Feedback Specificity, Learning Opportunities, and Learning

Jodi S. Goodman
University of Connecticut

Robert E. Wood
University of New South Wales

Although increasing feedback specificity is generally beneficial for immediate performance, it can undermine certain aspects of the learning needed for later, more independent performance. The results of the present transfer experiment demonstrate that the effects of increasing feedback specificity on learning depended on what was to be learned, and these effects were partially mediated through the opportunities to learn how to respond to different task conditions during practice. More specific feedback was beneficial for learning how to respond to good performance and detrimental for learning how to respond to poor performance. The former relationship was partially mediated by feedback specificity’s effect on learning opportunities during practice. The results have implications for designing feedback interventions and training to maximize the learning of various aspects of a task.

Textbooks routinely prescribe that supervisory activities such as performance appraisals (e.g., Cascio, 1998, p. 78), coaching, and development (e.g., Gomez-Mejia, Balkin, & Cardy, 2001, p. 248; Kreitner & Kinicki, 2001, p. 283) include specific feedback. Bernardin and Beatty (1984), for example, state that “. . . improvement is best fostered by specific verbal feedback provided by a supervisor or other appraiser, as close in time to the exhibited behavior as possible, and followed by suggestions on how future performance can be improved” (p. 197). The belief that greater feedback specificity leads to improved performance and learning has become an accepted generalization, despite a lack of evidence to support its positive impact on learning (Annett, 1969; Schmidt, 1991).

Feedback specificity refers to the level of information presented in feedback messages. As specificity increases, feedback focuses progressively on particular behaviors and provides more information on the locus of errors (Annett, 1969; R. A. Baron, 1988; Goldstein, Emanuel, & Howell, 1968; Payne & Hauty, 1955; Wentling, 1973). Specific feedback guides recipients to correct responses by helping them identify which behaviors are appropriate or inappropriate for successful performance (J. A. Adams, 1987; Anderson, 1982; Annett, 1969; Ilgen, Fisher, & Taylor, 1979; Payne & Hauty, 1955; Vroom, 1964). The increased guidance also decreases information-processing activities, such as error diagnosis, encoding, and retrieval (Christina & Bjork, 1991; Schmidt, 1991). Less active inferential processing is required to determine the links between actions and outcomes, because specific feedback does the work for recipients.

The majority of available evidence (see Goldstein et al., 1968, for an exception) shows that increasing the specificity of feedback has a positive monotonic relationship with immediate or short-term performance (Ilgen et al., 1979; Kluger & DeNisi, 1996). There is a clear body of evidence which shows that specific feedback that tells performers how far they are from the performance criterion leads to higher performance than less specific feedback which may just inform them of whether or not they are performing well (Kopelman, 1986). Feedback interventions in which feedback is augmented with supplemental information about the task, strategies, and appropriate behaviors also lead to better short-term performance during practice or training (see Kluger & DeNisi, 1996).

A less researched, and we believe more important, question is whether the augmentation of performance feedback delivers the expected learning benefits. That is, do increases in feedback specificity have a positive impact on performance in the longer term and when a task has to be performed more independently, with less specific or no supplemental feedback? This question is relevant for many situations, including those in which feedback is provided to new staff in training sessions or on the job, staff who are performing poorly because of skill deficits, and staff who have been recently promoted or assigned new tasks. A manager or trainer who gives very specific feedback to a staff member and then observes an immediate improvement in performance may mistakenly be led to believe that more specific feedback produces better learning.

The assumption that factors that enhance performance during the training or practice phase of a task also lead to superior learning is often untenable. Research has established that several interventions, including some feedback interventions, can simultaneously benefit practice performance and impair learning (Bjork, 1994; Christina & Bjork, 1991; Schmidt & Bjork, 1992). Also, factors that make practice more difficult and disrupt short-term performance lead to better retention and superior posttraining performance (Bjork, 1994; Christina & Bjork, 1991; Schmidt &
practice without the benefits of external facilitation forces trainees to engage in more exploration and information processing in their ability, errors, or reduced frequency of feedback often make trainees engage in more exploration and information processing in their attempts to identify correct responses. According to Schmidt and Bjork (1992), the demands of responding to task challenges during attempts to identify correct responses. According to Schmidt and Bjork (1992), the demands of responding to task challenges during attempts to identify correct responses decreases (Annett, 1969; Bjork, 1994; Ilgen et al., 1979; Payne & Hauty, 1955; Vroom, 1964). Conditions of practice that increase the guidance function of feedback lead to better performance during practice but suppress the information-processing activities that enhance learning (Bjork, 1994; Schmidt, 1991). At lower levels of specificity, feedback may stimulate the exploration and information processing that lead to inferences about the causes and effects of different responses. As feedback specificity increases, the feedback provides more guidance to participants for identifying and selecting responses. It does the inferential work for the recipient, so the amount of active exploration and information processing required to identify correct responses decreases (Annett, 1969; Bjork, 1994; Payne & Hauty, 1955; Schmidt, 1991).

The hypothesized relationships are summarized in Figure 1. In the discussion that follows, we present the arguments for the impact of feedback specificity on learning opportunities (Hypothesis 1) and of learning opportunities on learning (Hypothesis 2). The direct relationships between feedback specificity and learning (Hypothesis 3) and the mediation of those relationships by learning opportunities (Hypothesis 4) are corollaries of our arguments for Hypotheses 1 and 2.

Feedback specificity affects the relative exposure of recipients to instances of good versus poor performance during practice. At
lower levels of feedback specificity, performers are forced to experiment with different actions to discover which of the many possible responses are correct. This experimentation typically includes a range of actions, including errors and unsuccessful strategies that lead to declines in performance. As a result of the less informed and wider choice of actions taken while trying to learn a task, individuals who receive less specific feedback spend more of their practice time responding to poor performance than those who receive more specific feedback. At higher levels of feedback specificity, the guidance provided minimizes the frequency of errors and unsuccessful actions, leading to improvements in performance. As feedback specificity increases, it increasingly guides individuals more directly and more immediately to good performance conditions, increasing exposure to the correct responses associated with conditions of good performance and, correspondingly, reducing exposure to conditions of poor performance.

**Hypothesis 1**: Feedback specificity will affect learning opportunities, such that the higher the feedback specificity, the greater the percentage of instances of responding to good performance and the lower the percentage of instances of responding to poor performance during practice.

It stands to reason that performers must experience poor performance to learn decision rules for responding to poor performance and must experience good performance to learn decision rules for responding to good performance. Differential exposure to instances of good versus poor performance influences what is learned during practice, as associative learning is facilitated by the frequent coupling of actions with outcomes (Anderson, 1982). The greater number of instances of responding to good (poor) performance reinforces and aids retention and recall of the decision rules associated with the good (poor) performance. Thus, a person who has only managed high performing, competent subordinates may not know how to effectively manage incompetent subordinates, and vice versa. Similarly, a surgeon who has only performed surgeries that have gone well may not know how to effectively perform when complications arise; a machine operator who has only operated well-functioning machinery may not know what to do when a breakdown occurs; and a financial analyst who has managed funds only in a thriving economy may not know how to manage funds during a downturn, and vice versa.

To be effective across performance conditions, performers must learn to adjust their actions to different performance conditions. For example, managers must learn to adjust goal levels when setting challenging goals for different staff members. When a staff member is performing poorly, giving her a high goal may be discouraging and detrimental to performance if the goal is rejected, whereas a moderate level goal that is challenging relative to her performance should lead to performance improvements (Locke & Latham, 1990). For a staff member who is already performing well, the same high goal can be motivating and lead to further performance improvements, whereas a moderate level goal should be detrimental to performance (Locke & Latham, 1990). Managers who get experience managing only one level of performance (good or poor) have opportunities to learn one of these goal rules, but not the other. Those who are exposed to the different performance conditions of a task have opportunities to learn how to respond both when things are going well and when things are going wrong.

**Hypothesis 2**: Opportunities for learning will be related to learning outcomes, such that the percentage of instances of responding to good (vs. poor) performance during practice will be positively related to the learning of rules for correct responses to good performance and negatively related to the learning of rules for correct responses to poor performance.

Performers typically follow the guidance provided by feedback (Winstein & Schmidt, 1990). As they adjust their future behavior in accordance with specific feedback on their past behavior, performance increases, exploration and information processing are limited, and opportunities to experience errors decrease. Recipients receive more positive feedback as they repeat previous actions. Repetition aids the encoding and recall of what to do when
things are going well and facilitates the associative learning of correct responses to good performance. At the same time, the guidance provided by specific feedback decreases the extent to which performers develop hypotheses about how to correct errors and recover from poor performance.

Alternatively, by providing limited guidance, less specific feedback aids the learning of correct responses to poor performance by increasing opportunities to experience errors, encouraging the search for and evaluation of alternative actions in an attempt to identify appropriate responses, and increasing the instances of responding to poor performance. When feedback is less specific, the errors that result from experimentation during practice can also be a useful tool for learning, leading to new insights and creative solutions (Fresen et al., 1991). Although less specific feedback is beneficial for learning correct responses to poor performance, the limited guidance provided decreases the extent of associative learning of correct responses to good performance.

**Hypothesis 3:** Feedback specificity will be positively related to the learning of rules for correct responses for good performance and negatively related to the learning of rules for correct responses for poor performance.

On the basis of the arguments for the previous hypotheses, we predict that learning opportunities will serve as important mechanisms by which feedback specificity affects what is learned. As feedback specificity increases, recipients have more opportunities to learn decision rules for responding to good performance and fewer opportunities to learn decision rules for responding to poor performance, which in turn is beneficial for the learning of the rules for responding to good performance and detrimental to the learning of the rules for responding to poor performance.

Our predictions are consistent with evidence that repeated positive feedback increases the replication of existing scripts as compared to repeated negative feedback, which results in more strategy processing and variety of cognitive processes (Wofford & Goodwin, 1990). Specific feedback facilitates repetition of previous actions and avoidance of exploration and information processing, leading performers into good performance conditions, where they will continue to receive positive feedback as they repeat previous actions. This repetition facilitates memories of what to do when things are going well, but limits opportunities to experience poor performance conditions and learn what to do when things are not going well.

**Hypothesis 4:** The relationships between feedback specificity and learning of the rules for correct responses for good and poor performance will be mediated by learning opportunities. Specifically, (a) feedback specificity will be positively related to the percentage of instances of responding to good performance during practice and, subsequently, to the learning of correct responses to good performance, and (b) feedback specificity will be negatively related to the percentage of instances of responding to poor performance during practice and, subsequently, to the learning of correct responses to poor performance.

**Method**

**Overview**

We manipulated the specificity of feedback provided during a practice phase on a task that was unfamiliar to the participants. Participants completed a 19-trial practice task under one of three nested levels of feedback specificity, which ranged from outcome-only feedback at the low specificity end through to outcome-plus-specific-process feedback on all decisions at the high specificity end. We then tested the effects of the feedback provided during practice on performance of two transfer tasks (14 trials), completed 2 days later, without specific feedback. Learning was measured as the accuracy of responses to the different performance conditions encountered during the transfer tasks. This measurement strategy enabled us to examine the hypothesized existence of differences in the learning of the rules for correct responses for good versus poor performance across feedback specificity conditions.

**Participants**

Individuals were recruited from undergraduate management courses at a large midwestern university. They were given extra course credit for their voluntary participation. One hundred ninety-two individuals completed the practice task. Participants returned 2 days later and completed the testing tasks, which were used to assess the learning that had occurred at Time 1. Nine participants were lost as a result of attrition from Time 1 to Time 2, reducing the Time 2 sample size to 183 (95.3%). Participants were 61% male and 39% female and ranged in age from 18 to 41, with a mean age of 21.5 (SD = 2.85). Because we used a management decision-making simulation as our experimental task, we assessed participants’ management-related experience with a five-item, 5-point scale, with 1 = none and 5 = very much. We asked them to rate their previous experience with delegating work to others, assigning performance goals to others, setting their own performance goals, giving performance feedback to others, and rewarding others for their performance. As expected, participants had only a moderate amount of previous management-related experience (M = 2.67, SD = 0.79, \( \alpha = .80 \)), which facilitated the study of learning.

**Experimental Design**

We conducted a single-factor experiment, with repeated measures. Participants were randomly assigned to one of three treatment conditions: low, moderate, or high feedback specificity. The manipulations are described below. The repeated measures included multiple trials of the practice task and the two testing tasks.

**The Task**

The study was presented as a project in managerial decision making. Participants served as special order department manager in a business simulation called the Furniture Factory. This simulation has been used to study the effects of task complexity and goal setting on self-regulation, exploration, and performance (e.g., Wood, Atkins, & Bright, 1999; Wood, Bandura, & Bailey, 1990). This simulation was chosen because it has many different alternative responses and therefore requires exploration and information processing for the discovery of the decision rules that link actions to outcomes. This same feature also makes it possible to create feedback interventions that differ in the objective levels of task-relevant information. The task also provides records of all decisions made and whether each decision is correct or incorrect, which can be used to objectively measure participants’ learning of specific responses to different conditions of performance. Another relevant feature of the task is that the correct responses for some of the decision variables change between the good and poor performance conditions. This makes it possible to separately
assess the learning of the rules for correctly responding to good performance versus poor performance. As special order department manager, participants managed a small group of three workers during the practice period and five workers during the testing period over a series of performance trials. Each performance trial represented a week at the Furniture Factory. In their managerial role, participants assigned the workers to jobs and motivated them by providing goals, feedback, and rewards. The feedback just mentioned should not be confused with the feedback provided to the participants on their managerial performance and decisions, which we manipulated with our feedback specificity treatment conditions.

Participants received written instructions describing the simulation task; the skills and work-related preferences of each worker; the descriptions and requirements of each Furniture Factory job; and choices of types of goals, feedback, and rewards they could give to their workers. The instructions were presented on a computer screen and on a reference sheet to which participants could refer throughout the simulation. To learn how to manage their workers, participants had to figure out the correct responses for assigning goals, providing feedback, and allocating rewards to their employees, which are outlined in Table 1 and described below.

Participants made four decisions for each worker on each simulation trial. The actions taken on one trial influenced performance on subsequent trials, thus modeling the temporal effects of such actions in actual work environments. The available choices for the decisions represented the types of actions managers might take in an actual organization, and the impacts of different responses on performance were based on relevant findings from research in goal setting, feedback, expectancy, and equity theory. First, participants assigned each of the available workers to one of the production jobs required to complete the weekly order. Production of the weekly furniture orders required different jobs, such as milling timber, assembling parts, staining and glazing assembled frames, upholstering furniture, and preparing products for shipment. The workers differed in their skills and preferences to perform each job and in their potential maximum performance levels once assigned to a job. Once workers were assigned to jobs, participants set a goal for each worker from a set of five options. Participants could choose to give no goal, tell a worker to do his or her best, set a goal 25% worse than the standard (low goal), set a goal equal to standard (moderate goal), or set a goal 25% better than standard (high goal). Participant managers then chose among four types of performance feedback to provide each worker. They could choose to provide no feedback, discuss with the worker what he or she did correctly and incorrectly when performing the job (process feedback), inform the worker of his or her performance level in relation to the standard for the job (outcome feedback), or provide both process and outcome feedback. Participants then chose from a set of three reward levels for each worker. They could choose to give no reward, praise a worker (moderate reward), or publicly recognize a worker by posting a memo (high reward). Job assignment, goal, feedback, and reward decisions were made for each worker on each trial of the simulation. Participant managers could choose to maintain or change any or all of their decisions from trial to trial.

Participants had to learn the best responses for the job allocation, goal, feedback, and reward decisions through the testing of different responses and interpreting the feedback they received. The relationship between participants’ actions and their workers’ performance was governed by the set of rules outlined in Table 1, which are represented in the mathematical model used to calculate the hours taken to complete the assigned furniture order on each trial. Participants received no information regarding the mathematics of the model or the rules. Individual worker performance was computed with the mathematical model for the decision rules (see Wood & Bailey, 1985, for the mathematical model), and group performance was a simple addition of the performance of individual workers. The performance levels of workers and the group are reported as a percentage of the standard time for the job being performed. These standards were originally derived from time and motion studies of furniture factory jobs and have been modified over time to reflect the average time taken by participants who have completed the simulation task.

Optimal responses were contingent on a worker’s past performance (see Table 1). When the simulation began, the high goal of 25% better than standard was the optimal goal choice, because challenging goals tend to motivate high performance (cf. Locke & Latham, 1990). The high goal continued to be optimal until a worker experienced repeated, considerable failure. If a worker performed very poorly (i.e., at or below 20% worse than standard) for two consecutive trials, the moderate goal of achieving the standard became challenging and the optimal goal choice. If the worker performed well (i.e., at or better than standard) later, the high goal became the optimal choice again. The optimal choice for goal allocation could change back and forth across trials depending on a worker’s performance levels.

The simulation also began with outcome-plus-process feedback as the optimal feedback choice. This continued until a worker demonstrated job competence, operationalized as performing at or better than standard for three consecutive trials. At this time, process feedback may be viewed as over-supervision, so outcome feedback alone became the optimal feedback choice. If the participant manager changed the job assignment, outcome-

<table>
<thead>
<tr>
<th>Table 1</th>
<th>Decision Rules for Simulation</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Decision type</strong></td>
<td><strong>Rule description</strong></td>
</tr>
<tr>
<td>Employee job allocation</td>
<td>Assign each employee to a job on the basis of the match between job and employee characteristics.</td>
</tr>
<tr>
<td>Goal</td>
<td>Give the high goal initially.</td>
</tr>
<tr>
<td></td>
<td>Give the moderate goal after an employee performs very poorly for 2 consecutive weeks (≤ 20% worse than standard).</td>
</tr>
<tr>
<td></td>
<td>Give the high goal after an employee performs very well (≥ standard).</td>
</tr>
<tr>
<td></td>
<td>Giving no goal, or the low goal is never optimal.</td>
</tr>
<tr>
<td>Feedback</td>
<td>Give outcome-plus-process feedback initially.</td>
</tr>
<tr>
<td></td>
<td>Give outcome-only feedback after an employee performs well (≥ standard) for 3 consecutive weeks.</td>
</tr>
<tr>
<td></td>
<td>Giving no feedback, or process-only feedback is never optimal.</td>
</tr>
<tr>
<td>Reward</td>
<td>Give an employee no reward for poor performance (&lt; standard).</td>
</tr>
<tr>
<td></td>
<td>Give the moderate reward when performance is close to standard (standard ≤ performance &lt; 5% better than standard).</td>
</tr>
<tr>
<td></td>
<td>Give the high reward for good performance (≥ 5% better than standard).</td>
</tr>
</tbody>
</table>
plus-process feedback was again the optimal choice until the worker performed the new job at or better than standard for three consecutive trials. The optimal choice for feedback provision could change back and forth across trials depending on job allocation changes.

Optimal responses regarding rewards were based on a worker’s performance on the current trial. If a worker performed poorly (i.e., below standard), no reward should have been given. If a worker performed moderately well (i.e., standard ≤ performance ≤ 5% better than standard), the moderate reward should have been given. If a worker performed well (i.e., > 5% better than standard), the high reward should have been given. In addition, the mathematics linking chosen reward options to worker performance included a comparison of the reward-to-performance ratios of the different workers, as outlined in equity theory (J. S. Adams, 1963). As a result of this comparison, the impacts of chosen reward options on individual workers were partially dependent on rewards given to other workers.

**Manipulations**

Three nested feedback interventions of differing specificity were created, and participants were randomly assigned to one of the feedback specificity conditions. Participants in each condition were provided outcome feedback, and the specificity of process feedback on their decisions as special order department manager was manipulated across conditions. The information provided to participants by the simulation, and described below, ranged from outcome feedback and no process feedback in the low specificity condition, up to outcome-plus-specific-process feedback on each action taken in the high feedback specificity condition.

We recognize that increasing the specificity of feedback often entails varying other feedback dimensions, such as the amount of information. As feedback specificity increases, recipients often receive more information, but rather than producing information overload, the added information in more specific feedback reduces the active information processing required to infer correct responses. Although the manipulations described below differ in the amounts of information the participants had to read, more important, they differ in the extent to which active inference is required to determine the links between actions and outcomes. The lowest feedback condition provided no error information; thus it required the most active inference on the part of the recipient. The highest feedback condition provided the locus of error for each decision made by the participants, thus requiring the least amount of active inference. We chose to label our variable “feedback specificity” because of the variation in the level of information processing required and because our operationalization is similar to those used in past studies of feedback specificity (e.g., Goldstein et al., 1968; Goodman, Wood, & Hendrickx, 2004; Payne & Hauty, 1955; Wentling, 1973).

During the simulation, feedback was presented in writing on the computer screen. All feedback described for each condition was given for each performance trial and remained on the screen until it was updated for the next trial. A brightly colored message flashed to inform participants when feedback was updated, and feedback messages were written in red on a blue background to direct attention to the feedback.

**Low Feedback Specificity**

In the low feedback specificity condition, participants received outcome feedback that included the weekly job performance levels for each of their three workers and for the group as a whole. This included the hours taken by each worker and the work group to produce the order for that trial, the estimated standard hours for the order, and a comparison that showed the percentage above or below the standard. The outcome feedback was objective performance data that provided enough information for participants to quickly assess how each worker and the work group were performing relative to the estimated standards for the jobs and the total weekly furniture order, respectively. This was chosen as the baseline feedback specificity condition to test the positive and negative effects on learning of increasing the specificity of such feedback by adding supplemental process information.

**Moderate Feedback Specificity**

Participants in the moderate feedback specificity condition received the same type of outcome feedback as described for the low feedback specificity condition. This outcome feedback was supplemented with error signal feedback about their decisions. For each set of decisions about job allocation, goals, feedback, and rewards they provided to their workers on each trial, they were informed either that their decisions were made correctly for all workers or that their decisions were incorrect for at least one worker. For example, if a participant had given the correct goal to all workers on a particular trial, he or she was told, “You gave every worker the right type of goal.” If a participant made at least one goal allocation that was inconsistent with the correct choice for that decision, he or she was told, “You gave at least one worker the wrong goal.” If the participant had incorrectly assigned workers to jobs, he or she would get the message, “You have assigned at least one worker to the incorrect job.”

**High Feedback Specificity**

Participants in the high feedback specificity condition were provided the same type of outcome feedback as recipients in the low and moderate feedback specificity conditions. This was supplemented with more specific error signal feedback for each decision made for each of their workers. Participants were told whether the job allocation, goal, feedback, and reward they gave to each worker were right or wrong. Therefore, in addition to the outcome feedback, the information they received included a series of 12 statements (3 workers × 4 decisions) on each trial such as, “You assigned Jack to the wrong job,” and “You gave Neil the correct reward.”

**Procedure**

To study the effects of feedback specificity on both learning opportunities and learning, we conducted a transfer experiment that took place over two laboratory sessions.

**Time 1 Practice Period**

Participants were randomly assigned to one of the three feedback specificity conditions described above (i.e., low, moderate, or high). The experiment took place in a large computer laboratory, where data were collected from 20 participants at a time. The participants were seated at every other computer to avoid distraction and prevent them from seeing other participants’ computer screens. They began the experiment simultaneously and sat quietly until everyone had finished. When all participants had finished the experiment, they were dismissed together. Participants performed 19 trials (or “work weeks”) of the management simulation task to allow for practice, performance improvement, and rule learning. All participants were instructed that their goal during the practice period was to learn the rules for allocating jobs, goals, feedback, and rewards to each of their workers.

**Time 2 Testing Period**

Participants returned 2 days later for the second part of the experiment, to assess the learning that had occurred during the Time 1 practice period. Participants performed two testing tasks, lasting eight trials and six trials, respectively. They were provided the same type of outcome feedback that was provided to participants in all three feedback specificity conditions.
during the practice period. The testing tasks differed somewhat from the practice task, as described below; however, the rules and mathematical model that linked the participants responses to performance outcomes remained the same.

The purpose of the first testing task was to maximize the instances participants responded to good performance, thus providing them with ample opportunities to demonstrate the extent to which they had learned the rules for responding to good performance (i.e., give the high goal, outcome feedback, and the high reward). Compared to the practice period, participants supervised five, rather than three, workers, and a new set of workers was used to increase the chances that test performance was indicative of rule learning rather than memorization and repetition of responses made during the practice period.

The purpose of the second testing task was to ensure that participants experienced several instances of responding to poor performance, thus providing them with ample opportunities to demonstrate the extent to which they had learned the correct responses to poor performance (i.e., give the moderate goal, outcome and process feedback, and no reward). This was accomplished by assigning participants five workers who performed fairly poorly in the Furniture Factory jobs, even when correct management decisions were made. Together, the two testing tasks provided the variance in testing performance necessary to establish levels of learning for the good and poor performance conditions and to test the hypotheses.

**Measures**

**Learning Opportunities**

This variable was measured during the Time 1 practice period and was operationalized as the percentage of instances of responding to good versus poor performance to which participants were exposed during practice. We determined for which instances each participant should have given the high goal, outcome feedback, and/or the high reward to each of his or her workers during the practice period, as stipulated by the task rules listed in Table 1 and described earlier. The percentage of instances of responding to good (vs. poor) performance was computed as the total number of times a good performance rule should have been applied divided by 171 (3 workers $\times$ 3 decisions $\times$ 19 trials).

**Learning**

Learning was assessed during the testing period (Time 2). By measuring learning as performance at a later time, on a variant of the task, with minimal guidance, we were able to assess the extent to which the skills and task knowledge necessary for independent task execution were learned during the Time 1 practice period. Performance during practice may consist of true learning effects and transient effects (Christina & Bjork, 1991; Schmidt & Bjork, 1992). There is an inherent confound between transient and learning effects in measures of performance during practice (Wulf & Schmidt, 1994) that makes such measures ambiguous with respect to the amount learned (Schmidt & Bjork, 1992, p. 209). Through design, we were able to eliminate any transient feedback effects and focus on the learning effects that were evident across time and variations of the task.

We separately assessed participants’ learning of the rules for correctly responding to good performance and their learning of the rules for correctly responding to poor performance. Overall measures of learning that are typically used may obscure feedback specificity effects because of differential effects on the learning of correct responses to good versus poor performance. Separately assessing the learning of rules for correct responses for good and poor performance allowed us to identify whether feedback specificity is beneficial for, or detrimental to, each of the two aspects of learning.

**Good performance rule learning.** This variable was measured during the Time 2 testing period, for all 14 trials combined, as the percentage of instances the participant made the correct responses to good performance, as stipulated in the rules listed in Table 1. We determined for which instances each participant should have given the high goal, outcome feedback, and/or the high reward to each of his or her workers and for which of those instances these decisions were correctly made. The variable was computed by dividing the total number of times each good performance rule was correctly applied by the total number of times each good performance rule should have been applied.

**Poor performance rule learning.** This variable was measured during the Time 2 testing period, for all 14 trials combined, as the percentage of instances the participant made the correct responses to poor performance, as stipulated in the rules listed in Table 1. We determined for which instances each participant should have given the moderate goal, process and outcome feedback, and/or no reward to each of his or her workers and for which of those instances these decisions were correctly made. The variable was computed by dividing the total number of times each poor performance rule was correctly applied by the total number of times each poor performance rule should have been applied.

**Feedback Specificity Manipulation Check**

Twenty doctoral students ranked and rated the specificity of the feedback in the three experimental conditions. The doctoral students were given our definition of feedback specificity and examples of the feedback for each of the three conditions, but they had no knowledge of the purpose of the experiment or of the hypotheses. All 20 students ranked the three conditions in order of specificity, consistent with our intended order from low to high. The ratings, which were made on a 5-point scale, anchored by low specificity = 1 and high specificity = 5, were also consistent with the intended manipulations. The mean ratings differed significantly, $F(2, 57) = 226.76, p = .000$, from one another in the expected order (low $\bar{x} = 1.1, SD = 0.31$; moderate $\bar{x} = 2.3, SD = 0.47$; high $\bar{x} = 3.8, SD = 0.41$). The results of this manipulation check show that the doctoral students agreed with the objective distinctions in the information content of the three feedback specificity conditions.

**Data Analysis**

We used analysis of variance to test Hypothesis 1, which predicted that increasing feedback specificity would affect learning opportunities by increasing the percentage of instances of responding to good performance and decreasing the percentage of instances of responding to poor performance during practice. Hypothesis 2, which predicted that the percentage of instances of responding to good (vs. poor) performance during practice would be positively related to the learning of rules for correct responses to good performance and negatively related to the learning of rules for correct responses to poor performance, was tested with Pearson product–moment correlations. A repeated measures analysis of variance was used to test Hypothesis 3, that feedback specificity would have differential effects on the learning of the rules for responding to good and poor performance. Finally, R. M. Baron and Kenny’s (1986) regression approach was used to test Hypothesis 4, that learning opportunities during practice would mediate the relationship between feedback specificity and rule learning.

In supplemental analyses, the data revealed interesting relationships worthy of further consideration and possible replication in later studies. Therefore, we performed pairwise contrasts to explore differences among the feedback specificity conditions, to help interpret the omnibus $F$ tests. These supplemental results are presented with our hypothesis tests in the Results section that follows. Adjusting Type I error rates, using a per comparison rate of $\alpha_{pc} = .017$ (i.e., .05/3 comparisons) to account for these additional tests, did not alter our overall conclusions, but as noted, one result was only marginally significant ($p = .028$). The interpretations of the effects we found in these supplemental analyses, of course, are subject to replication of results in future studies.
Results

Assessment of Attrition Effects

Procedures recommended by Goodman and Blum (1996) were used to determine whether the attrition of 9 participants between the practice phase (Time 1) and the testing phase (Time 2) was likely to affect the results of analyses performed on data from Time 2. First, using Time 1 data, we performed multiple logistic regression to assess the presence of nonrandom sampling. The independent variables for the analysis were feedback specificity (two contrast variables) and learning opportunities during practice. The dependent variable was a dichotomous variable indicating whether or not each participant came to the Time 2 testing session of the experiment. The logistic regression analysis indicated no apparent nonrandom sampling, \( \chi^2(3, N = 192) = 2.60, p = .457 \). Neither feedback specificity nor learning opportunities during practice affected whether or not people participated at Time 2. In addition, no mean or variance enhancement or reduction was found between the Time 1 whole sample and the Time 2 participants. Because the Time 2 dependent variables, good and poor performance rule learning, could not be collected at Time 1, comparison between the Time 1 whole sample and the Time 2 participants was not possible. In addition, no mean or variance enhancement or reduction was found between the Time 1 whole sample and the Time 2 participants. Because the Time 2 dependent variables, good and poor performance rule learning, could not be collected at Time 1, comparison of the differences in structural relationships could not be performed; however, the analyses we were able to conduct were not indicative of attrition bias (Goodman & Blum, 1996).

Descriptive Statistics

Correlations among the study variables plus total sample and cell means and standard deviations are shown in Table 2.

Tests of Hypotheses

Hypothesis 1 was supported. Feedback specificity affected learning opportunities during practice. The higher the feedback specificity, the greater the percentage of instances participants responded to good performance and the lower the percentage of instances they responded to poor performance during practice, \( F(2, 189) = 24.89, p = .000, \eta^2 = .22 \). As shown in Figure 2, those in the high feedback specificity condition had more instances of responding to good performance and fewer instances of responding to poor performance during practice than those in the low (\( \psi = 0.17, SE = 0.03, t[1, 191] = 6.23, p = .000 \)), and moderate (\( \psi = 0.16, SE = 0.03, t[1, 191] = 6.01, p = .000 \)), conditions. There was no difference between the low and moderate feedback specificity conditions.

Turning to our predictions regarding rule learning, Hypothesis 2 was partially supported. As predicted, the percentage of instances of responding to good (vs. poor) performance during practice was positively related to the learning of rules for correct responses to good performance (\( r = .45, p < .001 \)). However, although we predicted a negative relationship between the percentage of instances of responding to good (vs. poor) performance during practice and the learning of rules for correct responses to poor performance, these variables were unrelated (\( r = -.01, p > .05 \)).

Hypothesis 3 was supported. Feedback specificity differentially affected the learning of rules for responding to good and poor performance, as evidenced by the interaction between feedback specificity and the type of rules (good vs. poor performance rules), \( F(2, 180) = 16.05, p = .000, \eta^2 = .15 \), shown in Table 3. As depicted in Figure 3, those in the high feedback specificity condition learned best what to do when workers were performing well (i.e., good performance rule learning), followed by those in the moderate (high vs. moderate: \( \Psi = 0.19, SE = 0.04, t[1, 182] = 5.38, p = .000 \)), then low (high vs. low: \( \Psi = 0.13, SE = 0.04, r[1, 182] = 3.66, p = .000 \)), feedback specificity conditions. The difference between the low and moderate conditions was not statistically significant. Alternatively, those in the low feedback specificity condition learned best what to do when workers were performing poorly. Those in the low feedback specificity condition

<table>
<thead>
<tr>
<th>Variable</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Feedback specificity(^a)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2. Learning opportunities: % instances responding to good (vs. poor) performance during practice</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3. Good performance rule learning</td>
<td>.40***</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4. Poor performance rule learning</td>
<td>-.16*</td>
<td>-.01</td>
<td>-.09</td>
<td></td>
</tr>
<tr>
<td>Total sample(^b)</td>
<td>.70</td>
<td>.64</td>
<td>.54</td>
<td></td>
</tr>
<tr>
<td>SD</td>
<td>.17</td>
<td>.21</td>
<td>.15</td>
<td></td>
</tr>
<tr>
<td>Low feedback specificity(^c)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5. M</td>
<td>.64</td>
<td>.56</td>
<td>.56</td>
<td></td>
</tr>
<tr>
<td>SD</td>
<td>.16</td>
<td>.22</td>
<td>.16</td>
<td></td>
</tr>
<tr>
<td>Moderate feedback specificity(^d)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5. M</td>
<td>.64</td>
<td>.62</td>
<td>.54</td>
<td></td>
</tr>
<tr>
<td>SD</td>
<td>.18</td>
<td>.18</td>
<td>.14</td>
<td></td>
</tr>
<tr>
<td>High feedback specificity(^e)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5. M</td>
<td>.81</td>
<td>.75</td>
<td>.50</td>
<td></td>
</tr>
<tr>
<td>SD</td>
<td>.10</td>
<td>.19</td>
<td>.16</td>
<td></td>
</tr>
</tbody>
</table>

\(^a\) Coded 1 = low, 2 = moderate, 3 = high. \(^b\) Time 1, n = 192; Time 2, n = 183. \(^c\) Time 1, n = 67; Time 2, n = 63. \(^d\) Time 1, n = 62; Time 2, n = 59. \(^e\) Time 1, n = 63; Time 2, n = 61. \(* p < .05\). \(** p < .001\).
exhibited better poor performance rule learning than those in the high feedback specificity condition ($V = -0.06, SE = 0.03, t[1, 182] = -2.12, p = .028, \eta^2 = .03$), which was only marginally significant when compared to the adjusted per comparison error rate of $p = .017$.

Hypothesis 4 was supported for the learning of rules for responding to good performance, but not for the learning of rules for responding to poor performance. As predicted, learning opportunities during practice partially mediated the relationship between feedback specificity and the learning of rules for responding to good performance, as shown in Table 4. When this mediator was added to the regression equation, the standardized regression coefficient for feedback specificity decreased by 46% but remained statistically significant. As expected, more specific feedback led to a greater percentage of instances of responding to good performance during practice, which in turn led to greater learning of the rules associated with good performance.

Alternatively, although feedback specificity was negatively related to the learning of rules for responding to poor performance ($r = -.16, p = .03$), learning opportunities during practice did not affect the learning of the rules for responding to poor performance. Therefore, we did not find support for mediation of the feedback specificity–poor performance rule learning relationship. 1

**Discussion**

The results of this experiment demonstrate that increasing the specificity of feedback during practice affects learning opportunities during practice, and that the effects of feedback specificity on learning depend on the aspects of the task to be learned. Increasing feedback specificity benefits the learning of responses for good performance and is detrimental to the learning of responses for poor performance. The effects of feedback specificity on the learning of rules for responding to good performance are partially mediated by learning opportunities.

As predicted by Hypothesis 1, increasing the specificity of feedback during practice increasingly guided performers to correct responses, steering those who received more specific feedback into good performance and encouraging them to stay there, whereas those who received less specific feedback spent more time responding to poor performance during practice. Consequently, individuals who received the more specific feedback had more opportunities to learn what to do when things were going well and fewer opportunities to learn what to do when things were going poorly. Alternatively, those who received less specific feedback had more opportunities to learn what to do when things were going poorly and fewer opportunities to learn what to do when things were going well than those who received more specific feedback.

The positive relationship between exposure to good (vs. poor) performance and the learning of rules associated with good performance, predicted by Hypothesis 2, suggests that performers must experience good performance to learn the associated decision rules. In addition, it supports the notion that an increased frequency in the coupling of actions with positive outcomes facilitates associative learning (Anderson, 1982) and retention and

---

1 Because learning opportunities, an informational variable, did not mediate the negative relationship between feedback specificity and the learning of the rules for responding to poor performance, we investigated a motivational explanation for this relationship. Self-efficacy was measured at the end of the practice period as a control variable to test for differences in the motivational levels of participants in the three feedback specificity conditions, and to control for motivational effects if necessary. Self-efficacy did not differ significantly across the feedback specificity conditions, $F(2, 189) = 1.37, p = .256, \eta^2 = .01$. Self-efficacy and feedback specificity had an additive impact on the learning of rules for responding to poor performance, $F(3, 179) = 4.21, p = .007, \eta^2 = .04$, but self-efficacy did not mediate the relationship between feedback specificity and the learning of rules for responding to poor performance. We therefore ruled out motivation as an explanation for the effect.
recall of the rules for correct responses to good performance. Alternatively, we did not find the expected negative relationship between exposure to good (vs. poor) performance and the learning of rules for responding to poor performance. Experiencing poor performance and the associations between actions and outcomes when responding to poor performance were not enough to affect the learning of the rules. The depth and types of information processing are likely important for explaining learning to respond to poor performance. This suggests a possible interaction between opportunities to learn the rules for responding to poor performance and characteristics of information processes used when responding to poor performance, which can be tested directly in future research. Later, in our discussion of Hypothesis 4, we present several features of poor performance that might lead to different information processes for learning to respond to poor versus good performance.

The differential impacts of feedback specificity on the learning of rules for correct responses to good versus poor performance, predicted by Hypothesis 3, demonstrate that the effects of feedback interventions and the nature of what can be learned about a task are more complex than indicated by previous research. It is not that one feedback specificity intervention is necessarily better than another; it is that different aspects of a task can be learned to a greater or lesser extent depending on the specificity of the feedback provided during practice. It is therefore important to think about exactly what we want performers to learn with regard to a task when designing feedback interventions. If we want them to learn rules for responding to both good and poor performance, the feedback needs to support ample opportunities for performers to test and learn the local rules for different aspects of a task. For example, if we want machine operators to learn how to respond when machinery is working properly and when it malfunctions, the

Table 3
Repeated Measures MANOVA: Feedback Specificity by Type of Rule (Good Versus Poor Performance Rules) on the Learning of Rules for Responding to Good Versus Poor Performance

<table>
<thead>
<tr>
<th>Source</th>
<th>SS</th>
<th>MS</th>
<th>F</th>
<th>df</th>
<th>η²</th>
</tr>
</thead>
<tbody>
<tr>
<td>Between subjects</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Feedback specificity</td>
<td>0.28</td>
<td>0.14</td>
<td>4.60*</td>
<td>2, 180</td>
<td>.05</td>
</tr>
<tr>
<td>Error</td>
<td>5.43</td>
<td>0.03</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Within subjects</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Type of rule (good vs. poor)</td>
<td>1.03</td>
<td>1.03</td>
<td>32.38***</td>
<td>1, 180</td>
<td>.26</td>
</tr>
<tr>
<td>Feedback specificity by type of rule (good vs. poor)</td>
<td>1.02</td>
<td>0.51</td>
<td>16.05***</td>
<td>2, 180</td>
<td>.15</td>
</tr>
<tr>
<td>Error</td>
<td>5.73</td>
<td>0.03</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Note. N = 183. MANOVA = multivariate analysis of variance.
* p < .05. *** p < .001.

Figure 3. Effects of feedback specificity on learning as indexed by the percentages of correct responses to good and poor performance during the testing period. Vertical lines depict standard errors of the means.
There are several features of poor performance that might differ-

tegrate that rule-learning process from the process of learning of

This result suggests that the explanation for learning how to respond to poor performance may be more complex than the explanation for learning how to respond to good performance. There are several features of poor performance that might differentiate that rule-learning process from the process of learning of rules associated with good performance. First, compared to good performance, poor performance generally places greater pressure on the individual to change his or her current line of actions and therefore requires deeper and more effortful processing of feedback information plus the consideration of alternatives. Exposure to poor performance during practice may be necessary, but insufficient, for learning rules for responding to poor performance. What performers do with those opportunities, how they search and process information, may be more important or at least different in poor performance versus good performance conditions. Second, poor performance is more likely to be judged as failure than good performance and to lead to negative self-evaluations that interfere with effective information processing and problem solving (Bandura, 1997). Future research should further address how people learn under poor performance conditions by examining, for example, how they search for information, for what information they search, and how information search and processing differ under conditions of good versus poor performance.

Contributions, Strengths, and Limitations

This experiment contributes to current understanding of feedback intervention effects on learning processes and outcomes. At a theoretical level, we have begun to unpack and examine the intervening processes through which feedback interventions affect learning. This appears to be a promising avenue for further research, along the lines suggested above. From a practical perspective, our results suggest that prescriptions about feedback specificity need to be qualified to reflect that increasing specificity can facilitate or inhibit learning, depending on what is to be learned. Further, the common suggestion that people should be given more specific feedback in the initial stages of learning a task should also be qualified.

Our research design choices present some strengths and limitations in regard to the generalizability of our findings. The simulated decision environment used in this experiment allowed for the control necessary for high internal validity, as well as the ability to objectively manipulate feedback specificity and measure learning
opportunities and learning. Our manipulation represented a range of feedback specificity levels by varying the content of task-relevant information in a feedback message in a way that was independent of the setting or source of the feedback intervention (Annett, 1969). This level of experimental control is appropriate when the aim is to identify the explanatory mechanisms for observed relationships, such as the differential effects of feedback on learning the rules for responding to good versus poor performance.

This emphasis on internal validity in our design raises the obvious question of whether these results will generalize to other more natural settings and other sources, particularly human sources, where the specificity of the feedback message will interact with other features of the source and the situation. Previous laboratory research on the effects of objective feedback has shown that feedback effects generalize to natural work settings, where the effects are often amplified when feedback interacts with variables that remain uncontrolled in field settings (Kopelman, 1986). This bodes well for the generalizability of feedback specificity effects. Future research to establish the generalizability of feedback specificity effects should focus on the interactions between feedback specificity and features of the feedback source, other dimensions of the feedback, and properties of the situation.

Of course, the external validity of any set of findings needs to be judged along multiple dimensions (Locke, 1986; Mook, 1983). In this study, we tested the common generalization that "the more specific feedback the better" and show that it does not always apply. We also provide some generalizable results about the differential effects of feedback specificity on learning good versus poor performance rules. The decision-making simulation used in this study was chosen to represent a team management task, as a means of engaging participants with the task, but we would not argue without further research that the results will generalize to team settings. The task was chosen because it incorporates the type of dynamic and complex information environment that characterizes many managerial and other types of decision-making tasks, in which there are multiple decision variables to be considered and multiple effects present in observed feedback. We can think of no reason to expect different results for many other complex tasks, particularly those for which rules are contingent on performance and in which people have to explore to learn which decision rules work and when. Further, as in our simulation, computers are a primary source of feedback in computer-based training, software programs, and high-tech equipment in organizations. So, although it is unknown whether our results will generalize to human feedback sources, they should apply to the numerous natural settings in which computers are primary feedback sources. Given the ease with which frequent, specific, immediate feedback can be built into computerized tasks and a tendency to believe that "more is better" when it comes to feedback, our results have important implications for decisions regarding the design of computer-provided feedback interventions.

On the limitations side, the decision options were well structured and readily available to the individuals working on the simulation. Also, the information received from the simulation is specific and timely and is missing the emotions and other content that are often included in information about employees and in the feedback received in many work group settings. Future research needs to consider how variations in the structure of the social context surrounding the decision task and other features of the feedback affect learning. After additional basic knowledge is gained in this area, future research should be performed in field settings.

Conclusion

Overall, our results show that those who receive feedback of varying specificity learn different things, through different means. Simple notions about feedback being beneficial or detrimental to learning need to be augmented by more complex models of the mechanisms by which varying feedback specificity creates conditions for learning different aspects of a task, partially through its impact on learning opportunities during practice. Future research should continue to explore the mechanisms by which variation in feedback interventions affects learning, with particular attention to how people learn the rules for responding to poor performance and differences in learning processes for good versus poor performance.

References


Goodman, J. S. (1998). The interactive effects of task and external feed-


Received September 17, 2002
Revision received September 23, 2003
Accepted October 20, 2003

New Editor Appointed for Journal of Occupational Health Psychology

The American Psychological Association announces the appointment of Lois E. Tetrick, PhD, as editor of Journal of Occupational Health Psychology for a 5-year term (2006–2010).

As of January 1, 2005, manuscripts should be submitted electronically via the journal’s Manuscript Submission Portal (www.apa.org/journals/ocp.html). Authors who are unable to do so should correspond with the editor’s office about alternatives:

Lois E. Tetrick, PhD
Incoming Editor, JOHP
George Mason University
Department of Psychology, MSN, 3F5
4400 University Drive, Fairfax, VA 22030

Manuscript submission patterns make the precise date of completion of the 2005 volume uncertain. The current editor, Julian Barling, PhD, will receive and consider manuscripts through December 31, 2004. Should the 2005 volume be completed before that date, manuscripts will be redirected to the new editor for consideration in the 2006 volume.