Feedback Specificity, Exploration, and Learning

Jodi S. Goodman  
University of Connecticut

Margaretha Hendrickx  
Binghamton University

Robert E. Wood  
University of New South Wales

Greater feedback specificity is generally considered to be beneficial for performance and learning, but the evidence for this generalization is limited. The authors argue that increasing the specificity of feedback is beneficial for initial performance but discourages exploration and undermines the learning needed for later, more independent performance. The results of their transfer experiment demonstrate that increasing the specificity of feedback positively affected practice performance, but its benefits did not endure over time or modification of the task. In addition, feedback specificity negatively affected levels of exploration during practice and interacted with exploration strategies to affect learning. The results suggest that those who received feedback of varying specificity may have learned through different but equally beneficial mechanisms.

The belief that greater specificity in feedback leads to improved performance and learning has become an accepted generalization, despite a lack of evidence to support the argument on which the prescription is based. Textbook discussions of supervisory activities such as performance appraisals (e.g., Cascio, 1998) and coaching and development (e.g., Gomez-Mejia, Balkin, & Cardy, 2001; Kreitner & Kinicki, 2001) routinely prescribe that effective performance feedback must be specific and timely. Textbook prescriptions often lack detail on what the specificity part means in practice, although the purpose appears to be captured by the argument 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” (Bernardin & Beatty, 1984, p. 197). That is, effective feedback is immediate and includes information on the behaviors that were performed incorrectly and how to correct them. However, is specific feedback truly beneficial? Based on the limited, available evidence, the answer appears to be “not always” (Annett, 1969; Kluger & DeNisi, 1998).

Feedback specificity refers to the level of information presented in feedback messages (Annett, 1969; I.L. Goldstein, Emanuel, & Howell, 1968). In its capacity as an information source (J.A. Adams, 1987; Anderson, 1982; Annett, 1969; Ilgen, Fisher, & Taylor, 1979; Payne & Hauty, 1955; Vroom, 1964), feedback “signals error information which releases corrective action” (Annett, 1969, p. 121), guiding performers to the correct response by allowing them to determine which behaviors are appropriate or inappropriate for successful performance. As feedback specificity increases, so does its capability to perform its information role (Payne & Hauty, 1955). There is a clear body of evidence which shows that specific, objective feedback that is matched to the performance criterion leads to higher performance than less specific, more subjective feedback (Kopelman, 1986). However, the available research evidence is much more equivocal on two other issues of practical importance. First, does increasing the information content of feedback messages up to the point of including a statement of corrective responses lead to increases in performance? The question of how much information to provide when giving feedback has direct relevance for the activities of managers and others who provide feedback in performance appraisal meetings, coaching and development activities, and daily acts of leadership. Current textbook prescriptions to be specific and the available research do not tell supervisors how specific they should be. Highly specific feedback that tells people “how to do it” when they are having difficulties should clarify expectations and provide structure to facilitate performance, but increasing the specificity of feedback has associated time and effort costs for those who must conduct the diagnoses and provide the information and may lead to negative reactions from recipients.

A second, and we believe more important, practical issue concerns the impact of increasing feedback specificity on performance in the longer term and when the supplemental information presented in more specific feedback is removed and a task has to be performed more independently, with less specific or no feedback. That is, does increasing the specificity of feedback aid or impede learning? The answer to this question is relevant for many situations, including those in which feedback is provided to new staff,
staff who are performing poorly and staff who have been recently promoted or assigned new tasks. For example, which of the following should a manager tell a staff member? (a) Your work group has performed at 90% of standard. (b) Your work group performed at 90% of standard because you allocated staff to the wrong jobs and failed to set realistic goals for them. To hit target, you need to assign Jack to Job X and Jane to Job Y and set both of them goals that are within 20% of their last month’s performance.

The latter, highly specific feedback example provides the information recommended by Bernardin and Beatty (1984), but it also removes the need for the recipient to search for or infer the correct responses. The less specific feedback in the former example provides no detail about the errors made or the corrective actions for improving performance. The recipient must discover these through his or her own exploration and information processing. Therefore, while the more specific feedback should lead to more immediate improvements in performance than the less specific feedback, it may also discourage the exploration processes that lead to learning and the development of the staff member’s capability to perform without the continuing support of the manager.

Our primary aim with this study was to test the effects of increasing feedback specificity on practice performance, exploration, and learning to investigate the generalization that effective feedback must be specific and begin to explore the mechanisms by which specific feedback operates during practice. We did this by varying the objective task-relevant information provided in a feedback intervention. The managerial decision-making task, which is described in the Method section, 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 a record of all decisions made, which can be used to objectively operationalize exploration patterns.

Development of Research Hypotheses: Effects of Feedback Specificity on Practice Performance, Exploration, and Learning

In this section, we develop the rationale for hypotheses regarding the effects of feedback specificity on practice performance, the level of exploration during practice, and learning. We also argue that systematic and unsystematic exploration differentially affect learning, that systematic exploration partially mediates the relationship between feedback specificity and learning, and that feedback specificity moderates the relationships between different types of exploration and learning.

The impact of feedback interventions on performance during practice or training is well studied in organizational behavior (cf. Kluger & DeNisi, 1996). Feedback interventions that direct attention to correct on-task behavior typically lead to performance improvements and decreased errors during practice and more rapid training. These effects are positively related to the frequency, immediacy, and focus (i.e., process vs. outcome feedback) of feedback (Earley, 1988; Earley, Northcraft, Lee, & Lituchy, 1990; I.L. Goldstein et al., 1968; Kim, 1984; Kluger & DeNisi, 1996; Payne & Hauty, 1955; Schmidt, 1991; Schooler & Anderson, 1990; Weinstein & Schmidt, 1990). They are also positively related to specificity when objective outcome feedback is compared with more ambiguous, subjective feedback (Kopelman, 1986). When outcome feedback is supplemented with additional task-relevant information that facilitates the identification of correct responses to that feedback, we expect it will have an immediate impact on performance. The more specific the information provided, up to the point of identifying the correct responses for the feedback recipient, the greater the immediate performance enhancing effects of feedback should be.

With few exceptions (e.g., Goodman, 1998), studies by organizational behavior researchers have focused on feedback effects on immediate performance and have not considered if and how feedback effects generalize over time and different versions of a task or how well effects endure on the removal of supplemental feedback. Feedback effects that endure over time or changes in conditions provide more convincing evidence for the relatively permanent acquisition of skills, understanding, and knowledge than practice performance (Christina & Bjork, 1991; Salorni, Schmidt, & Walter, 1984; Tolman, 1932/1967). Frequent and immediate feedback, which is known to have beneficial effects for practice performance, can undermine learning. Alternatively, feedback interventions that provide less frequent and delayed feedback and feedback summarized over multiple trials have been found to lead to poorer practice performance but better learning (Archer, Kent, & Mote, 1956; Christina & Bjork, 1991; Goodman, 1998; Schmidt, 1991; Schooler & Anderson, 1990; Weinstein & Schmidt, 1990).

For all but the simplest of tasks, learning the correct actions to take in response to different levels of performance and different situations typically requires some active search for and processing of information about the relationships between actions and outcomes (Klahr & Dunbar, 1988; Simon & Lea, 1974). When people repeat tasks over time and receive feedback on their performance, the search for information about the impacts of different actions is conducted through the testing of alternative responses. As tasks become more complex, the levels of exploration and processing required to develop effective representations of the decision rules for tasks increase (Klahr & Dunbar, 1988; Simon & Lea, 1974).

However, as the information provided in a feedback intervention becomes more specific, fewer exploration and information-processing activities are required to identify and select responses, and individuals are less inclined to explore the effects of different options (cf. Schmidt, 1991). Individuals come to rely on specific feedback for guidance, and they are less likely to remember previous responses (Christina & Bjork, 1991; Schmidt, 1991; Schmidt & Bjork, 1992). Essentially, the feedback does the work for performers, making it seemingly unnecessary for them to engage in the exploration, information-processing, and recall activities essential for learning. Specific feedback that guides individuals to correct responses also decreases errors that accompany experimentation with different options. On the surface this may seem beneficial; however, errors during practice can be a useful tool for learning. Errors can enhance one’s mental model of a task by leading to new insights and creative solutions (Frese et al., 1991). Errors also create opportunities to learn what to do when things are going wrong. By decreasing opportunities to experience errors, the guidance provided by specific feedback likely decreases the extent to which individuals learn how to correct errors and recover from poor performance.
As we have argued, we expect increases in feedback specificity to be accompanied by decreases in exploration activities. There are multiple types of exploration, which we discuss next, and while we expect the guidance provided by specific feedback to decrease the levels of all types, the different types of exploration will vary in their impacts on learning. Different types of exploration activities are not equally effective for the discovery of the correct decision rules for linking actions to outcomes, the impacts depend upon how exploration is done (Debowski, Wood, & Bandura, 2001). When responding to feedback on a multitrial task, individuals can explore by testing the effects of different responses or they can persist with existing strategies. When testing the effects of different responses over time or trials, how one looks or explores will determine the volume and content of information received and what can be learned from the processing of that feedback. Naturally occurring tasks typically require the sequential processing of feedback and the accumulation of information about the relationships between actions and outcomes. In general, the more systematic one is in the exploration process, the less confounded the information found, and the more one can learn from the feedback.

Varying one thing at a time (VOTAT; Tschirgi, 1980) is one example of a systematic exploration process that seeks to maximize the information value of feedback and facilitates the accumulation of information about the impacts of specific actions across trials. Changing one variable at a time while holding other variables constant provides unconfounded feedback on the impacts of the variable changed. In a study of goal setting, Vollmeyer, Burns, and Holyoak (1996) found that general, as opposed to specific, goals led to greater use of the VOTAT strategy, which resulted in superior learning. Alternatively, changing more than one thing at a time is a poor exploration strategy because it leads to confounded information. When more than one variable is changed simultaneously, it is not possible to determine which variable(s) led to an effect or whether the positive and negative effects of different variables canceled each other out. Results of confounded experiments will not be as useful for modifying or developing hypotheses about the relationships between actions and outcomes.

When individuals choose not to explore, or to persist with an existing set of responses, the implications for learning are less clear-cut. Of course, if individuals do not explore at all, they likely will learn little or nothing. However, if used sparingly, holding steady, or changing nothing across several trials of a task, is a strategy that may be beneficial to learning. Holding steady with a strategy may help create stable memory representations (Schmidt, 1991; Wulf & Schmidt, 1994) and add to evidence about the strategy that may be beneficial to learning. Holding steady with a strategy may help create stable memory representations (Schmidt, 1991; Wulf & Schmidt, 1994) and add to evidence about the strategy that may be beneficial to learning.

When responding to feedback on a multitrial task, individuals can explore by testing the effects of different responses or they can persist with existing strategies. When testing the effects of different responses over time or trials, how one looks or explores will determine the volume and content of information received and what can be learned from the processing of that feedback. Naturally occurring tasks typically require the sequential processing of feedback and the accumulation of information about the relationships between actions and outcomes. In general, the more systematic one is in the exploration process, the less confounded the information found, and the more one can learn from the feedback.

Varying one thing at a time (VOTAT; Tschirgi, 1980) is one example of a systematic exploration process that seeks to maximize the information value of feedback and facilitates the accumulation of information about the impacts of specific actions across trials. Changing one variable at a time while holding other variables constant provides unconfounded feedback on the impacts of the variable changed. In a study of goal setting, Vollmeyer, Burns, and Holyoak (1996) found that general, as opposed to specific, goals led to greater use of the VOTAT strategy, which resulted in superior learning. Alternatively, changing more than one thing at a time is a poor exploration strategy because it leads to confounded information. When more than one variable is changed simultaneously, it is not possible to determine which variable(s) led to an effect or whether the positive and negative effects of different variables canceled each other out. Results of confounded experiments will not be as useful for modifying or developing hypotheses about the relationships between actions and outcomes.

When individuals choose not to explore, or to persist with an existing set of responses, the implications for learning are less clear-cut. Of course, if individuals do not explore at all, they likely will learn little or nothing. However, if used sparingly, holding steady, or changing nothing across several trials of a task, is a strategy that may be beneficial to learning. Holding steady with a strategy may help create stable memory representations (Schmidt, 1991; Wulf & Schmidt, 1994) and add to evidence about the strategy that may be beneficial to learning.

Hypothesis 1A: The level of systematic exploration (VOTAT) used during practice will be positively related to learning.

Hypothesis 1B: Feedback specificity will be negatively related to practice performance.

Hypothesis 2: Feedback specificity will be negatively related to the levels of systematic and unsystematic exploration during practice.

Hypothesis 3: The relationship between exploration and learning will depend on the nature of exploration.

Hypothesis 3A: The level of systematic exploration (VOTAT) used during practice will be positively related to learning.

Hypothesis 3B: The level of unsystematic exploration used during practice will be negatively related to learning.

Hypothesis 4: The negative impact of feedback specificity on learning will be partially mediated through its negative effects on the level of systematic exploration during practice.

The partial mediation prediction in Hypothesis 4 is based on two considerations. First, the impacts of feedback specificity on exploration and learning are yet to be investigated and, while we believe and have argued that exploration is a major causal mechanism for the effects of feedback on learning, we also recognize that there are other possible causal mechanisms. For example, level of information-processing and memory effects might have effects independent of exploration. Second, as we have argued, the effects of exploration on learning will depend on the types of exploration conducted. As we argue next, as feedback and exploration are both sources of information about the task, the specificity of feedback and the type of exploration will interact in their effects on learning. Thus, in relation to the effects of feedback specificity on learning, we propose a partial mediational role for the level of exploration during practice and an interaction effect for type of exploration.

When individuals are making decisions about task actions, the utility of information available to them will be an interaction between the information provided in feedback interventions of differing specificity and the clarity of information provided by different exploration strategies. Different exploration strategies should be more or less beneficial for or detrimental to learning when coupled with the different levels of specific information in feedback interventions. Therefore, while unsystematic exploration will be detrimental to learning because of the confounding of effects for different actions in the outcome information received,
more specific feedback interventions that inform performers about correct and incorrect behaviors can be used to make inferences about which actions affected performance. Thus, specific feedback on correct and incorrect behaviors coupled with the multiple behavior changes may aid in the discovery of the correct decision rules for a task, making ordinarily uninformative cues gathered from unsystematic exploration informative and useful for learning. Alternatively, the positive effects of systematic exploration on learning should be greater as feedback specificity decreases. With less specific feedback interventions, the information from unconfounded tests in systematic exploration is imperative for identifying the correct decision rules.

Hypothesis 5: The relationships between exploratory behavior and learning will differ across feedback specificity conditions.

Hypothesis 5A: The positive relationship between systematic exploration and learning will decrease as feedback specificity increases.

Hypothesis 5B: Unsystematic exploration will be negatively related to learning when feedback specificity is lower and positively related to learning when feedback specificity is higher.

The hypotheses are summarized in the model in Figure 1.

Method

Overview

Participants completed an 18-trial practice task under one of four (nested) levels of feedback specificity, which ranged from outcome-only feedback at the low specificity end to specific guidance on the correct responses for decisions at the highly specific end. During the practice period, all participants were informed that their goal was to learn the decision rules underlying the task. Two days later, participants completed a 6-trial testing task. The task completed in the testing period was more complex than the practice task. All participants received outcome feedback in the testing period.

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 and sixty-one individuals completed the practice task. Participants returned 2 days later and completed the testing task, which was used to assess the learning that occurred at Time 1. Twelve participants were lost because of attrition from Time 1 to Time 2, reducing the Time 2 sample size to 149 (93%). Participants were 55.3% men, 44.7% women and ranged in age from 19 to 30, with a mean age of 21.20 (SD = 1.84). We assessed participants’ management-related experience with a five-item, 5-point scale, with responses ranging from 1 (none) to 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. Participants had a moderate amount of previous management-related experience (M = 2.77, SD = 0.74, α = .84).

Experimental Design

We conducted a single-factor experiment, with repeated measures, to test the effects of feedback specificity (four levels) on practice performance, exploration, and learning. Participants were randomly assigned to one of four treatment conditions: low, moderate, high, or very high feedback specificity. The manipulations are described below.

The Task

The study was presented as a project in managerial decision making. Participants served as special order department managers 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). 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 representing a week at the Furniture Factory. They assigned the workers to jobs and motivated them by providing goals, feedback, and rewards.

Participants received written instructions describing the simulation task, the skills and work-related preferences of each employee, the descriptions and requirements of each Furniture Factory job, and choices of types of goals, feedback, and rewards that they could give to their employees. The instructions were presented on the computer screen and on a reference sheet to which participants could refer throughout the simulation. To learn how to manage the group, they had to figure out the decision rules underlying the task structure, which are outlined in Table 1 and described below.

Critical features of the simulation contributed to its realism and the generalizability of our findings to other complex tasks. Like many complex
tasks, the simulation has underlying decision rules for linking actions to outcomes that have to be discovered through experimentation and information processing. In addition, the simulation requires multiple decisions about multiple employees and the learning of multiple rules regarding the effects of different decisions. This is more representative of the task situation in which performers typically receive feedback than a task on which feedback relates to a single action with a simple cause–effect relationship to a performance criterion.

Other simulation characteristics were designed to contribute to the participants’ experienced realism to foster their motivation to try to learn the task. First, employees differed in their skills and preferences to perform each job and in their overall motivation levels. Optimal matches of employees to jobs were based on fit between employee and job characteristics. In addition, participant managers’ decisions influenced overall motivation levels, which affect the magnitude of the effects of decision options for goals, feedback, and rewards on employee performance. Second, the choices for goals, feedback, and reward decisions, listed in the next paragraph, represented the types of actions managers might take in an actual organization. Third, the logic of the differential effects of decision options on employee performance was based on extant organizational behavior literature (viz., goal setting, feedback, expectancy, and equity theories and research). Fourth, decision choices made during one trial influenced performance on subsequent trials, thus modeling the temporal effects of such actions in actual work environments. Finally, standard performance estimates were derived from time and motion studies of furniture factory jobs. The standard performance estimates were used as points of reference for feedback to participants on their managerial performance and for production goals and feedback to Furniture Factory workers.

Participants made four decisions for each worker on each simulation trial. First, they 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. Next, participants set a goal for each worker from a set of five options. Participants could choose to give no goal, tell an employee to do his or her best, set a minimum goal of 25% worse than the standard derived from time and motion studies (easy goal), set a minimum goal equal to standard (moderate goal), or set a minimum goal of 25% better than standard (difficult goal). Participant managers then could choose among four types of performance feedback to provide each employee. They could choose to provide no feedback, discuss with the employee what he or she did correctly and incorrectly when performing the job (process feedback), inform the employee of his or her performance level in relation to the standard for the job (outcome feedback), or provide both process and outcome feedback. They then chose from a set of three reward levels for each employee. They could choose to give no reward, praise an employee (moderate reward), or publicly recognize an employee by posting a memo (high reward). Job assignment, goal, feedback, and reward decisions were made for each employee 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 decision rules regarding the best options for the job allocation, goal, feedback, and reward decisions through interacting with the task and interpreting the feedback they received. The relationship between participants’ decisions and their employees’ performance was governed by a set of decision rules, outlined in Table 1, that were 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 decision rules. To learn the decision rules shown in Table 1 and described below, participants had to test options, cognitively process the feedback they received as a result of their decisions, and continue to apply analytic strategies in ways that would reveal the governing rules. Individual employee performance was computed using the mathematical model for the decision rules (see Wood & Bailey, 1985, for the mathematical model); group performance was a simple addition of the performance of individual workers.

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

The simulation also began with both outcome and process feedback as the optimal feedback choice. This continued until an employee demonstrated job competence, operationalized as performing at or better than standard for three consecutive trials. At this time, process feedback could be viewed as oversupervision; thus outcome feedback alone becomes the optimal feedback choice. If the participant manager changed the job assignment, both outcome and process feedback was again the optimal choice until the employee 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 decisions regarding rewards were based on an employee’s performance on the current trial. If an employee performed poorly (i.e., below standard), no reward would be given. If an employee performed moderately well (i.e., standard ≤ performance < 5% better than standard), a moderate reward would be given. If an employee performed well (i.e., ≥ 5% better than standard), a high reward would be given. In addition, the mathematics linking chosen reward options to employee performance included a comparison of the reward with performance ratios of the different employees, as outlined in equity theory (J.S. Adams, 1963). As a result of this comparison, the impacts of chosen reward options on indi-

---

### Table 1

**Decision Rules for Simulation**

<table>
<thead>
<tr>
<th>Rule</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<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 a difficult goal initially. Give a moderate goal after an employee performs very poorly for 2 consecutive weeks ((\leq 20%)) worse than standard. Give a difficult goal after an employee performs well ((\geq)standard). Giving no goal or an easy goal is never optimal.</td>
</tr>
<tr>
<td>Feedback</td>
<td>Give both outcome and process feedback initially. Give only outcome feedback after an employee performs well ((\geq)standard) for 3 consecutive weeks. Giving no feedback or only process feedback is never optimal.</td>
</tr>
<tr>
<td>Reward</td>
<td>Give an employee no reward for poor performance ((&lt;)standard). Give a moderate reward when performance is close to standard (standard (\geq)performance &lt;5% better than standard). Give a high reward for good performance ((\geq 5%) better than standard).</td>
</tr>
</tbody>
</table>

---
individual employees were partially dependent on rewards given to other employees.

**Manipulations**

Four nested feedback interventions of differing specificity were created. The information provided ranged from the outcome feedback from the simulation, in the low specificity condition, to instructions on the correct responses for the actions taken, in the very high feedback specificity condition. Feedback was provided via computer and 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 to direct attention to the feedback.

**Low feedback specificity.** In the low feedback specificity condition, participants received outcome feedback for each of the three employee’s job performance and for the group. The outcome feedback at the individual and group levels included the hours taken 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 and provided enough information for participants to quickly assess how each employee and the group were performing relative to the estimated standards for the jobs and the total order, respectively.

**Moderate feedback specificity.** In addition to the outcome feedback received in the low feedback specificity condition, participants in the moderate feedback specificity condition were provided error signal feedback about their decisions. For each set of decisions about job allocation, goals, feedback, and rewards, they were informed that either their decisions were made correctly for all employees or their decisions were incorrect for at least one employee. For example, if a participant had given the correct goal to all employees on a particular trial, he or she was told, “You gave every employee 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 employee the wrong goal.”

**High feedback specificity.** Participants in the high feedback specificity condition were provided outcome feedback plus more specific error signal feedback for each decision made for each of their employees. Participants were told whether the job, goal, feedback, and reward given to each worker were right or wrong. Therefore, in addition to the outcome feedback, the information they received included a series of statements such as “You gave Jack the wrong goal” and “You gave Mary the correct feedback.”

**Very high feedback specificity.** In addition to the outcome and error signal feedback provided in the high feedback specificity condition, participants in the very high feedback specificity condition were instructed on the correct response whenever they made a decision error. For example, if a participant gave “Jack” a high reward in the face of poor performance, he or she was told, “You gave Jack the wrong reward. You should have given no reward.”

**Procedure**

To study the effects of feedback specificity on practice performance, exploration, 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 four feedback specificity conditions described previously (i.e., low, moderate, high, or very 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 to 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 18 trials (or “work weeks”) of the management simulation task to allow for practice, performance improvement, and learning. All participants were instructed that their goal during the practice period was to learn the decision 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 occurred during the Time 1 practice period. Participants performed six trials during which they were provided outcome feedback regarding performance levels for each of their workers and for the group as a whole—the same as that provided to participants in the low-sensitivity feedback condition during the practice period. Task complexity was increased during the testing period, with participants supervising five, rather than three, workers. In addition, there was no overlap between the groups of workers supervised at Times 1 and 2. However, the decision rules and mathematical model that linked the decision choices to performance outcomes remained the same. Any decision rules learned during the practice period could, therefore, be applied directly to the management of this new group of workers. Task complexity was increased, and workers were changed to increase the chances that testing performance was indicative of rule learning rather than memorization and repetition of decisions made at Time 1. Participants were told that their goal on this second task was to maximize the performance of their work group over multiple trials of the simulation.

Figure 2 is an example of what the participants saw on the computer screen during the simulation. This particular screen shot shows what the participant saw just after making a decision about job allocations. After job allocation decisions were made, the options were shown for goal assignment. After goals were chosen, the compared-to-estimate column indicated each employee’s performance on that trial, and additional feedback appeared below the matrix, its content dependent on the feedback specificity treatment condition in which the participant was randomly assigned. Feedback options were then presented to the participant, followed by the reward options. As the participant made decisions, his or her choices appeared in the corresponding spaces in the matrix. The participant did not actually see his or her employees performing the jobs, as is often the case in organizations.

**Measures**

**Practice performance.** Practice performance was operationalized as the percentage of hours above or below the previously set standard for the group of employees for each of the 18 practice period (Time 1) trials. It was computed as follows: practice performance = –[actual no. of hours for task completion by the group / standard no. of hours] – 1. For example, if a group of employees took 35 hr to complete a furniture order that should have taken 32 hr on the basis of time and motion studies, we computed the group’s performance as –1(35/32 – 1) = –.0938 or 9.38% worse than standard. Zero percent means that performance was equal to the standard, 10% means performance was 10% better than the standard, and –20% means performance was 20% worse than the standard.

**Learning.** Learning was the percentage of hours above or below the previously set standard for the group of employees for each of the six trials during the testing period (Time 2) and was assessed using the same formula as described for practice performance. 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 that generalize beyond the practice period or transient effects that are only evident under the practice conditions (Christina & Bjork, 1991; Schmidt & Bjork, 1992). Transient effects are temporary. There is an inherent confound between transient effects and more permanent 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). Through the transfer design, we were able to eliminate any transient effects and identify the learning effects that were due to feedback specificity during practice.

**Exploration.** Exploration was operationalized through analysis of changes in decisions for job, goal, feedback, and reward allocation from trial to trial for each worker during the practice period. For each of three workers, participants could make a systematic, unsystematic, or no change in their decisions about jobs, goals, feedback, and rewards. The three strategies described next—systematic exploration, unsystematic exploration, and hold steady—define the complete set of exploration options for the decisions made for each employee on each trial. For each employee within each trial, they are both exhaustive and mutually exclusive.

1. **Systematic exploration.** Systematic exploration was operationalized in a manner consistent with the logic of hypothesis testing within an unconfounded experimental design (Tschirgi, 1980; Vollmeyer et al., 1996); that is, changing only one decision per trial per employee. Within the design of the model, as described earlier, changing just one factor for a worker was equivalent to making an unconfounded change and thus allowed the participant to unequivocally identify the impact of the change made on the worker’s performance. When more than one factor was changed on a given decision cycle, the resulting feedback for that worker included the confounded effects for all of the factors changed. During the practice period, it was possible to engage in systematic exploration between 0 and 3 times per trial (up to one decision change for each of three employees). Across the entire practice period, it was possible to change decisions a minimum of zero times and a maximum of 48 times (3 employees \(\times\) 16 trial-to-trial opportunities for decision changes).

2. **Unsystematic exploration.** Unsystematic exploration was defined as engaging in confounded experiments and was operationalized as the number of times two or more decisions were changed for each employee across trials. During the practice period, it was possible to engage in unsystematic exploration between 0 and 3 times per trial (two, three, or four decision changes for each of three employees) and a maximum of 48 times across the practice period.

3. **Hold steady.** The hold steady strategy was operationalized as the number of times no decision changes were made from trial to trial for each employee. During the practice period, it was possible to engage in the hold steady strategy between 0 and 3 times per trial, with a maximum of 48 times across the practice period.

**Feedback specificity manipulation check.** Twenty doctoral students ranked and rated the specificity of the feedback in the four experimental conditions. The raters had no knowledge of the purpose of the experiment or of the hypotheses. The raters were given our definition of feedback specificity and examples of the feedback for each of the four conditions. All 20 doctoral students ranked the four conditions in order of specificity consistent with our intended order from low to very high. The ratings, which were made on a 5-point scale, anchored by 1 (low specificity) and 5 (high specificity), were also consistent with the intended manipulations. The mean ratings differed significantly, \(F(3, 76) = 351.94, p = .000\), from one another in the expected order: low \(M = 1.10, SD = 0.31\); moderate \(M = 2.30, SD = 0.47\); high \(M = 3.80, SD = 0.41\); and very high \(M = 4.85, SD = 0.37\). The results of this manipulation check show that the doctoral students agreed with the objective distinctions in the information content of the four feedback specificity conditions.

**Perceived feedback specificity.** Two items were developed to assess perceptions of the specificity of the objective feedback provided to participants \((r_{kk} = .80)\). These were completed at the end of the Time 1 practice session. The items were as follows: (a) “I received detailed feedback about my performance as Special Order Manager” and (b) “I was given specific feedback about my performance as Special Order Manager.” These items were measured on a 5-point Likert scale with responses ranging from 1 (strongly disagree) to 5 (strongly agree).

Perceived feedback specificity was measured for use as a control variable in the assessment of subjective feedback effects on learning. This was done to ensure that any observed effects were due to the subjective differences in the task-relevant information provided in the feedback interventions and were not due to differences in the encoding of that feedback. However, because perceived feedback was unrelated to learning \((r = .00)\), it was unnecessary to use it as a statistical control variable. Objective differences between the feedback specificity interventions were partially represented in the participants, perceptions of feedback specificity, \(F(3, 157) = 9.65, p = .000\). Participants in the high and very high conditions perceived the feedback they received to be more specific than did those in the low and moderate conditions: estimate \(\Psi = 1.67, SE = 0.32, t(157) = 5.19, p = .000\); \(\eta^2 = .15, 95\%\) confidence interval (CI) = 1.03 < estimate \(\Psi < 2.30\). Other differences were not statistically significant.

**Data Analysis**

To test Hypotheses 1A and 1B, a repeated-measures analysis of variance (ANOVA) was used to assess the between-subjects and linear within-subject effects of feedback specificity on practice performance (measured at Time 1, practice period) and learning (measured at Time 2, testing period), respectively. For Hypothesis 2, a repeated-measures ANOVA was
used to test the effects of feedback specificity on systematic and unsystematic exploration. For Hypothesis 4, we planned to use the Baron and Kenny (1986) method to test for the partial mediation of systematic exploration on the relationship between feedback specificity and learning. However, feedback specificity did not impact learning, so our data contradict the partial mediation hypothesis. To test Hypotheses 3A, 3B, 5A, and 5B, an ANOVA was used to assess the effects of exploration on learning and the interactive effects of feedback specificity and exploration on learning. For Hypotheses 3 and 5, learning was averaged across all six Time 2 testing period trials to yield a reliable measure of learning. The exploration variables were centered to decrease colinearity between the main and interaction effects associated with exploration. In addition, observed high correlations for unsystematic exploration and hold steady with practice performance ($r = - .83$ and $r = .82$, respectively) were due to the reciprocal relationships between the decisions made by participants and their performance on the task. Participants who were not performing well on trial 1 unsystematically changed their decisions on trial 1 + 1, which led to poor performance on trial 1 + 2 and so on. Participants who were performing well on trial 1 were less likely to change their decisions on trial 1 + 1. They held steady with their correct decisions on trial 1 + 1, which led to high performance on trial 1 + 2 and so on. Our concern was with how the different exploration strategies captured by the pattern of decisions made during practice influenced learning and not their concurrent relationships with practice performance. To ensure that any relationships between exploration and learning were not attributable to other determinants of practice performance, we included practice performance as a covariate in tests of Hypotheses 3 and 5. To decrease multicolinearity in our analyses, we used early practice performance, performance averaged across the first three practice trials, as our control variable instead of performance for all 18 of the practice trials. The bivariate correlations between early practice performance and unsystematic exploration and hold steady were $r = .41$, $p < .01$ and $r = .39$, $p < .01$, respectively.

In supplemental analyses, the data revealed interesting relationships worthy of further consideration and possible replication in later studies. Therefore, we performed orthogonal 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. We report the sample contrast coefficients (estimate $\Psi$), which are the weighted differences between the means being compared (Hays, 1988), the standard errors of the coefficients (SE), and the associated test statistics ($t = \text{estimate} \ \Psi / \text{SE}$), effect sizes ($\eta^2$), and 95% CIs. Significant $t$ tests indicate significant differences between means being compared. Adjusting Type I error rates, using a per comparison rate of $\alpha_{pc} = .017$ (i.e., .05/3 orthogonal comparisons), to account for these additional tests, did not alter our results or conclusions. 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 12 participants was likely to affect the results of analyses performed on data from Time 2. First, using Time 1 data, multiple logistic regression was performed to assess the presence of nonrandom sampling. The independent variables for the analysis were feedback specificity (three dummy variables) and Time 1 practice performance, systematic exploration, unsystematic exploration, and hold steady averaged or summed across all practice period trials. The dependent variable was a dichotomous variable indicating whether each participant came to the Time 2 session of the experiment. The logistic regression analysis indicated no apparent nonrandom sampling: Model $\chi^2(6, N = 161) = 9.02, p = .173$. None of the variables affected whether 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 variable—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 between the study variables and total sample and cell means and standard deviations are shown in Table 2. For efficient reporting in this table, the exploration variables were summed across the practice period trials, and practice performance and learning were averaged across practice period and testing period trials, respectively. The moderate relationship between practice performance and learning ($r = .36, p < .01$) supports the view that these are different constructs and highlights the importance of using a transfer design to assess learning.

**Tests of Hypotheses**

Hypotheses 1A and 1B predicted that feedback specificity would positively affect practice performance and negatively affect learning, respectively. Hypothesis 1A was supported. Feedback specificity had a positive effect on practice performance, as expected. Hypothesis 1B was not supported. While we predicted a negative main effect on learning, feedback specificity had no direct impact on learning. A significant between-subjects main effect, $F(3, 157) = 27.42, p = .000, \eta^2 = .36$, and a linear within-subject Feedback Specificity $\times$ Trial interaction, $F(3, 157) = 17.67, p = .000, \eta^2 = .25$, support Hypothesis 1A. As depicted in Figure 3A, the effects are driven by differences in practice performance levels between the high and very high specificity conditions and the low and moderate feedback specificity conditions: between subjects, $\Psi = 210.89, SE = 24.33, t(157) = 8.67, p = .000, \eta^2 = .32, 95\% \text{ CI} = 142.13 < \text{estimate} \ \Psi < 279.66$; linear Feedback Specificity $\times$ Trial interaction, estimate $\Psi = 94.32, SE = 13.32, t(157) = 7.08, p = .000, \eta^2 = .24, 95\% \text{ CI} = 68.01 < \text{estimate} \ \Psi < 120.63$.

A nonsignificant between-subjects main effect for feedback specificity over the six testing trials, $F(3, 145) = 2.48, p = .061, \eta^2 = .05$, and nonsignificant linear Feedback Specificity $\times$ Trial interaction, $F(3, 145) = 1.56, p = .202, \eta^2 = .03$, shows that feedback specificity did not have the negative direct impact on learning that we predicted with Hypothesis 3B. The latter result, depicted in Figure 3B, also shows that the no-feedback specificity group was advantaged in terms of the rate of further performance improvement during the testing period. However, comparison of the late Time 1 trials in Figure 3A and the Time 2 trials in Figure 3B shows that those in the low and moderate conditions learned far more than their Time 1 practice performance levels would suggest. Participants in these conditions performed far better in the Time 2 testing trials than they did during the last six trials of the Time 1 practice period, whereas those in the high and very high conditions performed similarly from late Time 1 to Time 2: within-subject
Table 2

**Correlations Among Variables, Means, and Standard Deviations**

<table>
<thead>
<tr>
<th>Variable</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Learning</td>
<td>0.00</td>
<td>0.36***</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2. Practice performance</td>
<td></td>
<td></td>
<td>0.42***</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3. Early practice performance</td>
<td></td>
<td></td>
<td></td>
<td>0.54***</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4. Systematic exploration</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.10</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5. Unsystematic exploration</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.32***</td>
<td></td>
<td></td>
</tr>
<tr>
<td>6. Hold steady</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.23**</td>
<td></td>
</tr>
<tr>
<td>7. Feedback specificity</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.19*</td>
<td></td>
</tr>
<tr>
<td>8. Perceived feedback specificity</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.00</td>
<td></td>
</tr>
</tbody>
</table>

**Note.** Early practice performance was included as a control variable in tests of Hypotheses 3 and 5, and perceived feedback specificity was collected as a potential control variable. \( n = 161 \) among Time 1 variables; \( n = 149 \) between learning (Time 2) and Time 1 variables.

* Coded: low = 1, moderate = 2, high = 3, very high = 4.

* \( p < .05 \). ** \( p < .01 \). *** \( p < .001 \).

Feedback specificity by time period, \( F(3, 145) = 25.87, p = .000, \eta^2 = .35 \); low and moderate versus high and very high, estimate \( \Psi = 62.02, SE = 7.17, t(145) = 8.65, p = .000, \eta^2 = .34, 95\% CI = 47.84 < \text{estimate } \Psi < 76.20 \).

Hypothesis 2 predicted that feedback specificity would have a negative impact on systematic and unsystematic exploration during the practice period. This hypothesis was supported. Between-subjects main effects demonstrated overall differences in systematic, \( F(3, 157) = 2.97, p = .034, \eta^2 = .05 \), and unsystematic, \( F(3, 157) = 16.96, n_s, \eta^2 = .24 \), exploration across the feedback specificity conditions. A similar difference was found for the hold steady strategy, \( F(3, 157) = 22.13, p = .000, \eta^2 = .30 \). Subsequent analyses revealed that those in the low and moderate feedback specificity conditions engaged in more systematic [estimate \( \Psi = -1.35, SE = 0.49, t(157) = -2.73, p = .007, \eta^2 = .05, 95\% CI = -2.33 < \text{estimate } \Psi < -0.38 \)] and unsystematic [estimate \( \Psi = -4.98, SE = 0.75, t(157) = -6.66, p = .000, \eta^2 = .22, 95\% CI = -6.45 < \text{estimate } \Psi < -3.50 \)] exploration and used the hold steady strategy [estimate \( \Psi = 6.33, SE = 0.83, t(157) = 7.65, p = .000, \eta^2 = .27, 95\% CI = 4.69 < \text{estimate } \Psi < 7.96 \)] less often than those in the high and very high conditions. Also, those in the high condition explored more [more unsystematic, estimate \( \Psi = -1.92, SE = 0.52, t(157) = -2.28, p = .024, \eta^2 = .03, 95\% CI = -2.23 < \text{estimate } \Psi < -0.16 \); less hold steady, \( \Psi = 1.53, SE = 0.58, t(157) = 2.64, p = .009, \eta^2 = .04, 95\% CI = 0.38 < \text{estimate } \Psi < 2.67 \)] than those in the very high condition.

The three sets of columns in Figure 4 show how often, out of a total of 48 opportunities across practice period trials, participants engaged systematic exploration, unsystematic exploration, or a hold steady strategy in each feedback specificity condition. Within-subject linear Feedback Specificity × Trial interactions also yielded the expected differences in systematic, \( F(3, 157) = 11.85, p = .000, \eta^2 = .18 \), and unsystematic, \( F(3, 157) = 4.75, p = .003, \eta^2 = .08 \), exploration, as well as in the use of the hold steady strategy, \( F(3, 157) = 17.75, p = .000, \eta^2 = .08 \). Subsequent analyses indicate that systematic [estimate \( \Psi = -2.14, SE = 0.39, t(157) = -5.48, p = .000, \eta^2 = .16, 95\% CI = -2.91 < \text{estimate } \Psi < -1.37 \)] and unsystematic [estimate \( \Psi = -1.17, SE = 0.42, t(157) = -2.76, p = .000, \eta^2 = .05, 95\% CI = -2.02 < \text{estimate } \Psi < -0.33 \)] exploration decreased more over time and use of the hold steady strategy [estimate \( \Psi = 3.11, SE = 0.46, t(157) = 7.25, p = .000, \eta^2 = .25, 95\% CI = 2.41 < \text{estimate } \Psi < 4.21 \)] increased more over time for those in the high and very high feedback specificity conditions than for those in the low and moderate conditions. Systematic exploration decreased more over time for those in the very high feedback specificity condition compared with those in the high condition: estimate \( \Psi = -0.56, SE = 0.27, t(157) = -2.06, p = .041, \eta^2 = .03, 95\% CI = -1.10 < \text{estimate } \Psi < -0.02 \). As expected, the form of the interactions suggests that the more informative the content of the feedback is about performance processes and outcomes, the less participants explore the task space, overall and over time.

Hypothesis 3A, which predicted a positive relationship between systematic exploration and learning, was supported. As shown in Table 3, when controlling for early practice performance and feedback specificity, systematic exploration had a significant relationship with learning, \( F(1, 143) = 3.97, p = .048, \eta^2 = .03 \). Hypothesis 3B, which predicted a negative relationship between unsystematic exploration and learning, was also supported. As shown in Table 3B, when controlling for early practice perfor-
Performance and feedback specificity, unsystematic exploration was significantly related to learning, $F(1, 143) = 3.95, p = .049, \eta^2 = .03$. The hold steady strategy was positively related to learning ($r = .23, p < .01$) but, after controlling for early practice performance and feedback specificity, this effect was no longer significant, $F(1, 143) = 0.22, p = .639, \eta^2 = .00$.

Hypothesis 4 predicted that systematic exploration would partially mediate the effect of feedback specificity on learning. Because we found no main effect of feedback specificity on learning (Hypothesis 1B), Hypothesis 4 was not supported.

Table 3 also includes the results of tests of Hypothesis 5, which predicted that exploration and feedback specificity would interact to affect learning. Hypothesis 5A was not supported. The interaction of systematic exploration and feedback specificity did not have a significant impact on learning, $F(3, 140) = 0.51, p = .676, \eta^2 = .01$.

Hypothesis 5B was partially supported. Controlling for early practice performance, unsystematic exploration interacted with feedback specificity to affect learning, $F(3, 140) = 3.17, p = .027, \eta^2 = .07$, as shown in Table 3. We expected that unsystematic exploration would be negatively related to learning when feedback specificity was lower and positively related to learning when feedback specificity was higher. The shape of interaction was as expected for the low, moderate, and high feedback specificity conditions, but not for the very high condition. Figure 5A shows that unsystematic exploration was negatively related to learning in the low and moderate feedback specificity conditions and had a positive relationship with learning in the high feedback specificity condition, as anticipated. However, unexpectedly, it was negatively related to learning in the very high feedback specificity condition.

Use of the hold steady strategy also interacted with feedback specificity to affect learning as shown in Table 3, $F(3, 140) = 2.94, p = .046, \eta^2 = .06$. The form of the interaction, as shown in Figure 5B, is essentially the opposite of the form of the interaction for unsystematic exploration, as would be expected. Holding steady proved to be a beneficial strategy for learning for those in the low, moderate, and very high feedback specificity conditions and a harmful strategy in the high feedback specificity condition.

Discussion

The results of this experiment demonstrate that increasing the specificity of the feedback during practice differentially affects initial performance and later performance that requires learning and reduces the level of exploration during practice. In addition, the various exploration strategies are differentially related to learning and interact with the specificity of the feedback to affect learning. By examining the impacts of differing levels of feedback specificity, up to a level that includes corrective responses, we have provided the first full test of the generalization that specific feedback leads to higher performance and learning and have identified some potential qualifying conditions for textbook prescriptions to be specific when giving feedback. The theoretical and practical implications of these results, plus limitations and recommendations for follow-up research, are discussed later in this section.

As predicted, feedback specificity differentially affected practice performance and learning—although, for the latter, not in the

![Figure 3](image-url)

Figure 3. Panel A: Effects of feedback specificity on practice performance (Time 1 testing period). Panel B: Effects of feedback specificity on learning (Time 2 testing period). Low is indicated by a diamond; moderate is indicated by a square; high is indicated by a triangle; very high is indicated by a circle.

![Figure 4](image-url)

Figure 4. Effects of feedback specificity on Time 1 exploration. There were 48 opportunities to use systematic exploration, unsystematic exploration, or a hold steady strategy.
exact fashion we predicted. It positively affected practice performance, as expected, but did not have the anticipated negative main effect on learning. Instead, feedback specificity had no main effect on learning. However, consistent with our arguments, the most significant improvements in performance levels from the initial practice period to later performances of the task in the testing period were by those in the low and moderate feedback specificity conditions. While the results of many studies using psychomotor tasks show negative effects of feedback on learning (see Schmidt, 1991, and Salmoni et al., 1984, for reviews), some studies provide evidence of no lasting effect (e.g., M. Goldstein & Rittenhouse, 1954; Stockbridge & Chambers, 1958), as we did in this experiment for differing levels of feedback specificity on a managerial decision-making task. There are two possible explanations for the lack of effect of feedback specificity on learning in our study. First, the initial effect on performance during the practice period could be entirely transient, with no lasting positive or negative impact. Alternatively, those in different feedback specificity conditions could have learned through different but equally effective means, possibly learning how to manage the task under different conditions. The learning effects for the interactions between feedback specificity and type of exploration are consistent with the latter explanation, as discussed later.

Also as predicted, feedback specificity was negatively related to the level of exploratory behavior. Increasing feedback specificity led to less systematic and unsystematic exploration behaviors. As expected, systematic exploration was positively related to learning, and unsystematic exploration was negatively related to learning. While we were unable to test for mediation, these results suggest an indirect relationship. It is possible that a feedback-induced decrease in systematic exploration could be detrimental to learning, and a feedback-induced decrease in unsystematic exploration could be beneficial to learning. Of importance, these results point to the limitations of measures of strategy that do not adequately differentiate the quality of the search strategy used and the implications for performance and learning on the task being studied. This is one of many limitations of self-report measures of strategy, which may confuse search effort with strategy quality on complex tasks. Our use of objective measures of exploration strategies is one of the strengths of our study.

We also predicted that feedback specificity would interact differently with systematic and unsystematic exploration strategies in their impact on learning because of the differing combinations of information that become available with each type of exploration. While there was no significant interaction with systematic exploration, we found a significant interaction between feedback spec-

Table 3
Analysis of Variance for Tests of Hypotheses 3 and 5: The Effects of Exploration on Learning and the Interactive Effects of Feedback Specificity and Exploration on Learning

<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><strong>Systematic Exploration</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Step 1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Early practice performance</td>
<td>2,316.34</td>
<td>2,316.34</td>
<td>14.38***</td>
<td>1, 143</td>
<td>.09</td>
</tr>
<tr>
<td>Feedback specificity</td>
<td>789.64</td>
<td>263.21</td>
<td>1.63</td>
<td>3, 143</td>
<td>.03</td>
</tr>
<tr>
<td>Systematic exploration</td>
<td>639.72</td>
<td>639.72</td>
<td>3.97*</td>
<td>1, 143</td>
<td>.03</td>
</tr>
<tr>
<td>Error</td>
<td>23,039.61</td>
<td>161.12</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Step 2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Feedback specificity by systematic exploration</td>
<td>249.37</td>
<td>83.12</td>
<td>0.51</td>
<td>3, 140</td>
<td>.01</td>
</tr>
<tr>
<td>Error</td>
<td>22,790.24</td>
<td>568.77</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Unsystematic Exploration</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Step 1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Early practice performance</td>
<td>1,291.48</td>
<td>1,291.48</td>
<td>8.01**</td>
<td>1, 143</td>
<td>.05</td>
</tr>
<tr>
<td>Feedback specificity</td>
<td>180.51</td>
<td>60.17</td>
<td>0.37</td>
<td>3, 143</td>
<td>.01</td>
</tr>
<tr>
<td>Unsystematic exploration</td>
<td>635.93</td>
<td>635.93</td>
<td>3.95*</td>
<td>1, 143</td>
<td>.03</td>
</tr>
<tr>
<td>Error</td>
<td>23,043.40</td>
<td>161.14</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Step 2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Feedback specificity by unsystematic exploration</td>
<td>1,463.72</td>
<td>487.91</td>
<td>3.17*</td>
<td>3, 140</td>
<td>.07</td>
</tr>
<tr>
<td>Error</td>
<td>21,579.69</td>
<td>154.14</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Hold Steady</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Step 1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Early practice performance</td>
<td>1,916.20</td>
<td>1,916.20</td>
<td>11.59***</td>
<td>1, 143</td>
<td>.07</td>
</tr>
<tr>
<td>Feedback specificity</td>
<td>316.98</td>
<td>105.66</td>
<td>0.64</td>
<td>3, 143</td>
<td>.01</td>
</tr>
<tr>
<td>Hold steady</td>
<td>36.64</td>
<td>36.64</td>
<td>0.22</td>
<td>1, 143</td>
<td>.00</td>
</tr>
<tr>
<td>Error</td>
<td>23,642.69</td>
<td>165.33</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Step 2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Feedback specificity by hold steady</td>
<td>1,401.10</td>
<td>467.03</td>
<td>2.94*</td>
<td>3, 140</td>
<td>.06</td>
</tr>
<tr>
<td>Error</td>
<td>21,579.69</td>
<td>720.09</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Note. The dependent variable is learning averaged over the six Time 2 testing period trials. Early practice performance was included as a control variable. The exploration variables were centered. n = 149.  
* p < .05. ** p < .01. *** p < .001.
The interactions, coupled with the nonsignificant main effect of feedback specificity on learning, suggest that those who get feedback of varying specificity may learn through different but equally beneficial mechanisms. For those in the high feedback specificity condition, the more decision changes made the better. The outcome feedback and feedback about correct and incorrect decisions helped participants interpret the confounded effects of multiple decision changes on performance levels. The confounded information proved to be less interpretable to those in the low and moderate feedback specificity conditions.

Those in the low and moderate feedback specificity conditions also benefited from limiting their exploration. Holding their decisions steady across some trials was helpful to learning, possibly because on such a complex task, performance levels will fluctuate somewhat even when the same decisions are made on subsequent trials. Holding steady provided additional information on the effects of particular decisions and decision combinations on employee performance levels. Those in the high feedback specificity condition did not use this information to their advantage. For them, it appears that holding steady resulted in wasted trials.

Taken together, the effects of feedback specificity on exploration in the low, moderate, and high conditions and the relationships between exploration and learning in those conditions may seem contradictory on the surface but are consistent with our arguments and point to the complexity of feedback effects. Feedback specificity influenced the levels of exploration while the effects of the different types of exploration on learning differed across the different levels of feedback specificity. The lower guidance in the low and moderate feedback specificity conditions encouraged greater exploration, but it was unsystematic exploration, which did not contribute to learning in those feedback specificity conditions. Alternatively, the greater guidance in high feedback specificity led to less exploration, but more exploration of any type, even if it was unsystematic, produced better learning outcomes when feedback specificity was high.

**Theoretical and Practical Implications**

Establishing the link between feedback specificity and exploration behaviors during the learning phases of task performance also points to a new direction for theorizing and research on the effects of feedback on performance. Previous explanations for the effects of feedback on performance have focused on micromotivational and cognitive processes, with particular emphasis on how feedback affects the focus of attention (e.g., Kluger & DeNisi, 1996; Locke & Latham, 1990). These hypotheses are rarely tested directly in organizational behavior research, which is typically concerned with more molar and more observable behaviors that can be more directly evaluated and managed in work settings. Exploration, search, or information acquisition behaviors are more observable than attention processes and are an important activity in many problem-solving and learning processes. Also, like feedback, they are a potential source of task-relevant information. Future research can extend researchers’ understanding of the relationship between feedback specificity and exploration behaviors and how they affect learning. It should also consider the effects of other aspects of feedback, as outlined in the later discussion of possible future research.

Another new avenue for theorizing comes from results that are consistent with the argument that participants in different feedback specificity conditions may have learned correct responses for different aspects of the task. This hypothesis cannot be directly tested with our data set because our measure of learning is not sensitive enough to determine exactly what correct responses were learned and by whom. However, it could explain why such vast differences in practice performance across feedback specificity conditions would lead to no apparent differences in the learning measure, as performance levels for those in the low and moderate conditions increased while performance levels decreased for those in the high and very high conditions, comparing late practice period to early testing period trials. Recall that the decision rules for the simulation are dynamic, such that whether a particular choice of action is correct depends on how well employees are performing and for
how long. For example, employees should be given a difficult goal until or unless they perform very poorly for two consecutive trials, at which point, a moderate goal should be given. Then, when performance rises above standard, a difficult goal should be given. Given the practice performance differences during the Time 1 practice period, those in the low and moderate feedback specificity conditions would have had more opportunities to learn what to do under conditions of poor employee performance (e.g., give moderate goals, process and outcome feedback, and low rewards). They spent more time responding to poor performance. Alternatively, those in the high and very high feedback specificity conditions would have had more opportunities to learn what to do under conditions of good employee performance (e.g., give difficult goals, outcome feedback, and high rewards). Participants spent more time responding to good performance. The positive effects of learning how to respond to good versus poor levels of performance likely canceled each other out in the overall measure of learning used in this experiment. These results and observations point to an opportunity for theory development regarding the differential learning opportunities created by different levels of feedback specificity, which also has practical implications for designing optimal feedback interventions to provide opportunities to learn all aspects of effectively executing a task.

From a more practical perspective, our aim was to test the accepted generalization that increasing the specificity of feedback will have a positive impact on performance and learning and not, necessarily, to generalize the results of this specific study to specific natural settings (Ilgen, 1986; Locke, 1986; Mook, 1983). By focusing on the information content of feedback messages, we have isolated the key element of feedback specificity and have shown that it can have effects inconsistent with the generalization that underpins unqualified textbook prescriptions to “give specific feedback.” Our results do not support a conclusion that supervisors and other sources should never provide specific feedback, but they do lead us to question that being more specific when giving feedback is always a good thing. As predicted, increasing the specificity of feedback had a positive impact on immediate performance but had no effect on longer term transfer performance that required learning. Also, providing more specific feedback reduced the levels of exploration, which, if done correctly, will contribute to problem solving and learning on most tasks. Supervisors who continually provide highly specific feedback, including advice on how to “get it right,” to new staff and those experiencing performance problems might therefore undermine the very behaviors that are believed to be integral to learning (Condry, 1977). However, further research is needed before we can make a general statement on what levels of specific information should be provided in feedback messages and when.

**Limitations and Future Research**

The use of an experimental simulation to test our hypotheses raises the obvious question of whether these results will generalize to other more natural settings and other sources, particularly human sources, in which 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 versus subjective feedback has been shown to generalize to natural work settings, where the effects are often amplified when feedback interacts with other variables that are controlled in laboratory settings (Kopelman, 1986). Our manipulations represented the range of specificity levels for objective feedback by varying the levels 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). Future research to establish the generalizability of the effects of feedback specificity should focus on the interactions between feedback specificity and features of the feedback source, other dimensions of the feedback, and properties of the situation.

We used a computer as our feedback source. Computer-provided feedback is becoming more common as computerized training and computer-based tasks become more prevalent in organizations. However, we cannot merely assume that impersonal computerized feedback and feedback from human sources operate similarly. Issues fairly unique to human sources, such as motive interpretation (Ilgen et al., 1979), source credibility (Fedor, 1991), and inconsistencies in feedback provision, may lead to different learning effects for human versus computer feedback interventions. Future research should compare the interactive effects of human versus computer and other impersonal feedback interventions of differing levels of specificity on exploration, learning processes, and outcomes.

Future research could also consider the consequences for exploration and learning of other dimensions of feedback, such as the sign, frequency, and timing. Evidence suggests that these dimensions also differentially affect practice performance and learning (cf. Goodman, 1998; Schmidt, 1991). For example, we expect that the sign of feedback will influence the levels of exploration and that this relationship may be moderated by the specificity of the feedback intervention. Negative feedback, because it indicates that actual performance is below the desired standard, will typically stimulate some search to identify a corrective response to close the gap. Over the longer term, the exploratory activity stimulated by negative feedback may have significant developmental benefits, and these effects may be moderated by individual differences in resilience to negative feedback and uncertainty (Dean, Russell, & Muchinsky, 1999). Our results suggest that the potential developmental benefits of exploratory search following negative feedback may also be reduced when it is highly specific feedback that includes information on the causes of the poor performance and the appropriate corrective responses. Such feedback is often received from people who are trying to be helpful and who may confuse it with support. Positive feedback is more likely to be interpreted as evidence that current strategies are okay and is less likely to stimulate search for new strategies, even under conditions of low feedback specificity.

Earlier, we discussed the implications for theory development of the possibility that participants in different feedback specificity conditions may have learned correct responses for different aspects of the task. Future research should test hypotheses about the effects of feedback specificity on exposure to different situations, such as good versus poor performance, during practice and tease apart learning to assess exactly what is learned in each feedback specificity condition as a result of differential learning opportunities. Such research may lead to knowledge about optimal feedback interventions that provide opportunities to learn all aspects of effectively performing a task.
Finally, the observed relationship between the practice performance and learning ($r = .36$, $p < .01$) and differential effects of feedback specificity on practice performance (positive) and learning (none) highlights the importance of using a transfer design to examine the impact of feedback interventions on longer term performance and performance under conditions that differ from the practice period. A study design that only examined practice performance would have erroneously concluded that low and moderate feedback specificity conditions were detrimental to learning. Despite the poor performance during practice for participants in these feedback specificity conditions, they learned just as much as participants in the high and very high conditions and experienced greater improvements in their performance from the practice period to the testing period. Future research should also use a transfer design to examine the impacts of feedback, as well as other conditions of practice, on learning.

**Conclusion**

The results of our transfer experiment show that increasing the specificity of the feedback intervention positively affected practice performance, but these benefits did not endure over time, modification of the task, and removal of the intervention. In addition, we found that increasing feedback specificity led to less exploration of all kinds, greater use of systematic exploration during practice was positively related to learning, greater use of unsystematic exploration was detrimental to learning overall, and use of unsystematic exploration and hold steady strategies interacted with feedback specificity to affect learning. These results point to the need for greater specificity and possible qualifications to the recommendation to “be specific when giving feedback.” Our results suggest that those who got feedback of varying specificity may have learned through different but equally beneficial mechanisms, possibly learning how to manage the task under different conditions. Future research should continue to explore the mechanisms by which variance in different dimensions of feedback interventions, in interaction with other features of the task setting, affect learning.

**References**


Revised: June 18, 2002

Would you like to know when the next issue of your favorite APA journal will be available online? This service is now available to you. Sign up at http://watson.apa.org/notify/ and you will be notified by e-mail when issues of interest to you become available!