers will evaluate his or her performance. Presumably, such apprehension arises out of the worker’s concern about whether co-workers will approve or disapprove of his or her performance.

Evaluation apprehension may also increase when performance is monitored by one’s supervisor, who can punish or reward job behaviors. A potential explanation for the goal-setting effect is that it introduces a concrete referent (the specific goal) against which performance can be easily judged and evaluated. If goals are not attained, tangible negative consequences may follow.

The present study examines the effects of goal setting in two types of evaluative situations. In a peer evaluation condition subjects are led to believe that their performance scores will be displayed publicly. In a compliance condition subjects are led to believe that the experimenter will terminate their participation in the study if performance is poor, thereby eliminating their opportunity to gain extra course credits. If goal setting is effective because it heightens concern about evaluation, there should be an interactive effect on performance of goals and situations: In the standard baseline condition in which evaluation apprehension is relatively low, goal setting should improve performance because it heightens evaluative concerns. However, when evaluative concerns are already maximized in the situation by other means, the goal-setting effect should be weakened.

**Task Characteristics**

Although goal attributes have been discussed and researched (e.g., Ivancevich & McMahon, 1977; Steers & Porter, 1974), relatively little direct study of task characteristics has been undertaken.

Terborg and Miller (1978) suggested that goals are most likely to affect performance on complex tasks where goals serve to cue the subject toward effective behaviors such as generating, testing, and implementing alternative strategies. If goal setting is effective partly because it encourages strategy development, it should be most effective for tasks that are relatively complex or high on variety. For simpler more repetitive tasks, the potential benefits of strategy development should be lessened and so goal setting should be less useful for improving performance. Two task characteristics were examined in the present study: the skill or ability (cognitive vs. manual) necessary to perform the task and task variety. Goal-setting theory predicts that goals should increase performance regardless of the type of skill required; the benefits of goal setting should be greatest for the high-variety versions of each task.

**Satisfaction With Performance**

Whereas numerous studies have investigated the effects of goals on objective performance, relatively few of these have included attitudinal measures, such as satisfaction with performance, despite speculation that goal setting should increase satisfaction (Steers & Porter, 1974). Studies that have assessed satisfaction under various goal conditions report inconsistent findings (see Ivancevich, 1976, 1977; Umstot, Bell, & Mitchell, 1976). In the study reported here, measures of satisfaction with performance and with the task were collected in order to provide further data relevant to the question of the relationship between goal setting and satisfaction.

**Method**

**Design**

Two hundred and sixty-three students (52% males and 48% females) enrolled in introductory psychology classes participated as volunteer subjects in a 4 (goal conditions) x 3 (evaluative situations) x 2 (task variety) x 2 (order of task presentation) factorial design; all subjects worked on two tasks (manual and cognitive). For each
subject, quantity of performance was assessed for a cog-
nitive and manual task. The four experimental factors
and their respective levels were as follows: Type of goal
(control, peer evaluation, compliance), task variety (low,
medium, high), and order of task completion (cognitive task first,
manual task first). Subjects’ satisfaction with their per-
formance and their reactions to the tasks were assessed in
a questionnaire completed at the end of the experi-
ment.

General Procedure

Most subjects (84%) participated in this experiment in
groups of three; some subjects (14%) participated in
groups of two; and a few (2%) worked alone. Tests of
the impact of group size revealed no significant effect.
All subjects in the peer evaluation condition participated
in groups of three. Within each session, all subjects
worked under the same experimental condition, which
was randomly assigned. Due to the complexity of run-
nin this experiment, equal cell sizes were difficult to
obtain. A check for confounding relationships among
the independent variables revealed no significant cor-
relations among treatment conditions. Therefore, statis-
tics appropriate for nonorthogonal designs were em-
ployed.

Each subject worked at a separate table, which faced
away from the other subjects in the room. Upon arrival
to the experiment, subjects were told they would be
working on several tasks similar to the kinds of tasks
people work on in various jobs. The subjects’ first task
was explained to them and they were given 10 minutes
to practice the task. After the practice session, subjects
were told they would be working on the task for 25
minutes and that they would then stop and begin a new
task. Subjects who were given goals were told their goals
at this point. After working on the first task for 25 min-
utes, subjects were stopped and the second task was ex-
plained. As for the first task, subjects were given 10
minutes to practice the task. Goals were then assigned,
when applicable, and a 25-minute work session followed.
Upon completion of the second task, subjects were given
the posttest questionnaires. The experimenter then ex-
plained the purpose and design of the study and thanked
the subjects for their participation.

Tasks

Two tasks similar to those used in previous goal-set-
ting studies were chosen to represent manual and cog-
nitive tasks in general.

A task similar to Terborg and Miller’s (1978) model-
buiding task was designed to represent manual tasks.
This manual task involved snapping together approxi-
mately 20 small interlocking plastic Lego pieces to form
a three-dimensional model vehicle (e.g., tractor, jeep,
aircraft). The pieces required to build one model were
enclosed in an envelope with an instruction sheet. The
instructions consisted of two pictures, one picture show-
ing the model half-built and one showing the model
completely built. By checking the models they built
against these pictures, subjects obtained feedback about
the quality of their performance. Completed models re-
ained on their tables until the end of the 25-minute
testing session, thereby providing performance feedback
to the subjects about the quantity of models they had completed.

The cognitive task used in this study was an adap-
tation of Weed and Mitchell’s (1980) task. In the present
study, subjects were given a simple floor plan of a one-
story, three-room building. All rooms were four-sided.
Although the dimensions of each room were not indi-
cated on the floor plan, sufficient information was pro-
vided to enable the subject to determine the dimensions
of each room. The subject’s task was to calculate the
number of units of carpeting to purchase (a) for each
room, assuming a different color would be used in each
room and (b) for the entire building, assuming the same
color would be used in all rooms. To provide feedback
to subjects about the quality of their performance on
this task, the solutions for each floor plan were enclosed
in the immediately succeeding envelope along with the
next floor-plan sketch. Quantity feedback was easily de-
termined visually since completed floor plans were
stacked on the subjects’ tables.

Manipulations

Task variety. Two versions of the manual and cog-
nitive tasks were designed in order to manipulate variety.
For the manual task, subjects in the low-variety condi-
tion (n = 125) were given a single type of vehicle to
assemble. Subjects in the high-variety condition (n = 138)
assembled five different types of vehicles. For the cog-
nitive task, subjects in the low-variety condition worked
on floor plans for buildings that were all the same shape
(rectangular), but with differing dimensions. The carpet
to be used in these buildings was always sold in the same
size unit (a 10’ X 20’ roll). In the high-variety condi-
tion, the three rooms in the building were arranged to
form shapes other than rectangles. The units of carpet
to be used in the high-variety conditions also varied from
one building to the next.

Goals. Four types of goal conditions were created:
no goal (n = 57), do-your-best goal (n = 81), easy goal
(n = 61), and difficult goal (n = 64). In the no-goal con-
dition, subjects were simply told to work on the task.
In the do-your-best condition, subjects were instructed
to “try to do your best on this task just do the best you
can.” In the easy- and difficult-goal conditions, subjects
were told to try to complete a specific number of models
and floor plans. Appropriate goals were determined by
a pilot study in which 25 subjects worked on the tasks
under a no-goal instruction set. Easy goals were defined
as half a standard deviation above the mean number of
models/floor plans completed by the pilot subjects.
Difficult goals were defined as 1.5 standard deviations
above the mean. Goals were established separately for low-
variety and high-variety versions of each task to allow for
differences in performance due to characteristics of the
task.

Evaluative context. Three types of situations were
created in order to examine the impact of heightened
concern about the evaluations of one’s peers (peer eval-
uation, n = 80) versus concern about complying to an
authority who has the power to withhold tangible re-
wards (compliance, n = 100). In a control condition (n = 83)
no attempt was made to heighten concern about the
evaluations of subjects’ peers or of the experimenter,
though some concern about each of these audiences can
be assumed to exist in any such experimental setting.
To heighten concern about peer evaluations, subjects were led to believe that their performance scores would be recorded on a chalkboard displayed in full view of all participants. When subjects arrived for the experiment, this board contained the names and scores of three putative subjects. To insure that subjects did not use these scores as a basis for setting goals, the fictitious scores were for tasks dissimilar from those on which the subject would be working. Before subjects began their first task, the board was erased and the subjects’ names were written on it. Subjects were told the following:

Before you continue working on the floor plans (models) in the box, I’d like to have each of you tell me your name so I can write it here on the board. This is where I’ll be recording how many floor plans (models) each of you completes so you can compare yourself to the others in the group.

For conditions in which goals were set, the experimenter stated the goals after the names had been written on the chalkboard. Then subjects began the 25-minute test period. After subjects finished working on the first task, the experimenter moved their completed models or floor plans to a scoring table. After receiving instructions and completing a practice session for the second task, subjects were told, "While you are working on the models (floor plans), I’ll be scoring your models (floor plans) so everyone can see how he or she did." During the time that subjects were working on the second task, the experimenter checked their work from the first task. The chalkboard was turned so subjects could not see it and the experimenter pretended to write scores next to the chalkboard. Then subjects began the 25-minute test period. Subjects were told the following:

For conditions in which goals were set, the experimenter stated the goals after the names had been written on the chalkboard. Then subjects began the 25-minute test period. After subjects finished working on the first task, the experimenter moved their completed models or floor plans to a scoring table. After receiving instructions and completing a practice session for the second task, subjects were told, "While you are working on the models (floor plans), I’ll be scoring your models (floor plans) so everyone can see how he or she did." During the time that subjects were working on the second task, the experimenter checked their work from the first task. The chalkboard was turned so subjects could not see it and the experimenter pretended to write scores next to the subjects’ names. In fact, subjects never actually saw their scores.

A compliance condition was created to simulate real-life concerns about the evaluations of those in control of valued rewards. All subjects who participated in this experiment were enrolled in a psychology course that required students to accrue three research credits. Various ways were available for accruing these credits, the favored way being to participate as a subject in ongoing research projects for a total of 3 hours. Subjects in the compliance condition had signed up to participate in a study for which they had been told they would earn two or three credits. The experimenter described the conditions of their participation as follows:

The sign-up sheet for this study indicated that the number of credits you can earn for this project is two or three which means 2 or 3 hours of participation. Because of the type of research I am doing, during the third hour of this study I can only use people who are able to perform tasks at a certain level. Therefore, as 2 hours of participation approaches, I will make a decision about whether each of you will be able to continue for the third hour and thus earn the full three subject credits. If I decide your performance is not satisfactory, as judged by a combination of several factors, you will be asked to quit the study. If you are terminated after the first 2 hours, you will receive 2 hours of subject credit. Do you understand the conditions of your participation here today?

Order. To control for the possible effects of task order, order of task completion was counterbalanced such that approximately one half of the subjects worked on the manual task first (n = 121) and half worked on the cognitive task first (n = 142).

Dependent Measures

Performance. Both quantity and quality of performance on the two tasks were assessed. Quantity of performance was defined as the number of task units (i.e., number of models built and number of carpet-purchasing problems solved) completed during the 25-minute testing sessions. For the manual task, credit was given for partially completed models based on the number of pieces assembled. Quality of performance was defined as the number of correctly completed task units. Because quantity and quality of performance were strongly correlated for these tasks (r = .90 and r = .92 for the manual and cognitive tasks, respectively), only the results for quantity will be reported.

Posttest questionnaires. After subjects had completed both tasks, they responded to two posttest questionnaires in order to report their perceptions and feelings related to each of the two tasks. Subjects’ perceptions of the tasks were assessed using five, 7-point bipolar scales. Subjects described each of the two tasks on the dimensions of easy/difficult, enjoyable/unenjoyable, simple/complex, repetitious/nonrepetitious, and interesting/boring. In addition, degree of autonomy, variety, and feedback from the tasks were assessed using Hackman and Oldham’s (1980) three-item indexes of task characteristics. Satisfaction with performance was assessed by asking, "How satisfied are you with how well you did on the (task):" Response categories ranged from 1 (completely unsatisfied) through 7 (completely satisfied).

Finally, the posttest questionnaires included several manipulation checks. To determine whether the experimenter’s goal was clearly communicated, subjects were asked, "Did the experimenter set a goal for you when you were doing the (task)?" Subjects who indicated "yes" then described the goal in their own words. Two questions were included as manipulation checks for the evaluative conditions: "How important is it to you that the other participants think you performed well on the (task) you worked on today?" and "How concerned were you with how the experimenter would evaluate your performance?"

Results

Manipulation Checks

Manual versus cognitive tasks. Subjects’ descriptions of the tasks on the posttest questionnaire were examined to check whether subjects perceived the manual and cognitive tasks as differing on dimensions other than the prima facie dimension of manual versus cognitive (see Table 1). Two-tailed, paired t tests revealed that the manual task was perceived as somewhat easier, less complex, more enjoyable, and more interesting. Subjects perceived the manual task as giving
less autonomy; the manual and cognitive tasks were perceived as no different on dimensions of feedback about performance or variety.

**Goal conditions.** Subjects' responses to the question, "Did the experimenter set a goal for you when you were doing the task?" were analyzed using a chi-square test, which revealed strong differences in responses across the four conditions, $X^2(3) = 42.93, p < .05$. For both tasks, 95% of the subjects in the easy- and difficult-goal conditions perceived that goals were set. In the do-your-best condition, 42% of the subjects perceived that goals were set. Sixteen percent of subjects in the no-goal condition perceived that goals were set. Subjects who perceived a goal set were asked to describe the goal. Their responses were coded as "nonspecific quantity and/or quality goal" (e.g., "do as many as I can correctly"), "do your best," or "specific quantity" (e.g., "do 7"). Looking at subjects who perceived that a goal was set, 91% of those in the easy- and difficult-goal conditions correctly described the set goal; of those in the do-your-best condition, 52% described the goal as "do your best" and 42% described the goal as "nonspecific quantity and/or quality". The few subjects in the no-goal condition who indicated a goal had been set described the goal as either "nonspecific quantity and/or quality" (7 subjects) or as "do your best" (2 subjects) and so were not dropped from the analyses.

Table 1

**Results of Manipulation Checks Comparing Tasks and Experimental Conditions**

<table>
<thead>
<tr>
<th>Manipulation check</th>
<th>Manual task</th>
<th>Cognitive task</th>
<th>Paired $t'$</th>
</tr>
</thead>
<tbody>
<tr>
<td>How easy</td>
<td>2.52</td>
<td>3.27</td>
<td>5.73*</td>
</tr>
<tr>
<td>How complex</td>
<td>2.48</td>
<td>3.12</td>
<td>5.36*</td>
</tr>
<tr>
<td>How enjoyable</td>
<td>2.74</td>
<td>4.45</td>
<td>2.30*</td>
</tr>
<tr>
<td>How interesting</td>
<td>3.47</td>
<td>4.70</td>
<td>9.19*</td>
</tr>
<tr>
<td>Autonomy</td>
<td>3.70</td>
<td>4.45</td>
<td>5.21*</td>
</tr>
<tr>
<td>Feedback</td>
<td>4.67</td>
<td>5.05</td>
<td>2.62</td>
</tr>
<tr>
<td>Variety</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Low variety condition</td>
<td>2.13</td>
<td>2.50</td>
<td></td>
</tr>
<tr>
<td>High variety condition</td>
<td>3.43</td>
<td>3.26</td>
<td></td>
</tr>
<tr>
<td>$F(1, 261)$</td>
<td>51.51*</td>
<td>19.16*</td>
<td></td>
</tr>
<tr>
<td>Repetitiveness</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Low variety condition</td>
<td>5.81</td>
<td>5.05</td>
<td></td>
</tr>
<tr>
<td>High variety condition</td>
<td>4.57</td>
<td>4.51</td>
<td></td>
</tr>
<tr>
<td>$F(1, 261)$</td>
<td>40.01*</td>
<td>7.12*</td>
<td></td>
</tr>
<tr>
<td>Complexity</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Low variety condition</td>
<td>2.53</td>
<td>3.14</td>
<td></td>
</tr>
<tr>
<td>High variety condition</td>
<td>3.24</td>
<td>4.02</td>
<td></td>
</tr>
<tr>
<td>$F(1, 261)$</td>
<td>3.69*</td>
<td>7.18*</td>
<td></td>
</tr>
<tr>
<td>Concerned about experimenter's evaluation</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Control condition</td>
<td>3.52</td>
<td>3.42</td>
<td></td>
</tr>
<tr>
<td>Peer evaluation condition</td>
<td>2.95</td>
<td>3.07</td>
<td></td>
</tr>
<tr>
<td>Compliance condition</td>
<td>3.84b</td>
<td>4.07b</td>
<td></td>
</tr>
<tr>
<td>$F(2, 260)$</td>
<td>6.91*</td>
<td>9.05*</td>
<td></td>
</tr>
<tr>
<td>Concerned about peer's evaluations</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Control condition</td>
<td>3.26</td>
<td>3.00</td>
<td></td>
</tr>
<tr>
<td>Peer evaluation condition</td>
<td>2.77</td>
<td>2.94c</td>
<td></td>
</tr>
<tr>
<td>Compliance condition</td>
<td>3.87</td>
<td>3.58</td>
<td></td>
</tr>
<tr>
<td>$F(2, 260)$</td>
<td>10.22*</td>
<td>4.34*</td>
<td></td>
</tr>
</tbody>
</table>

**Note.** $N = 263$.

*df = 262.*

* $p < .05$.

Planned comparisons revealed that the compliance mean is significantly higher than the other conditions.

Planned comparisons revealed that the peer evaluation mean is significantly lower than the other conditions.

Planned comparisons revealed that the peer evaluation mean is significantly lower than the other conditions.
Evaluative context. Compared to subjects in the control and peer evaluation conditions, subjects in the compliance condition were more concerned about "how the experimenter would evaluate your performance (on the task)." A contrast comparing the control and peer-evaluation conditions to the compliance condition found that subjects were most concerned in the compliance condition.

For both tasks, a main effect of evaluative context was found for the posttest question, "How important is it to you that the other participants think you performed well on the (task) you worked on today?" Inspection of the means for each condition revealed that subjects in the peer-evaluation condition reported relatively low concern about their peers' opinions, rather than the moderate amounts of concern reported by subjects in the control and the compliance conditions. This paradoxical result may reflect a self-protective stance of subjects in the peer-evaluation condition. The result was surprising for two reasons. First, during their debriefing sessions, subjects indicated they had believed that their scores would be posted on the chalkboard for others to see. Second, in a pilot study, subjects had been given the instructions for the experiment but were not exposed to the 25-minute test sessions. Compared to pilot subjects in the control condition (n = 7), subjects in the peer-evaluation condition (n = 15) reported on a pilot questionnaire feeling more "nervous" (M = 2.0 vs. M = 3.5), more "worried" about their performance (M = 2.3 vs. M = 3.9), more "motivated to try" (M = 3.4 vs. M = 5.0), and more "threatened" (M = 2.3 vs. M = 3.3).

Major Analyses

The two major dependent variables in this study were quantity of performance on a manual (model-building) and a cognitive (carpet-purchasing) task. Three major independent variables of interest in this study were type of goal, evaluative context, and task variety. Two independent variables of secondary interest were the order in which tasks were completed, which was manipulated as a precautionary control, and subjects' sex. Although subjects were not assigned to conditions so as to control for sex as a factor, sex of subject was examined as a fifth independent variable.

A 4 (goal) x 3 (situations) X 2 (task variety) x 2 (order) x 2 (sex) multivariate analysis of variance (MANOVA) for two dependent variables (quantity performance on two tasks) and unequal cells was performed using a classical experimental approach to test the relationship between independent and dependent variables. This analysis revealed significant main effects for goals, evaluative context, task variety, and sex, as well as a significant Goal x Task Variety x Order interaction (see Table 2). These effects and related supplementary analyses are discussed below.

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Goal</td>
<td>3, 176</td>
<td>4.22</td>
<td>1.46</td>
<td>13.10</td>
<td>4.92*</td>
<td>.10</td>
<td>6.348</td>
<td>3.00*</td>
<td></td>
</tr>
<tr>
<td>Evaluative context</td>
<td>2, 176</td>
<td>9.70</td>
<td>5.04*</td>
<td>8.53</td>
<td>4.81*</td>
<td>.09</td>
<td>6.348</td>
<td>2.66*</td>
<td></td>
</tr>
<tr>
<td>Task variety</td>
<td>1, 176</td>
<td>1,267.83</td>
<td>79.61*</td>
<td>68,042.13</td>
<td>355.51*</td>
<td>2.15</td>
<td>2, 175</td>
<td>189.00*</td>
<td></td>
</tr>
<tr>
<td>Sex order</td>
<td>1, 176</td>
<td>8.34</td>
<td>8.66*</td>
<td>7.24</td>
<td>8.16*</td>
<td>.08</td>
<td>2, 175</td>
<td>6.83*</td>
<td></td>
</tr>
<tr>
<td>Goal x Task variety x Order</td>
<td>3, 176</td>
<td>8.34</td>
<td>2.89*</td>
<td>2.41</td>
<td>.91</td>
<td>.07</td>
<td>6,348</td>
<td>2.16*</td>
<td></td>
</tr>
</tbody>
</table>

Note. N = 263.
*p <.05.
EXPLAINING PERFORMANCE VARIABILITY

creases the occurrence of self-set goals. Similarly, heightening evaluation apprehension may increase personal goal setting. Yes/no responses to the question, "When you were working on the (task), did you have a personal goal in mind for your performance?" were analyzed to test this possibility. A chi-square test revealed differing rates of self-set goals across evaluative conditions. For the manual task, subjects were more likely to set personal goals in the compliance condition (73%) than in the peer evaluation (45%) and control (45%) conditions, \( \chi^2(2) = 8.93, p < .05 \). This same pattern of greater fre-

\[ \text{Table 3} \]

Effect of Goal, Evaluative Context, and Sex on a Manual and a Cognitive Task Under High- and Low-Variety Conditions

<table>
<thead>
<tr>
<th>Condition</th>
<th>Manual task</th>
<th>Cognitive task</th>
</tr>
</thead>
<tbody>
<tr>
<td>Goal</td>
<td></td>
<td></td>
</tr>
<tr>
<td>No goal</td>
<td>37</td>
<td>57</td>
</tr>
<tr>
<td>Low variety</td>
<td>23</td>
<td>10.27 5.87</td>
</tr>
<tr>
<td>High variety</td>
<td>34</td>
<td>7.03 2.62</td>
</tr>
<tr>
<td>Do your best</td>
<td>81</td>
<td>6.68 2.42</td>
</tr>
<tr>
<td>Low variety</td>
<td>40</td>
<td>12.56 4.47</td>
</tr>
<tr>
<td>High variety</td>
<td>41</td>
<td>7.52 2.59</td>
</tr>
<tr>
<td>Easy goal</td>
<td>61</td>
<td>4.23</td>
</tr>
<tr>
<td>Low variety</td>
<td>33</td>
<td>2.08 5.35</td>
</tr>
<tr>
<td>High variety</td>
<td>28</td>
<td>15.86 4.35</td>
</tr>
<tr>
<td>Difficult goal</td>
<td>64</td>
<td>7.16 3.00 2.64</td>
</tr>
<tr>
<td>Low variety</td>
<td>29</td>
<td>12.68 3.87</td>
</tr>
<tr>
<td>High variety</td>
<td>35</td>
<td>7.92 3.65</td>
</tr>
<tr>
<td>Control</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Low variety</td>
<td>37</td>
<td>12.93 5.39</td>
</tr>
<tr>
<td>High variety</td>
<td>46</td>
<td>7.77 2.85</td>
</tr>
<tr>
<td>Peer evaluation</td>
<td>80</td>
<td>7.02 2.24</td>
</tr>
<tr>
<td>Low variety</td>
<td>45</td>
<td>13.04 4.18</td>
</tr>
<tr>
<td>High variety</td>
<td>35</td>
<td>7.58 3.51</td>
</tr>
<tr>
<td>Compliance</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Low variety</td>
<td>43</td>
<td>10.22 4.71</td>
</tr>
<tr>
<td>High variety</td>
<td>57</td>
<td>13.81 4.99</td>
</tr>
<tr>
<td>Sex of subject</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Male</td>
<td>141</td>
<td></td>
</tr>
<tr>
<td>Low variety</td>
<td>62</td>
<td>12.04 4.57</td>
</tr>
<tr>
<td>High variety</td>
<td>79</td>
<td>7.98 3.43</td>
</tr>
<tr>
<td>Female</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Low variety</td>
<td>63</td>
<td>12.03 5.23</td>
</tr>
<tr>
<td>High variety</td>
<td>59</td>
<td>6.69 2.02</td>
</tr>
</tbody>
</table>

For the cognitive task, contrasts of the univariate means were conducted to determine the source of the overall effect of goal condition. Performance was significantly higher among subjects who were given specific goals (easy and difficult conditions) compared to subjects given a general do-your-best goal or no goal, \( t(262) = 3.12, p < .05 \). Performance in the difficult-goal condition was significantly higher than performance in the no-goal, and do-your-best conditions; however, performance in the difficult-goal condition was not significantly different from performance in the easy-goal condition. For the manual task, performance in the difficult-goal condition was significantly higher than performance in the no-goal condition, but no other contrasts were significant. Table 3 presents the performance means for each goal condition.

Evaluative context. The predicted importance of evaluative context in determining performance was supported by a main effect in the MANOVA. Univariate tests indicated that situations affected performance on both the manual and cognitive tasks \( \omega^2 = 5\% \), and \( \omega^2 = 2\% \), respectively. However, inspection of performance means for the two tasks for each evaluative condition revealed an unexpected pattern. As shown in Table 3, performance was lowest, rather than highest, in the compliance condition. For both tasks, performance in the compliance condition was significantly worse than in the control and peer-evaluation conditions combined. For both tasks, performance in the control and peer-evaluation conditions were not significantly different.

Locke et al. (1981) hypothesized that competition improves performance because it in-

\[ \]
SUSAN E. JACKSON AND SHELDON ZEDECK

Table 4
Effect of Goal Condition on Satisfaction With Performance for Two Tasks

<table>
<thead>
<tr>
<th>Goal condition</th>
<th>Cognitive task</th>
<th>Manual task</th>
</tr>
</thead>
<tbody>
<tr>
<td>No goal</td>
<td>.23</td>
<td>.33</td>
</tr>
<tr>
<td>Do your best</td>
<td>.47</td>
<td>.39</td>
</tr>
<tr>
<td>Easy goal</td>
<td>-.31</td>
<td>-.34</td>
</tr>
<tr>
<td>Difficult goal</td>
<td>-.50</td>
<td>-.48</td>
</tr>
</tbody>
</table>

Note. Entries represent deviations of the cell means from the grand mean after controlling for objective performance.

SUSAN E. JACKSON AND SHELDON ZEDECK

The frequency of self-set goals in the compliance condition (74%) compared to the peer-evaluation (51%) and control (49%) conditions occurred for the cognitive task, but the result was only marginally significant, $X^2(2) = 5.13$, $p < .10$. Inspection of the types of self-set goals reported by subjects revealed that three fourths of such goals were general goals for performance quality and/or quantity.

Task variety. Inspection of the univariate $F$ tests reveals significant effects of task variety for both tasks ($w^2 = 29\%$ for the manual task, and $w^2 = 55\%$ for the cognitive task). The main effect of task variety is of limited theoretical interest in goal-setting research, however; variety was manipulated to determine whether goal setting has differential impact for tasks characterized by low and high variety. The lack of a significant interaction between goals and variety in either the multivariate or univariate analyses indicates that goal setting has the same impact on tasks low and high in variety.

Sex of subject. The MANOVA revealed that males significantly outperformed females on the tasks used in this experiment. Univariate $F$ tests indicate that sex affects performance on each task ($w^2 = 4\%$ for performance on the manual task, and $w^2 = 1\%$ for performance on the cognitive task). As already noted, sex of subject was not of theoretical interest to the authors, but was included as a factor in the analyses reported here so that its effects on the dependent variables would not be confounded with effects of the other independent variables.

Order. No significant effects of order of task completion were found.

Goal x Task Variety x Order. The overall MANOVA revealed an unexpected three-way interaction between goal condition, task variety, and order of working on the tasks. Univariate $F$ tests revealed that this effect was significant only for the manual task. Because this interaction is the only interaction out of 11 to reach significance, it is likely to be a chance result and will therefore not be interpreted here.

Satisfaction

Predictably, satisfaction with performance was correlated with actual performance ($r = .22$ and $.42$, $p < .05$, for the manual and cognitive tasks, respectively). Therefore, the impact of goals and evaluative context on satisfaction were examined using univariate analysis of covariance, treating performance as a covariate. Evaluative context was unrelated to satisfaction with performance but goal condition was related to satisfaction, $F(3, 178) = 6.04$, $p < .05$, $w^2 = 6\%$, for the manual task; and $F(3, 178) = 4.78$, $p < .05$, $w^2 = 5\%$, for the cognitive task. Overall, satisfaction was higher for the manual task, paired $t(262) = 9.11$, $p < .05$, but the pattern of results across goal conditions was the same for the two tasks. As shown in Table 4, goal had a negative effect on satisfaction. Satisfaction was highest in the do-your-best condition; it was slightly lower in the no-goal condition; it was lower still in the easy goal condition.

Finally, a significant effect of task variety was found for the cognitive task, with satisfaction being greater in the low-variety condition, $F(1, 178) = 24.09$, $p < .05$. Task variety had no effect on satisfaction with performance on the manual task.

Discussion

In the present study, specific (difficult and easy) goals led to better performance on a cognitive task than did a general goal (do your best) or no goal, but this effect was not found for a manual task. In addition, both evaluative context and task variety made independent contributions to performance. Finally, satisfaction was decreased by the presence of goals but was unaffected by the evaluative context.

In the past, when differences have been found across studies in the effectiveness of goal setting for different tasks, it has been
possible to link the differential effects of B°ai setting to task differences because of other important factors that also vary across studies. For example, one explanation for differential effects of goals has been differences in the ways researchers have operationalized easy and difficult goals (Locke et al., 1981). Easy goals have been operationalized as 10% above previous performance levels (London & Oldham, 1976), as a .75 probability of success (Motowidlo, Loehr, & Dunnette, 1978), and in the present study as one-half standard deviation above the mean performance level of pilot subjects. Difficult goals have been operationalized in equally as many ways, including a .26 probability of success (Frost & Mahoney, 1976), a .20 probability of success (Motowidlo et al., 1978), as 40% above previous levels of performance (London & Oldham, 1976), and in this study as one and one-half standard deviations above the mean performance level of pilot subjects. Inconsistencies in the methods that have been used to set goals, combined with the frequent omission of this information in published reports, are a serious problem in the research literature on goal setting, making meaningful comparisons across studies ambiguous at best, and often impossible. However, because goal levels were operationalized identically for both tasks in the study reported here, differences in the operationalization of goals cannot explain the Goal X Task interaction found in the present study. Nevertheless, these results should be interpreted cautiously, because although we have described the results as if the treatment levels for each factor are representative of a population of such levels, the levels may not be representative. Therefore, the results may have been different if other treatment levels had been chosen. In addition, a single study cannot explore all potentially important facets of either the evaluative situations created or the characteristics that distinguish the manual and cognitive tasks from each other. It appears that one or more characteristics of the tasks themselves are the cause of the differential effectiveness of goal setting. Responses to the posttest questionnaires indicated that subjects felt the manual task was easier, less complex, more enjoyable, and more interesting. Overall, the two tasks were rated as equal in variety. Evidence from the present experiment suggests that the dimensions of task difficulty and complexity may be less important in accounting for this Goal X Task interaction than are the subjects' interest in the task. Future research should test whether goal setting is more effective on tasks for which intrinsic motivation is initially low. When intrinsic motivation to perform is initially high, the potential for goal-setting techniques to increase motivation may be attenuated, rendering goal setting less effective (see Mossholder, 1980).

That the impact of goals was influenced by the nature of the task (cognitive vs. manual) suggests that closer examination of task characteristics is needed. For example, for the tasks used in the research reported here, relatively little information was provided to subjects about how to complete the tasks. Providing more complete information about how to do the tasks might affect the impact of the goal-setting manipulations. Another potentially important task characteristic is novelty. Given extensive experience with a task, a performance plateau may be reached so that performance improvements are less likely, and hence differences between goal conditions are negligible. Furthermore, tasks requiring a multitude of skills and knowledges may be more conducive to goal-setting effects than tasks requiring one or two skills such as those in the present study. Varying the degree to which the final outcome is known or familiar, the degree to which the steps for solution are prescribed, or the number of correct solutions to a problem may affect the impact of goal-setting procedures. In sum, the present results should encourage researchers to explore the impact of task characteristics on goal-setting effectiveness. Research on task taxonomies is much needed and may be beneficial to future goal-setting research.

In the present study, the evaluative context within which behavior occurred had a slightly stronger effect on performance than did goal setting. For both the manual and cognitive tasks, subjects performed worst in a compliance condition in which they believed they would be punished for poor performance. For both tasks, performance was best in a peer evaluation condition, and performance was intermediate in a control condition. The hypothesis that peer evaluation enhances
performance because it leads to spontaneous goal setting by subjects was not supported to the present study. Analyses of self-reported effort ratings (which are not reported here but are available from the authors) also failed to explain differences in performance across evaluative conditions. This study and other research (e.g., White et al., 1977) clearly point to the evaluative context as an important determinant of performance. More research in this area is badly needed in order to improve our understanding of the social dynamics that affect productivity in the workplace.

As noted above, the nature of the relationship between goal setting and satisfaction has also not been well-researched. In the present study, goals were associated with decreased satisfaction with performance. Our findings fit well with those reported by Rakesstraw and Weiss (1981), who found that satisfaction with one’s own performance was negatively related to the size of the discrepancy between one’s own performance and the performance of a model worker. Rakesstraw and Weiss argued that the model’s performance served as a standard against which subjects evaluated their own performance. Failure to meet this standard led to low satisfaction. Similarly, a specific performance goal serves as a standard for performance. In the easy-goal condition, about 50% of the subjects attained the assigned goal, whereas only about 10% of the subjects in the difficult condition attained the goal. For both tasks, success in attaining the assigned goal was correlated with satisfaction (r = .36 and .39 for the manual and cognitive tasks, respectively). The implication of these results for applying goal-setting theory is clear: Supervisors will need to learn how to strike a balance in order to set goal levels high enough to result in improved performance without setting them so high that they are unrealistic and likely to cause dissatisfaction.

Reference Note


References


Received April 5, 1982
Revision received July 12, 1982