Performance Monitoring: How It Affects Work Productivity

James R. Larson, Jr., and Christine Callahan
The University of Illinois at Chicago

Performance monitoring refers to gathering information about the work effectiveness of others. Traditionally, monitoring has been assumed to influence people's work behavior only to the extent that it is accompanied by additional managerial action (e.g., delivering positive or negative performance consequences). It seems likely, however, that performance monitoring may also have an independent effect on work behavior by influencing the perceived importance of the monitored task. This hypothesis was tested in a laboratory experiment in which subjects worked at 2 tasks for 2 hr. Subjects' work output was significantly affected both by monitoring alone and by monitoring coupled with contingent performance consequences, and the effect of monitoring alone was mediated by perceptions of task importance. The latter result is discussed in terms of social-information-processing theory, especially as it relates to role-making processes in organizations.

The productivity of individuals and groups at work is generally assumed to be due in part to the quality of the supervision and leadership they receive. Guided by this assumption, researchers have long sought to identify the specific supervisory and leadership behaviors that contribute to managerial effectiveness (e.g., Bowers & Seashore, 1966; Dowell & Wexley, 1978; Fleishman, 1953; House & Mitchell, 1974; Lord, 1977; Luthans & Lockwood, 1984; Stogdill, 1963; Vroom & Yetton, 1973; Yukl & Nemeroff, 1979).

One supervisory behavior that has recently begun to attract some attention is performance monitoring. Performance monitoring refers to the gathering of information about the work effectiveness and productivity of individuals, groups, and larger organizational units. This might be done by observing employees' work behavior, inspecting their work output, asking them to report verbally about their work progress (e.g., during weekly staff meetings), or by reading documents that summarize key performance indicators. These tactics vary widely in their obtrusiveness. In some cases, a manager's monitoring behavior will be very apparent to those whose performance is being monitored, whereas in other cases it will not.

The impact of performance monitoring on individual and organizational effectiveness has generally been assumed to be a conjunctive one (cf. Mintzberg, 1973; Yukl, 1989). That is, performance monitoring is typically thought to influence work productivity only to the extent that it is coupled with subsequent managerial action (e.g., providing feedback, rewarding good performance, eliminating barriers to work effectiveness).

Illustrative of this viewpoint is a recent study by Komaki (1986). She had trained observers systematically record and code the behavior of 24 managers working in two divisions of a large insurance firm. Twelve of these managers were judged by their superiors to be very effective in motivating others; the remaining 12 were judged to be relatively ineffective in this regard. Komaki (1986) found that the only behavior category to distinguish between these two groups of managers was that of performance monitoring. The effective managers monitored their subordinates' performance almost 50% more often than did those who were judged to be ineffective, primarily by making more frequent use of direct work sampling (as opposed to relying on either self-reports about work progress or reports from other individuals). Interestingly, there were no differences between the two groups in the frequency with which they provided either performance antecedents (defined as behaviors that instruct, remind, or convey expectations about performance) or performance consequences (defined as behaviors that convey a knowledge of performance, such as giving performance feedback and providing verbal rewards or punishments).

In discussing her results, Komaki (1986) noted that performance monitoring should enhance managerial effectiveness because it enables managers to evaluate subordinate performance more accurately, thus increasing the managers' ability to provide consequences that are contingent upon performance. It has been well documented that performance improves when consistent, contingent consequences are provided (e.g., Honig & Stadden, 1977; Kazdin, 1980; Komaki, Desselles, & Schepman, 1988; Luthans, 1979; Scott & Podsakoff, 1985). Thus, despite the fact that there was no significant difference in the simple frequency with which performance consequences were provided by the two groups of managers in her study, Komaki (1986) postulated that the consequences delivered by the effective managers were nevertheless more successful in influencing their subordinates' performance because the delivery of those consequences was based on more accurate, specific performance information. Said differently, performance monitoring was assumed to contribute to managerial effectiveness because it increased the contingency between subordinate work performance and the performance consequences managers delivered.

An Alternative Perspective

Although this view of performance monitoring is compelling, both personal experience and prior theory suggest that
performance monitoring may also affect work productivity in ways that do not require it to be coupled with subsequent managerial action. Specifically, social-information-processing theory postulates that an individual's attitudes and beliefs can be significantly influenced by subtle social (i.e., behavioral) cues that affect the way in which events at work are perceived (Pfeffer, 1980; Salancik & Pfeffer, 1977, 1978). Virtually any behavior can serve this cueing function, even when it is not explicitly intended to do so. For example, the frequency with which a manager monitors a subordinate's performance may help shape that subordinate's beliefs about the relative importance of his or her various work activities. It seems reasonable to expect that the more important a particular task is to a manager, the more closely performance on that task will be monitored (other things being equal). If so, and if the subordinate is aware of the manager's performance-monitoring behavior, then that behavior should serve as a cue indicating the relative importance of the task in question. It may also signal whether or not the subordinate can expect to be rewarded for performing well on that task or punished for performing poorly. Such outcome expectancies are an essential element of many cognitive models of work motivation (e.g., Naylor, Pritchard, & Ilgen, 1980; Porter & Lawler, 1968; Vroom, 1964) and are assumed to guide behavior before (sometimes long before) the outcomes actually materialize.

Thus, the implication of social-information-processing theory is that performance monitoring may also have a relatively independent effect on subordinate productivity through its influence on perceived task importance. This effect is expected to be separate from, and in addition to, the previously described conjunctive effect, which occurs when monitoring is coupled with subsequent managerial action.

The Present Study

The purpose of the present study was to test whether performance monitoring can indeed affect work behavior even in the absence of subsequent managerial action. In particular, the following two hypotheses were tested:

Drawing on social-information-processing theory, we predicted that monitoring, by itself, would increase performance on a task in comparison with when performance on that task was not monitored. This was expected to occur even when monitoring was not followed by the delivery of contingent performance consequences. The only requirement was that the monitoring activity be apparent to the individual or group whose performance was being monitored.

On the basis of conventional assumptions about the conjunctive effects of performance monitoring, we also predicted that when monitoring was followed by the delivery of contingent performance consequences, performance on the monitored task would increase significantly in comparison with when monitoring was not followed by the delivery of contingent performance consequences.

These hypotheses were tested in a laboratory experiment in which subjects performed two different work tasks during a 2-hr period. One third of the subjects worked on these tasks for the entire 2 hr without interruption, and another third was interrupted every 20 min as the experimenter (Christine Calla-
circle each word they thought was misspelled. There was no time limit for this test because it was intended as an index of power, not speed. The percentage of words correctly classified was taken as the pretest measure of spelling ability.

Finally, subjects were given an alphabetizing test. The materials for this test consisted of two word lists, each printed on a separate piece of paper and each containing 15 proper names. The subjects were to sort the names in each list into one of two adjacent columns on the page according to the first letter of the name. Letters of the alphabet were randomly assigned to columns, and this assignment differed for each word list. Subjects then had to sort the two resulting columns of names into alphabetical order. The number of names correctly sorted and alphabetized was used as the measure of alphabetizing ability.

Cover story. After these three measures were obtained, the cover story for the experiment was presented. Subjects were told that they were participating in a study investigating the effects of noise on task performance, that they were in the "no-noise" condition, and that their performance would serve as a standard against which the performance of individuals in two "noise" conditions would be compared. In fact, noise was held constant at a low level for all subjects.

Experimental tasks. The two tasks the subjects were to perform were then explained in detail. Each subject was handed a copy of a typed manuscript of a chapter from an introductory organizational behavior textbook. Part of the subject's job was to proofread this chapter for spelling errors (e.g., ther, manager, cabinet). Specifically, subjects were to read through the chapter carefully and circle any spelling errors they encountered. They did not actually have to correct the errors, however. When describing this task, the experimenter stressed the importance of both proofreading as much of the chapter as possible (i.e., quantity of work) and catching all of the errors in whatever work they did proofread (i.e., quality of work).

In addition to the proofreading task, subjects were given an alphabetizing task similar to the one they had done during the pretest phase. Thus, they were handed a stack of pages with a different list of 15 words on each. They were to sort each list into two columns according to the first letter of the word and then to arrange the columns in proper alphabetical order. Again, the experimenter stressed the importance of both alphabetizing as many word lists as possible (i.e., quantity of work) and being accurate in doing so (i.e., quality of work).

Treatment manipulations. At this point, the instructions varied by treatment condition. Subjects in the monitoring-only condition were told

"While you are working on these two tasks, I will come into the room every 20 min or so to collect your work and to answer any questions you may have."

In addition to the previous statement, those in the monitoring-plus-consequences condition were told

"I will also give you information about your performance up to that point in the experiment."

Subjects in the control condition received neither of the above instructions.

All subjects were then told

"Your goal is to do the best you can on both tasks. While working, you can switch back and forth between the two tasks as much as you wish. For example, if you get tired of working on one task, switch to the other one for a while. When you get tired of that, switch back, and so on. Just remember that your goal is to do your best on both tasks."

The subjects were questioned to make sure that they understood the procedure and then were moved to separate rooms, where they worked independently on these tasks for 2 hr. Subjects in the control condition worked uninterrupted for the entire 2-hr period. In the monitoring-only and monitoring-plus-consequences conditions, the experimenter appeared once every 20 min to collect work that had been completed, and, if appropriate, to give feedback.

Although the subjects worked at two different tasks, the experimenter monitored (and gave feedback about) their performance on only one of the tasks. Proofreading was designated as the task to be monitored for half of the subjects, whereas alphabetizing was so designated for the remaining subjects. Each time the experimenter collected the subjects' work in the monitoring-only and monitoring-plus-consequences conditions, she did so only for the designated task. Thus, output from only the proofreading task was collected for half of the subjects, whereas output from only the alphabetizing task was collected for the other half of the subjects. Furthermore, each time the experimenter collected work from a subject, she made a statement indicating that she was monitoring the quantity of work that had been completed. For example, at the first 20-min monitoring point the experimenter said, "Let's see how many lines you've proofread" when performance on the proofreading task was being monitored. No work from the task designated as the nonmonitored task was ever collected.

In the monitoring-plus-consequences condition, the experimenter not only monitored the subjects' performance on the designated task but she also provided the subjects with specific, accurate performance feedback. This feedback was focused on the quantity of work completed and was accompanied by an appropriately worded evaluative statement. For example, if performance on the proofreading task was being monitored, the experimenter gave feedback by stating

"During the last 20 min period you completed a total of ______ lines. This is better (worse) than your previous average, which was about ______ in 20 min. That's very good (That means you probably didn't do as well as you could have)."

Finally, in the control condition, the experimenter never interrupted the subjects either to monitor their performance or to give feedback. Consequently, there was no difference in the way output from the two experimental tasks was treated. The monitored and nonmonitored designations were used in the control condition solely for the purpose of subsequent between-group statistical comparisons.

At the end of the 2-hr period, all subjects completed a postexperimental questionnaire, after which they were debriefed and dismissed.

Dependent Measures

The primary dependent variables in the study were the quantity and quality of the subjects' performance on each task. In the proofreading task, performance quantity was defined as the total number of lines read. Performance quality was indexed by the number of misspelled words that were circled minus the number of correctly spelled words that were circled, divided by the total number of lines read. In the

1 Note that the alternative wording here reflects the fact that the experimenter's evaluative statements were based on the subject's actual performance on the monitored task. The wording does not constitute another independent variable manipulation. Over the course of the 2-hr experiment, subjects in the monitoring-plus-consequences condition often received both positively worded and negatively worded statements, though they generally received more of the former than of the latter. Also, if a subject in this condition did no work on the monitored task during a particular 20-min period (as occasionally happened, especially early in the experimental session), feedback consisted merely of a reminder at the next feedback delivery point that no work at all had been completed on the monitored task for the period in question.
alphabetizing task, the total number of word lists alphabetized was taken as the quantity measure, and the overall percentage of words correctly sorted and placed in proper alphabetical order served as the quality measure. Finally, because the raw scores for these two tasks were not directly comparable, they were converted to standard scores (using the grand mean and standard deviation for each measure across all subjects) before being analyzed.

In addition to the behavioral measures of performance quantity and quality, manipulation-check items were included in the postexperimental questionnaire to determine whether the subjects perceived the experimenter's monitoring and feedback behavior as anticipated. Several questions were also included to assess the subjects' perceptions of the relative importance of the two tasks they performed.

Results

Manipulation Checks

In the postexperimental questionnaire, we asked subjects to indicate how often the experimenter checked to see how well they were doing (1 = never and 7 = frequently). A planned comparison revealed that, as expected, subjects in the monitoring-only (M = 4.67) and monitoring-plus-consequences (M = 5.07) conditions reported that the experimenter monitored their performance significantly more often than subjects in the control condition reported (M = 1.36), t(80) = 9.76, p < .001. There were no other significant between-group differences for this measure. The monitoring portion of the supervisory-behavior-pattern manipulation thus appears to have been successful.

To assess the effectiveness of the consequences portion of the supervisory-behavior-pattern manipulation, we asked subjects two questions about the amount of performance feedback they got from the experimenter (1 = none and 7 = a great deal). Responses to these two questions were averaged to form a single feedback quantity index. As expected, a planned comparison using this index revealed that subjects in the monitoring-plus-consequences condition (M = 4.77) reported getting significantly more feedback from the experimenter than did subjects in either the monitoring-only (M = 1.40) or control (M = 1.68) conditions, t(80) = 11.81, p < .001. There were no other significant between-group differences on this measure.

As a further check on the consequences portion of the supervisory-behavior-pattern manipulation, we asked subjects in the monitoring-plus-consequences condition to evaluate the focus of the feedback they received on a single bipolar scale (1 = focused on accuracy and 7 = focused on quantity). The mean rating for this item (M = 4.92) was significantly above the scale midpoint, t(25) = 2.42, p < .03, indicating that subjects perceived the feedback they got to be focused more on the amount of work they did than on its quality. Subjects in the monitoring-plus-consequences condition were also asked to evaluate the accuracy of the feedback they received, again on a single bipolar scale (1 = accurate and 7 = inaccurate). The mean rating for this item (M = 2.56) was significantly below the scale midpoint, t(26) = 4.76, p < .001, indicating that subjects generally perceived the feedback they received to be accurate. In combination, then, the results from all of these measures suggest that the consequences portion of the supervisory-behavior-pattern manipulation was also successful.

Finally, to assess subjects' overall affective reaction to the experimenter and her behavior, we asked subjects in all conditions to rate the experimenter on three 7-point bipolar scales (annoying—reassuring, nerve-racking—relaxing, and helpful—distracting). There were no between-group differences on any of these measures, and in each case the mean was in the positive direction and significantly above the scale midpoint (p < .001 for all). Thus, and as intended, the experimenter was perceived in a positive manner across all treatment conditions.

Pretest Measures

A preliminary set of analyses was performed to determine whether any of the three pretest measures were significantly related either to the independent variables or to the primary dependent variables. To assess whether the pretest measures were systematically related to the independent variables, we conducted three separate 3 X 2 (Supervisory Behavior Pattern X Task Monitored) analyses of variance (ANOVAs). Unexpectedly, one analysis revealed a significant supervisory-behavior-pattern main effect on the reading-speed measure, F(2, 84) = 4.39, p < .02. Even though we had used a random procedure to assign subjects to treatment conditions, subjects in the monitoring-only condition had significantly lower reading-speed scores than did subjects in the monitoring-plus-consequences condition (p < .05), with subjects in the control condition falling in between.

To determine whether any of the pretest measures were related to the primary dependent variables, we computed partial correlations that controlled for all treatment main and interaction effects. Once again, there was a significant relationship involving the reading-speed measures: Reading speed was significantly correlated with the proofreading quantity score, r(85) = .32, p < .003.

In light of these two findings, the proofreading quantity scores were adjusted to remove the influence of the subjects' pretested reading speed. This was done by regressing the proofreading quantity scores onto the reading-speed scores and then using the standardized residuals (i.e., that portion of each subject's proofreading quantity score that was independent of his or her pretested reading speed) in all subsequent analyses.

Tests of Main Hypotheses

The means and standard deviations for the performance quantity and quality measures are presented in Table 1. These data are reported separately for the monitored and nonmonitored tasks within each treatment condition. Because the quantity and quality scores were uncorrelated when all treatment main and interaction effects were held constant, r(85) = .05, ns, these two sets of scores were analyzed separately. A 3 X 2 X 2 mixed-design ANOVA was performed on each set of scores, followed where appropriate by planned comparisons to test the specific research hypotheses under investi-

2 Several subjects inadvertently failed to answer one or more questions on the postexperimental questionnaire. This accounts for the lower-than-expected degrees of freedom in certain manipulation-check and supplementary analyses.
In both analyses, supervisory behavior pattern (monitoring-only, monitoring-plus-consequences, or control) and type of task monitored (proofreading or alphabetizing) were between-subjects factors, and the source of the performance measure (the monitored or nonmonitored task) was a within-subjects factor.

For the performance quantity scores, the overall 3 x 2 x 2 ANOVA revealed both a significant main effect for source, F(1, 84) = 22.80, p < .001, η² = .21, as well as a significant Source x Supervisory Behavior Pattern interaction, F(2, 84) = 9.38, p < .001, η² = .18. The interaction is presented graphically in Figure 1. As can be seen, although subjects generally completed more work on the monitored task than on the nonmonitored task, the magnitude of this effect depended significantly on the supervisory behavior pattern exhibited by the experimenter. The effect was largest in the monitoring-plus-consequences condition, somewhat smaller but still present in the monitoring-only condition, and absent altogether in the control condition.

To test more precisely the two research hypotheses under investigation, we conducted two follow-up planned comparisons to assess the significance of the differences among the three supervisory-behavior-pattern means for the monitored task. Hypothesis 1 predicted that monitoring, by itself, would affect subjects' performance on the monitored task. In support of this hypothesis, we found that subjects in the monitoring-only condition (M = .82) completed significantly more work on the monitored task than did subjects in the monitoring-only condition, t(84) = 2.41, p < .01, η² = .06. These results thus clearly support the idea that performance monitoring by managers and supervisors can affect subordinate task performance both independently, and in conjunction with the delivery of contingent performance consequences.

The only other significant result from the overall analysis of the performance quantity scores was a main effect for the type of task monitored, F(1, 84) = 5.81, p < .002, η² = .06. This effect indicates that when proofreading was designated as the task to be monitored, subjects completed more work on both the proofreading and the alphabetizing tasks than when alphabetizing was designated as the task to be monitored. No other effects from the 3 x 2 x 2 analysis approached significance.

It was of interest to know whether these changes in the quantity of work completed by subjects were accompanied by any changes in the quality of their work. Accordingly, an overall 3 x 2 x 2 ANOVA was conducted on the performance quality scores. None of the main effect or interaction terms from this analysis approached significance, nor were any of the individual cell means significantly different from one another when tested by the Student-Newman-Keuls procedure, p > .05 for all comparisons.

Thus, it appears that the observed increases in the quantity of work completed on the monitored task in the monitoring-only and monitoring-plus-consequences conditions came at the expense of the quantity of work done on the nonmonitored task in those conditions and not at the expense of the quality of that work. Said differently, rather than working more quickly at both tasks and risking lower quality for each, subjects in the monitoring-only and monitoring-plus-consequences treatment conditions apparently redistributed the relative amount of ef-

Table 1
Dependent Measure Means and Standard Deviations in Each Treatment Condition

<table>
<thead>
<tr>
<th>Measure</th>
<th>Control (n = 15)</th>
<th>Monitoring only (n = 15)</th>
<th>Monitoring plus consequences (n = 16)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Proofreading</td>
<td>Alphabetizing</td>
<td></td>
</tr>
<tr>
<td>Quantity</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>M</td>
<td>−0.07</td>
<td>0.46</td>
<td>1.13</td>
</tr>
<tr>
<td>SD</td>
<td>0.63</td>
<td>1.09</td>
<td>0.88</td>
</tr>
<tr>
<td>Quality</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>M</td>
<td>−0.14</td>
<td>0.23</td>
<td>0.32</td>
</tr>
<tr>
<td>SD</td>
<td>0.99</td>
<td>0.40</td>
<td>0.62</td>
</tr>
</tbody>
</table>

Note. Means are in standard score form. Proofreading quantity scores have been adjusted to remove the effect of pretest reading speed. *Neither task was actually monitored in this condition. Rather, subjects were randomly divided into two subgroups, and the monitored and nonmonitored labels were assigned solely for purposes of statistical analysis.
As might be expected, the monitored task was generally perceived as more important than the nonmonitored task. How-

\[ F(2, 79) = 15.64, p < .001, \eta^2 = .25 \]

there was both a significant main effect for object, 
\[ F(1, 79) = 26.21, p < .001, \eta^2 = .25 \]

Supplementary Analyses

Because the independent (i.e., nonconjunctive) effect of monitoring was expected to affect subjects' work performance through its influence on their perceptions of task importance, several items were included in the postexperimental questionnaire to assess the extent to which subjects felt it was important to do well on both the monitored and nonmonitored tasks (1 = not at all important and 7 = very important). Responses to these items were averaged and then analyzed by means of an overall 3 x 2 x 2 mixed-design ANOVA in which the object of the importance ratings (the monitored or nonmonitored task) was included as a within-subjects factor.

The results of this analysis were very similar to those reported earlier for the performance quantity scores. Specifically, there was both a significant main effect for object, 
\[ F(1, 79) = 33.46, p < .001, \eta^2 = .29 \]

As expected, the task-importance ratings accounted for a significant amount of variance in the present analysis, 
\[ F(1, 78) = 33.46, p < .001, \eta^2 = .29 \]

To determine whether these task-importance perceptions were likely to have mediated the effects of the independent variables on the quantity of work completed by subjects, we performed a hierarchical regression analysis in which the task-importance ratings were entered in the first step, followed in the second step by the simultaneous entry of vectors for the three experimental variables (supervisory behavior pattern, type of task monitored, and the source of the performance measure) plus their interactions. This analysis is equivalent to the original 3 x 2 x 2 ANOVA performed on the quantity scores, with the exception that all of the variance in the dependent variable that could be accounted for by the importance ratings was removed first. To the extent that task-importance perceptions did mediate the already reported effects of the independent variables on the quantity of work completed by subjects, those effects should be nonsignificant in the present analysis.

As expected, the task-importance ratings accounted for a significant amount of variance in the present analysis, 
\[ F(1, 78) = 33.46, p < .001, \eta^2 = .29 \]

For the planned comparison testing Hypothesis 1 (i.e., between the monitoring-only and control conditions for the monitored task), 
\[ t(80) = 1.39, \text{ ns} \]

For the planned comparison testing Hypothesis 2 (i.e., between the monitoring-plus-consequences and monitoring-only conditions for the monitored task), on the other hand, the removal of the variance accounted for by the importance ratings slightly reduced the magnitude of the effect, but did not eliminate it, 
\[ t(80) = 1.99, p < .05, \eta^2 = .05 \]

Taken as a whole, these results suggest that the perceived importance of the experimental tasks mediated the impact of the independent (i.e., nonconjunctive) effect of monitoring but not its conjunctive effect (i.e., its effect when coupled with the delivery of contingent performance consequences). This conclusion is implied by the interrelationship among three findings: (a) The treatment manipulations had a significant effect on the importance ratings; (b) the importance ratings were significantly related to the quantity of work completed by subjects; and (c) when the variance in work quantity accounted for by the importance ratings was controlled, the effect of monitoring alone (but not monitoring coupled with contingent performance consequences) on the quantity of work completed by subjects was eliminated.

Discussion

The purpose of this study was to investigate whether performance monitoring by itself can influence individuals' work
behavior without being coupled with subsequent managerial action (e.g., delivering performance feedback, giving verbal rewards and punishments). We postulated that a supervisor's monitoring activity can serve as a cue signaling the relative importance of various tasks at work and that the perceived importance of a task in turn affects the amount of effort devoted to it. This importance-mediated impact of performance monitoring was hypothesized to be independent of, and in addition to, the conjunctive effect that monitoring traditionally has been assumed to have when coupled with subsequent managerial action.

The results of the study strongly support this conceptualization. The amount of work completed on the experimental tasks increased significantly when performance on those tasks was monitored (in comparison with when it was not), and this effect appeared to be mediated by the tasks' perceived importance. In addition, there was a further increase in the amount of work completed when monitoring was followed by the delivery of contingent performance consequences. This latter effect, however, seemed not to be mediated to any large extent by the perceived importance of the task.

The independent (i.e., nonconjunctive) effect of monitoring on the subjects' task performance can be explained by a social-information-processing view of work behavior, especially as it relates to role-making processes in organizations (Graen, 1976; Katz & Kahn, 1978). Organizational work roles are sets of behaviors that position occupants are expected to enact. Expectations about role behavior can be held and communicated by many different organizational members, not the least of which is one's immediate work supervisor. An important implication of social-information-processing theory is that these role expectations can be communicated (albeit with varying degrees of accuracy) through a wide variety of behavioral cues. Included here are supervisory actions, such as performance monitoring, that may have been initiated for other purposes, but that nevertheless convey to the position occupant expectations for his or her behavior. Individuals will often be acutely attuned to such cues, especially when they are experiencing high levels of role ambiguity (cf. Ashford & Cummings, 1983).

Within this framework, the importance ratings gathered in the present study can be viewed as indices of one aspect of the experimental work role perceived by subjects. As such, different patterns of importance ratings signify different perceptions of that work role, and it is these perceptions that in part guide the subjects' work behavior. The subjects' motivation to enact their perceived work roles can be readily accounted for in expectancy theory terms, because compliance with role expectations often has intrinsic as well as extrinsic value for individuals (Graen, 1976).

Turning to the conjunctive effect of monitoring, at least two mechanisms may have contributed to this effect. One is operant conditioning (cf. Komaki, 1986) and is suggested by the fact that each performance consequence delivered to subjects included an appropriately worded positive or negative evaluative statement. Thus, subjects may have been reinforced by the positive evaluative statements, causing them to maintain good performance, and punished by the negative evaluative statements, causing them to change (i.e., improve) poor performance. The second mechanism that may have contributed to the conjunctive effect of monitoring involves goal setting. It is possible that the feedback subjects received about their performance early in the experiment prompted them to set specific goals for improvement (cf. Klein, 1989) and that these goals in turn led to subjects' increased productivity (cf. Locke, Shaw, Saari, & Latham, 1981).

The present study was not designed to distinguish the relative contribution of these two mechanisms to the conjunctive effect of performance monitoring. Nevertheless, it is clear that that effect, whatever its source, was sufficiently unrelated to the perceived importance of the two experimental tasks to remain visible even when perceived importance was held constant. At the very least, therefore, this suggests that the processing of social cues for the role expectations they may imply is only one among several potentially important mechanism underlying motivated work behavior.

Although it is clear from the present study that monitoring can have a beneficial impact on performance even when it is not followed immediately by the delivery of contingent performance consequences, the monitoring–productivity relationship is probably not a simple linear function. On the contrary, although some monitoring appears to be good for performance, more monitoring may not be better. Indeed, it seems highly likely that monitoring can be carried to such an extreme that it may actually have a negative effect on productivity because it is seen as representing excessively close supervision. In this regard, it should be noted that although Komaki (1986) found that effective managers monitored the performance of their subordinates nearly 50% more often than did those who were less effective, monitoring still accounted for only 2.9% of the effective managers' time. Moreover, what was being monitored in the present study was productivity, not behavior. It seems quite probable that either a dramatic increase in the rate of monitoring or a shift toward monitoring behavior instead of performance will reduce employees' feelings of autonomy on the job, thwart efforts toward self-management, and generally have a deleterious effect on motivation and performance. The trick, then, is to discover the optimal level of monitoring for maximizing performance. This optimal level may well vary according to task, setting, and individual subordinate characteristics (cf. Komaki et al., 1989).

In the present study, improvements in performance quantity and quality on the monitored task came at the expense of performance quantity (but not quality) on the nonmonitored task. This finding is perhaps most easily interpreted in terms of Naylor et al.'s (1980) contention that motivation is essentially a resource allocation process. If performance on a task is to improve, barring any changes in skills, abilities, or situational constraints, that improvement must come as a result of devoting more personal resources to the task. This means spending more time on the task, exerting more effort per unit of time, or both. Spending more time on a task speaks to the directional component of motivated behavior, whereas exerting more effort per unit of time focuses on the intensity component. In the present study, subjects apparently chose to redirect their behavior (as opposed to simply increasing its intensity) in order to improve productivity on the monitored task without risking any decrease in performance quality. That is, although no direct measure of "time in task" was obtained, it seems reasonable to
speculate that the observed increases in performance on the monitored task in the monitoring-only and monitoring-plus-consequences conditions, and the complementary decreases in performance on the nonmonitored task in those conditions, came about through progressive shifts in the allocation of time such that increasingly more time was devoted to the monitored task and increasingly less was spent on the nonmonitored task.

That subjects elected to improve their productivity on the monitored task by changing the direction rather than the intensity of their work behavior is perhaps not so surprising. As Eden (1988) suggested, employees may often "groove in" to an overall level of effort on the job (i.e., the total amount of effort expended across all activities combined) and strive to maintain that level over time (cf. Gardner & Cummings, 1988). This, of course, does not mean that they cannot, or will not, ever increase the total amount of effort they expend. Rather, it merely implies that, when alternative strategies are available for attaining performance objectives, subjects may prefer not to increase their overall level of effort. As the results of the present study suggest, one such alternative may be to allocate time to tasks in proportion to their perceived importance. This should pose no particular problem as long as the perceived importance of various tasks parallels their actual importance to the organization. When this parallelism does not exist, however, overall job performance may not improve, and one set of performance deficits may simply be substituted for another. Such is likely to be the case when behavioral cues that can signal the relative importance of various tasks (e.g., monitoring) themselves do not mirror the tasks' actual importance. This situation might occur, for example, when performance on tasks considered to be less important is nevertheless much easier to monitor than is performance on more important tasks.

The implication for managers and supervisors, then, is that they must be careful that their monitoring and delivery of contingent performance consequences does not improve productivity in one task area at the expense of other areas that are also critical to the job. The time and effort that employees redirect to the monitored task(s) must come from activities that objectively are either less essential or irrelevant (e.g., daydreaming, schmoozing) to organizational functioning. Further research should investigate how managers might better influence not only where resources are directed (e.g., by monitoring and providing contingent performance consequences) but also where those resources are drawn from.

Because the present study has several limitations, replication and extension are needed to demonstrate the generalizability of the obtained results. First, although the unit of analysis was the individual subject, the unit of random assignment was the experimental session. We chose this procedure to accommodate logistical constraints. Though we have no evidence that this procedure in and of itself compromised the integrity of the results as presented, it does violate an assumption underlying our statistical analyses. A replication that eliminates this problem is therefore in order. Along the same lines, our random assignment procedure did not yield completely equivalent groups on one of the pretest measures. Although we compensated for this by statistically controlling for the observed difference, a replication that accomplishes the same end solely through random assignment would be useful.

Second, additional research is needed to show that the obtained results hold over longer time intervals and for different task configurations. It would be especially valuable from a theoretical point of view to know what happens when the work role consists of only a single task. The conceptual analysis presented here implies that, for subjects to improve productivity in such a situation, they would have little choice but to increase the intensity (as opposed to changing the direction) of their work behavior. This implication needs to be verified.

Finally, additional field research on the impact of performance monitoring is also needed. Although field studies have shown monitoring to distinguish effective from ineffective managerial performance in samples as varied as insurance claims office managers (Komaki, 1986) and racing yacht skippers (Komaki et al., 1989), at least one study failed to detect such a difference. Luthans, Rosenkrantz, and Hennessey (1985) found that more effective managers engaged in less monitoring behavior than did less effective managers, but this difference was not statistically significant. Although it is always difficult to interpret the meaning of null results, Luthan et al.'s finding may indicate a boundary condition that delimits the generalizability of monitoring effects. Perhaps, for example, performance monitoring is beneficial only for certain types of employees, in certain types of work settings, or with certain operationalizations of the variables of interest. Further research in the field is needed to explore the possibility of such boundary conditions.

1 Unlike the other studies cited in this article, Luthans, Rosenkrantz, and Hennessey (1985) used a promotional index to define managerial effectiveness. This procedure may have confounded managerial effectiveness with organizational level and job demands, making the results of their study especially difficult to interpret.

References


References

The Publications and Communications Board of the American Psychological Association announces the appointment of Robert J. Sternberg, Yale University, as editor of Psychological Bulletin for a 6-year term beginning in 1991. Beginning immediately, manuscripts should be directed to

Robert J. Sternberg
Yale University
Department of Psychology
P.O. Box 11A Yale Station
New Haven, Connecticut 06520-7447

Received October 28, 1988
Revision received March 20, 1990
Accepted March 21, 1990