TESTING A NEW SUPERVISORY PROCESS FOR IMPROVING INSTRUCTION

GERALD H. ELGARTEN, The City University of New York

The most frequently used technique to evaluate teachers and to improve instruction involves a classroom observation, followed by a post-observation conference and sometimes preceded by a pre-observation conference. Although supervisors might engage in some or all of these components, the general purpose is to increase teachers' effectiveness. Some researchers have proposed that supervisors can increase teacher effectiveness by identifying the techniques and behaviors that are not enhancing instruction and by giving the teacher alternative approaches and strategies that might accomplish lesson objectives. Research shows, however, that teachers believe they learn little from this process.

Therefore, we need to overcome deficiencies in the traditional supervisory process by developing alternatives that will enable teachers to learn more from their interactions with supervisors. This exploratory study constructed a new supervisory process for improving instruction and tested the process with secondary school supervisors and mathematics teachers.

The test of this new model of supervision consisted of comparing two supervisory processes. One, the local traditional supervisory process, consisted of an observation and post-observation conference where the supervisor and teacher collegially, through verbal interaction, developed a list of suggestions for improving the teacher's performance. The other was the

---


experimental supervisory process, which was the same as the traditional model, except that after observing the teacher's lesson the supervisor taught the same lesson to the teacher's next class. The only deviations in the supervisor's lesson were the demonstrations of behavioral changes that the supervisor believed would make the teacher more effective. While observing the supervisor teach, the teacher recorded the differences in his own and the supervisor's lessons. These differences in the two lessons became the focus of the post-observation conference and then, in turn, the suggested behavioral changes for the teacher. In effect, the study determined whether supervisors more effectively transfer suggested behavioral changes to teachers verbally or through demonstration.

Research conducted in an industrial setting indicates that observational learning (also referred to as modeling) effectively changes adult behavior.\(^5\) Observational learning is defined as acquiring cognitive skills and behavior patterns by observing others' performance.\(^6\) According to Bandura's theory, four subprocesses govern observational learning: (1) attentional, referring to observers' ability to focus on the relevant aspects of the modeled behaviors, (2) retention, referring to observers' ability to remember the aspects of the intended modeled activities, (3) production, referring to observers' ability to convert symbolic concepts into appropriate behavior, and (4) motivation, referring to observers' ability to use what they observed. Bandura believes that effective modeling requires the demonstrator to attend to all four subprocesses.

**RESEARCH QUESTION**

Is Bandura's observational learning theory applied to supervising instruction effective in changing teacher behavior?

**Hypothesis 1.** Teachers who are supervised with the use of modeling will show a significantly greater rate of implementation of a supervisor's suggested changes in classroom behavior than teachers who are supervised without modeling.

**Hypothesis 2.** Teachers who leave the post-observation conference with a list of behavioral changes will show a significantly greater rate of implementation of a supervisor's suggested changes in classroom behavior than teachers who leave without a list.

---


Hypothesis 3. An interaction effect will occur: Teachers who leave the post-observation conference that used modeling with a list of suggestions will carry out a significantly greater number of the suggestions than teachers in any other treatment group (modeling without a list, traditional with a list, traditional without a list).

I used the following rationale for the study:

- evidence suggesting that teachers' classroom behavior can be changed to increase student learning
- Bandura's observational learning theory
- substantial empirical evidence showing that observational learning effectively changes adult behavior

A RESEARCH MODEL OF OBSERVATIONAL LEARNING THEORY

The experimental modeling supervisory process is a research model of Bandura's theory. (1) A one-to-one correspondence exists between the constructs in the theory and the activities of the constructs in the modeling supervisory process, (2) to preserve the structure of the theory, its constructs appear in the supervisory process in the same order as they appear in the theory. Table 1 outlines the modeling process.

Design of the Exploratory Field Experiment

In each of the six participating schools, the mathematics supervisor administered both the experimental modeling and the local traditional supervisory processes, eliciting from each teacher four suggested changes in behavior. The local traditional supervisory process is the same as the experimental process without the supervisor’s demonstration. The traditional
Table 1. Experimental Supervisory Process Using Modeling

<table>
<thead>
<tr>
<th>Supervisor</th>
<th>Teacher</th>
</tr>
</thead>
<tbody>
<tr>
<td>A. Observation and Analysis</td>
<td>A. Teach the lesson.</td>
</tr>
<tr>
<td>1. Observe lesson.</td>
<td></td>
</tr>
<tr>
<td>2. Identify behaviors that need to be introduced or changed.</td>
<td></td>
</tr>
<tr>
<td>B. Teach the same lesson to the teacher's next class. Keep close to the teacher's lesson and style.</td>
<td></td>
</tr>
<tr>
<td>1. Model strengths of the teacher's lesson.</td>
<td></td>
</tr>
<tr>
<td>2. Model new behaviors that would increase learning in the classroom.</td>
<td></td>
</tr>
<tr>
<td>C. Post-observation conference</td>
<td></td>
</tr>
<tr>
<td>1. Ask the teacher to respond to questions 1, 2, and 3.</td>
<td></td>
</tr>
<tr>
<td>2. Ask the teacher to record behavior changes to carry out in future lessons.</td>
<td></td>
</tr>
<tr>
<td>3. Ask the teacher to record changes in behavior to carry out immediately.</td>
<td></td>
</tr>
<tr>
<td>4. Ask the teacher</td>
<td></td>
</tr>
<tr>
<td>a. How will you carry out the suggestions?</td>
<td></td>
</tr>
<tr>
<td>b. How will you remember to carry out the suggestions?</td>
<td></td>
</tr>
<tr>
<td>(Motivation: Feedback)</td>
<td></td>
</tr>
<tr>
<td>D. The supervisor informally observes the teacher once during the next week to give feedback.</td>
<td></td>
</tr>
<tr>
<td>E. The supervisor informally asks the teacher about progress in carrying out the changes every three days.</td>
<td></td>
</tr>
</tbody>
</table>

(Attentional) (Retention: Symbolic coding)  
B. Observation and analysis—the teacher responds to three questions:  
1. What do I (supervisor) do that you don't do too often?  
2. What do we both seem to do?  
3. What good things do you do that I don't seem to do?  
(Retention: Symbolic rehearsal)  
C. Post-observation conference  
1. The teacher responds to questions 1, 2, and 3.  
2. The teacher records changes in behavior to carry out in future lessons.  
3. The teacher records changes in behavior to carry out immediately.  
4. The teacher responds to questions.  
(Production: Reinforcement)  
D. The teacher practices behavior changes with the behavior list developed at the post-observation conference in view when teaching.  
E. The teacher remains in contact with the supervisor, providing feedback about progress in incorporating new teaching behaviors into lessons.
process is the comparison treatment because it is the dominant approach all supervisors in the school system use. I made no attempt to change the common supervisory practices being used.

Sample

In each school, 8 mathematics teachers were randomly assigned to one of four treatments. experimental modeling supervisory process, with the teacher leaving the post-observation conference with a list of suggested behavioral changes, experimental modeling, with the teacher leaving without a list; local traditional supervisory process, with the teacher leaving the post-observation conference with a list of suggested behavioral changes; and local traditional, with the teacher leaving without a list (for all treatment groups, \( n = 12 \)). For the 48 teachers, the following contextual data were collected. 29 were male (60 percent), and 19 were female (40 percent); the median years teaching, 6; the median age, 37, 41 were liberal arts majors (85 percent) and 7 education majors (15 percent); the supervisors rated the teachers' general performance level as excellent for 3 teachers (6 percent), good for 17 teachers (35 percent), average for 23 teachers (48 percent), and weak for 5 teachers (11 percent).

Therefore, this sample was representative of all high school mathematics teachers in New York City on teachers' gender but not the median years teaching (13). The mathematics chairs participating in the study had an average of 6 years' experience as supervisors. The average experience of mathematics supervisors in the school system is about 7 years.

One week after the post-observation conference, the respective supervisor revisited the teacher's class and compiled an implementation score ranging from 0, indicating that the teacher carried out no suggestions, to 4, indicating that the teacher carried out all suggestions.

Threats to Internal Validity

The supervisors' observation was the instrument used for gathering data. This instrument presents potential problems, such as possible rater bias to a particular treatment and variability in how the raters recorded the data. In this study, the raters (mathematics supervisors who administered all the treatments in their schools) determined, by a yes or no rating, whether teachers had carried out the suggestions. This design introduced another problem: The raters might want to show that they were effective supervisors by being lenient in recording the teacher's use of their suggestions. The supervisors' training experience and lack of involvement in designing the experimental supervisory process minimized these potential threats to the accuracy of collecting data (instrumentation).
To control the variability in how the raters recorded the data, the training program conducted before the experiment clearly delineated what to accept as implementation of a behavior change. Simulations identified what constituted an implemented behavior. Also, because all 6 raters observed all groups, any variability should occur equally among all groups and should therefore not be a factor.

The supervisors in this experiment have successfully passed the New York City licensing examination for secondary school supervisors of mathematics. These experienced supervisors suggested changes in teacher behavior and then acted as rating teachers in their own departments to determine whether the teacher carried out the suggestions. These activities were normal job responsibilities of these mathematics supervisors.

All supervisors were aware that their teacher ratings would be compiled with all other schools and not analyzed separately. Also, because the supervisors were carrying out all treatments and were not involved in designing any of them, they should have no bias toward a particular treatment. In addition, before the experiment, the supervisors indicated that they did not find a high degree of implementation, by teachers, of their suggestions. This admission indicates that they would not be too lenient in evaluating the implementation of their suggestions as a way of concealing their own weaknesses. Therefore, the possible threat to instrumentation caused by rater bias to a treatment, or by raters trying to show effectiveness by giving high ratings, was minimal.

DATA ANALYSIS

Table 2 summarizes the mean implementation scores for teachers in the four treatment groups. To determine whether there were any significant differences in the means reported in Table 2, I computed a two-way analysis of variance (ANOVA). Table 3 summarizes the results.

DISCUSSION AND CONCLUSIONS

Hypothesis 1. The two-way ANOVA indicated significant differences between the groups on the number of suggestions carried out. The differences were due to the type of supervision alone. The mean for the modeling group \( (M = 3.17) \) was significantly different from that of the traditional group \( (M = 2.13) \), \( F(1, 44) = 15.88 \) \( (p < .001) \). Based on the significant difference in the teachers' implementation rate of their supervisors' suggestions for the modeling and traditional supervisory processes, we can conclude that teachers tended to carry out more suggestions when they saw the suggestions modeled (demonstrated) by their supervisors. This finding supports Hypothesis 1 and is consistent with the findings of empirical studies conducted in an industrial
### Table 2. Implementation Scores for Teachers in Four Treatment Groups

<table>
<thead>
<tr>
<th>Type of supervision</th>
<th>With list</th>
<th>No list</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Modeling</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>M</td>
<td>3.42</td>
<td>2.92</td>
<td>3.17</td>
</tr>
<tr>
<td>SD</td>
<td>.67</td>
<td>.90</td>
<td>.82</td>
</tr>
<tr>
<td>n</td>
<td>12</td>
<td>12</td>
<td>24</td>
</tr>
<tr>
<td>Traditional</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>M</td>
<td>2.25</td>
<td>2.00</td>
<td>2.13</td>
</tr>
<tr>
<td>SD</td>
<td>1.14</td>
<td>.85</td>
<td>.99</td>
</tr>
<tr>
<td>n</td>
<td>12</td>
<td>12</td>
<td>24</td>
</tr>
<tr>
<td>Total</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>M</td>
<td>2.83</td>
<td>2.46</td>
<td></td>
</tr>
<tr>
<td>SD</td>
<td>1.09</td>
<td>.98</td>
<td></td>
</tr>
<tr>
<td>n</td>
<td>24</td>
<td></td>
<td>24</td>
</tr>
</tbody>
</table>

NOTE. Scores ranged from 0, indicating the teacher carried out no suggestions, to 4, indicating the teacher carried out all the suggestions.

### Table 3. Two-Way Analysis of Variance of Implementation Scores for Teachers in Four Treatment Groups

<table>
<thead>
<tr>
<th>Source</th>
<th>Sum of squares</th>
<th>df</th>
<th>Mean square</th>
<th>F</th>
</tr>
</thead>
<tbody>
<tr>
<td>Main Effects</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Type of supervision</td>
<td>13.021</td>
<td>1</td>
<td>13.021</td>
<td>15.878*</td>
</tr>
<tr>
<td>With list/no list</td>
<td>1.687</td>
<td>1</td>
<td>1.687</td>
<td>2.058</td>
</tr>
<tr>
<td>Interaction</td>
<td>.187</td>
<td>1</td>
<td>.187</td>
<td>.229</td>
</tr>
<tr>
<td>Residual</td>
<td>36.083</td>
<td>44</td>
<td>.820</td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>50.979</td>
<td>47</td>
<td>1.085</td>
<td></td>
</tr>
</tbody>
</table>

* *p < .001
environment, as well as in the training of student teachers. Both of these studies reported that modeling improved the subjects' performances in the behaviors being modeled.

For effective observational learning, Bandura identifies four subprocesses that the model must attend to. The first subprocess is attentional and refers to the model's ability to focus observers' attention on the behaviors being modeled. When the supervisors taught the same lesson to the teachers' next classes, they were incorporating the attentional subprocess into the experimental supervisory model. The supervisors asked the teachers, "What do you see me doing that you don't do too often?" Their responses, developed while they observed their supervisors teaching their classes the same lesson they had just taught, indicate that the teachers were able to identify the modeled behaviors. Follow-up interviews with the participating supervisors also confirmed this finding.

The second subprocess is retention and relates to observers' ability to remember the important behaviors being modeled. Gerst points out that developing symbols to represent observed behaviors aids retention. The raw data from the modeling groups indicate that from the same process used for identifying modeled behaviors (teachers observing their supervisors teach their classes the same lesson), the teachers produced symbolic representations of the modeled behaviors—their own written description of the behaviors. The teachers wrote the behavioral descriptions in response to the supervisors' question, "What do you see me doing that you don't do too often?"

Bandura has found that when symbolic codes are rehearsed, observers increase their retention of the behaviors. Follow-up interviews with the participating supervisors reveal that in the groups using modeling, the

---


teachers' lists of identified good teaching behaviors became the focus of the post-observation conferences. The supervisors' focus on the teachers' lists of potential behavior changes during the post-observation conferences shows that the model attended to the retention process of Bandura's theory.

The production process refers to observers' ability to convert the symbolic codes into appropriate behaviors. Bandura points out that people will model less than they observed if they are not actively encouraged and do not receive positive feedback. Follow-up interviews reveal that 5 of the 6 participating supervisors did not actively encourage or give positive feedback to their teachers. Therefore, 5 of the 6 models (supervisors) did not attend to the production process. The teachers' tendency to carry out a significantly greater number of suggestions when the supervisory process used modeling may contradict Bandura's emphasis on positive feedback during the production process. Alternative explanations exist for the teachers' high implementation rate of suggestions without positive feedback from their supervisors: The lack of positive feedback did not overcome modeling's powerful effect when examining the implementation one week after the post-observation conference; a decrease in the teachers' implementation rate may occur several weeks after the post-observation conference; the teachers may have carried out even more suggestions if their supervisor had used positive feedback; or the feedback was not necessary.

The motivation process refers to observers' ability to put what they observed into practice. Bandura, Adams, Hardy, and Howells and Schunk believe that when observers see a successful behavior carried out, they tend to believe that they can also perform the behavior. All the supervisors said they believed a major strength of the experimental supervisory process was that the teachers were able to see the suggestions actually carried out with their own students. This process of demonstrating behavioral changes in the context of their own class may have increased the teachers' sense of efficacy and, in turn, may have motivated the teachers to carry out the supervisors' suggestions.

Hypothesis 2. The two-way ANOVA indicated no significant differences between the groups leaving the post-observation conference with or without a list of suggestions on the number of suggestions carried out. The mean for

---

15Ibid.
17Dale H. Schunk, "Self-efficacy Perspective on Achievement Behavior" (Rockville, MD, August 1982, ERIC Document No. 226 293).
the group leaving the post-observation conference with a list of suggestions (M = 2.83) did not significantly differ from that of the group leaving the post-observation conference without a list of suggestions (M = 2.46), F(1, 44) = 2.058 (p < .05).

Therefore, giving teachers a list of suggestions at the end of a post-observation conference does not tend to affect how well they carried out the suggestions. This finding does not support Hypothesis 2, which is based on my prediction, from the results of a pilot study, that giving teachers a list of suggested behavior changes would increase their implementation rate of their supervisors' suggestions. In follow-up interviews, however, the supervisors all agreed that they found no evidence in their teacher contacts following the post-observation of the teachers actually keeping the list of suggestions. The supervisors said they did not see any of the lists in view when they revisited the teachers' classrooms. Therefore, this study did not determine the effects on the teacher's implementation rate of their supervisor's suggestions that would stem from keeping a list of suggestions in view while teaching.

Hypothesis 3. The two-way ANOVA indicated no interaction effects between the groups in the number of suggestions carried out, F(1, 44) = .229 (p < .05). This finding does not support Hypothesis 3, which was based on my prediction, from the results of a pilot study, that giving teachers a list of suggested behavior changes would significantly increase their implementation rate of the suggestions for both the traditional and the modeling supervisory processes.

A possible reason for the lack of a significant interaction effect was that the teachers did not use the list of suggestions as a reminder to carry out the behavioral changes. Perhaps expecting professionals to carry a list around with them was unreasonable. This behavior appears not to fit with most teachers' self-images. If teachers had the option of leaving the post-observation conference with or without a list, however, those selecting the list may be more inclined to use it.

Limitations

1. The qualifications of the persons using the proposed supervisory model. All supervisors in this study were subject-area specialists. Any attempt to generalize these results to supervisors who are not specific subject specialists would be questionable.

2. The classroom situation where the supervisory process was being carried out. All modeling behavior changes occurred in mathematics classrooms with the specific teacher's class, teaching the same lesson. Any attempt at using the results of this study in a teaching situation that differs from this pattern would be questionable.
Recommendations

1. When modeling is used in the supervisory process, teachers carry out significantly more suggestions than when the traditional supervisory process is used. Secondary school principals might encourage their mathematics supervisors and other subject-area specialists to use modeling in the supervisory process. Principals should also consider using modeling in the supervisory process when observing classes in their own subject area of specialization. Schools with no subject-area specialists or supervisors should consider training master teachers to use modeling with their peers.

2. Two evaluation studies by the Rand Corporation have reported a significant relationship between teachers' sense of efficacy and student achievement. Based on the findings of the Rand studies, researchers from the University of Florida conducted a teacher-efficacy study. The purpose of the research was to further investigate the relationship between the teachers' sense of efficacy and student achievement. In a final report in 1983, Ashton, Webb, and Doda cite the major findings as supporting the Rand results. They report that the teachers' sense of efficacy was (1) significantly related to students' achievement in high school basic skills classes, (2) related to school organizational structures (team organization and teacher involvement in decision making); and (3) related to a warm, accepting classroom climate.

In their conclusions, Ashton and others have supported the findings of Bandura, Guskey, Hillman, and Schunk that perceived self-efficacy is not a fixed construct. Instead, self-efficacy is negotiated daily with each teacher interaction with students, parents, colleagues, and administrators. These researchers suggest that future research explore the processes by which instructional techniques and administrative strategies can "reduce the threats and..."
increase the support of teachers' sense of efficacy. Based on the empirical evidence of the strong relationship between teachers' sense of efficacy and student achievement and further evidence that our perceived self-efficacy can change, I recommend that future studies using modeling in the supervisory process examine the effects of modeling on teachers' sense of efficacy.

3. The use of modeling in the supervisory process is effective in changing mathematics teachers' classroom behaviors. Modeling might have similar effects on principals. Therefore, I recommend applying the same supervisory process outlined in this study to principals' inservice training. For example, an experienced principal might be assigned to help a novice principal improve faculty conference presentations. After observing the novice principal make a presentation at a morning faculty conference, the more experienced principal would make the same presentation, modeling some suggested behavior changes, at an afternoon conference. During the afternoon conference, the novice principal would be responding to the same three questions asked in the experimental supervisory process. (1) What do I [the supervisor] do that you don't seem to do? (2) What do we both seem to do? (3) What good things do you do that I don't seem to do? The responses would then become the focus of the post-observation conference.

GERALD H. ELGARTEN is Associate Professor of Mathematics Education, The City College, The City University of New York, North Academic Center 6/207, Convent Avenue at 138th Street, New York, NY 10031.

---

