Representational Guidance of Action Production in Observational Learning: A Causal Analysis

Wayne R. Carroll
University of Arizona

Albert Bandura
Stanford University

ABSTRACT. This experiment tested the hypothesis that the number of model presentations and verbal coding of modeled actions affect reproduction accuracy through their effect on cognitive representation. Subjects viewed a complex action pattern either two or eight times with or without verbal coding to highlight the dynamic structure of the component actions and their temporal sequencing. They then received, in order, a recognition test and a pictorial-arrangement test to assess the accuracy of their cognitive representations of the modeled actions. Subsequently, all subjects were tested for their ability to reproduce the action pattern from memory. Results showed that increased exposure to modeled actions enhanced the accuracy of both the cognitive representation and the behavioral reproduction. Verbal coding also increased cognitive and reproduction accuracy, but only when combined with multiple opportunities to observe the modeled actions. A causal analysis confirmed that the effects of multiple exposures and verbal coding were entirely mediated by changes produced in the accuracy of cognitive representation.

OBSERVATION OF SKILLED ACTIONS provides an effective means of accelerating their acquisition. Although the power of observational learning has been amply documented (Gould & Roberts, 1982), the reported studies, for the most part, have not been programmatic or adequately guided by theory. Moreover, they have suffered from a methodological problem in that the level of observational learning has been inferred solely from performance of the modeled actions (Newell, Morris, & Scully, 1985). People do not always enact everything they learn. Without independent measures of cognitive representations of modeled actions, it is difficult to determine how much learning has

This research was supported by Public Health Research Grant MH-5162—25 from the National Institute of Mental Health. We gratefully acknowledge the assistance of Jean Di Francis, Scott Irving, and Mark Wiseley in the conduct of the experiment.

Requests for reprints should be sent to Albert Bandura, Stanford University, Building 420, Jordan Hall, Stanford, CA 94305.
occurred through observation and how cognition guides skilled enactment (Adams, 1987).

According to the social cognitive theory of observational learning (Bandura, 1986) that has guided our own research, acquisition of modeled patterns of behavior is governed by four constituent processes. Information conveyed by modeled performances about the dynamic structure of action patterns is extracted through selective attention to spatial and temporal features and is transformed into a cognitive representation by symbolic coding and cognitive rehearsal. The cognitive representation both guides the production of skilled action and provides a standard against which to make corrective adjustments in performance.

In addition to attentional and representational processes that determine acquisition and retention of cognitive representations, a conception-matching process governs the translation of representations into action. Monitored performance provides the necessary information for detecting and correcting mismatches between conception and action. And finally, motivational processes facilitate acquisition of cognitive representations through their effect on attentional and retentional processes and regulate performance by motivating observers to execute what they have learned observationally.

The proposition of social cognitive theory, that modeled actions are first organized cognitively before being translated into action, has received some empirical support. Previous research has shown that visual monitoring of subjects' attempts to reproduce a sequence of modeled actions consisting of many subcomponents has no beneficial effect until a cognitive representation, assessed independently, has achieved more than rudimentary development (Carroll & Bandura, 1982, 1985, 1987). Correlational analyses have further shown that the better the cognitive representation the more accurate are the subsequent reproductions of action patterns (Carroll & Bandura, 1985).

The preceding studies were primarily designed to investigate the relationship between development of cognitive representations and self-monitoring of one's enactments. All subjects, therefore, alternated between observing the modeled actions and performing them, prior to being tested for their ability to reproduce the actions from memory. Consequently, one cannot separate the effects of exposure to the modeled actions on reproduction accuracy and development of cognitive representations from the possible contributory effects of performance elements.

The present experiment was designed to examine the role of representational guidance in observational learning by inducing differential levels of cognitive representation without any intervening behavioral enactments. Cognitive representation was measured before subjects were tested for their ability to reproduce the modeled actions and was manipulated by varying the number of presentations of the modeled actions. Multiple exposures provide observers with opportunities to discern the structure of the modeled actions, to organize and verify what they know, and to give special attention to problematic aspects in subsequent exposures. Under limited exposure, observers are much less likely to acquire knowledge of all the crucial features of the modeled actions.

A second way of manipulating cognitive representation was by verbal coding of the modeled actions as they were being exhibited. When people simply observe entire performances, they often fail to grasp crucial details that may vary in salience, discriminability, and complexity. Verbal coding that depicts in words what is being modeled behaviorally helps to channel observers' attention to essential features of modeled activities and makes them more recognizable (Bandura, 1986; Sheffield & Maccoby, 1961). Additionally, verbal coding strategies provide subjects with a means of transforming modeled information into easily remembered linguistic codes (Bandura & Jeffery, 1973; Bandura, Jeffery, & Bachica, 1974).

The present experiment used a factorial design in which subjects received either two or eight model presentations of a complex action pattern. Within each of these exposure conditions, subjects observed the modeled actions either with or without concurrent verbal coding that described each of the action components and how to execute the transitional movements between them. Following the exposure trials, the accuracy of subjects' cognitive representation of the modeled activity was measured. In the final phase of the experiment, subjects reproduced the modeled action pattern from memory. It was predicted that the conditions combining multiple model presentations with verbal coding would produce the highest level of cognitive representation and that minimal modeling without the benefit of verbal coding would be least effective in this regard. It was further hypothesized that increased accuracy of cognitive representation would be accompanied by an increased level of behavioral production of the modeled action pattern.

A second, and theoretically more important, goal of the present study was to provide a stringent test of the mediating role of representational guidance of action production in observational learning. It was hypothesized that both the effect of multiple model presentations and verbal coding on accuracy of reproducing the modeled action pattern would be mediated by changes in cognitive representation.

Validation of this causal model requires supporting evidence for several forms of linkage between inducing conditions, cognitive representation, and action (Judd & Kenny, 1981). First, it must be shown that multiple model presentations and verbal coding affect reproduction accuracy. Without this evidence, it would be pointless to investigate the mediating role of cognitive representation. Second, it is necessary to show that the treatment conditions affect the cognitive representation. Third, there must be an effect of cognitive representation on reproduction accuracy when the effect of the treatment variables is statistically controlled. Finally, it is necessary to show that the effect of multiple model presentations and verbal coding on reproduction ac-
accuracy is eliminated when cognitive representation is statistically controlled. Support for this last effect, if the other effects are also corroborated, provides the most stringent test of the validity of the causal contribution of cognitive representation to action. Evidence of complete mediation of the treatment effects would establish cognitive representation as not only a necessary, but a sufficient factor in the causal chain.

Method

Subjects

Twenty-eight male and 28 female right-handed undergraduates, enrolled in introductory psychology classes, were randomly assigned in equal numbers to the four experimental conditions. The subjects were fulfilling a requirement for research participation.

Modeling Stimuli and Apparatus

The modeled action pattern and paddle device used by subjects for the reproduction test were the same as those used in a prior experiment (Carroll & Bandura, 1982). Hence, only a brief description of these aspects of the procedure will be presented here.

Each subject watched a video monitor showing a male model performing a complex action pattern containing nine different response components that varied in the spatial configuration and movement of the arm, wrist, and paddle. The first response component or starting position of the modeled action pattern was presented for 5 s. The eight subsequent response components were then each modeled for 2 s, with a 1.5-s transitional movement between the component actions. The complete action pattern took 33 s to execute. This complex action pattern was constructed so as to encompass common aspects of intricate activities. It required both correct patterning and temporal sequencing of actions, some of which were readily codable, whereas others contained features that were highly subtle.

The modeled display presented only the extreme right portion of the body, as videotaped from behind the model, so that observers would not have to transform the modeled actions. A video monitor (48.26 cm) was used to play back the action pattern. The camera angles for recording the model's and subjects' subsequent performance were kept approximately equal, so that the visual stimuli resulting from the model's performance would closely approximate those that subjects, themselves, would receive when attempting to reproduce the demonstration. Previous research has shown that a marked discrepancy between these camera angles tends to retard observational learning of intricate activities (Rosal et al., 1961) and avoidance responses (Grenewald & Albert, 1968).

A second videocassette recorder was used to record subjects' reproductions of the modeled action pattern. When subjects were not observing the demonstration or being tested for their ability to reproduce the modeled actions, a neutral gray, imageless raster appeared on the monitor.

Experimental Conditions

Subjects were seated before the video monitor. The angle of the video camera, located behind them, was adjusted so as to make visual feedback from each subject's performance similar to visual feedback from the modeled action pattern when the test for reproducing it was administered. Subjects were informed by a male experimenter that the study dealt with the learning of movement patterns. He also instructed them to watch the video monitor at all times.

All subjects observed the same sequence of modeled actions, but half of the subjects observed the actions twice whereas the remainder observed them eight times. Within each of these exposure conditions, half of the subjects received concurrent verbal coding that had been recorded on the audio track of the videocassette used to present the modeled action pattern. The remaining subjects viewed the action pattern without verbal coding. The interval between model presentations was kept constant at 10 s.

The verbal coding described for each of the component actions the movement of the paddle as well as that of the wrist, arm, and shoulder. Pilot research indicated that presenting movement information using kinesiological terminology, such as flexion and extension, was not effective. Consequently, the verbal information was given in a simplified form, for example, "Without changing the angle of the paddle face, move it forward and down."

After subjects had completed viewing the scheduled number of demonstrations, they received a recognition test of their knowledge of the correct component responses and a pictorial-arrangement test of their knowledge of the correct sequence of component responses. Upon completion of these two tests that assessed the accuracy of cognitive representation of the modeled actions, all subjects were instructed to pick up the paddle and to reproduce the demonstration as accurately as possible. Because previous research has shown that subjects perform the action pattern used in the present study more accurately if allowed to visually monitor their actions (Carroll & Bandura, 1982, 1985, 1987), subjects were able to see their actions on the video monitor as they performed them.

Cognitive Representation

In the recognition test, subjects were presented photographs of the nine response components of the action pattern along with photographs of three highly similar distractors for each component. The response components and distractors were photographed directly from the video monitor to ensure equivalence of the former to the recorded
modeled pattern. Each component and its distractors were mounted on a separate page, according to a predetermined random sequence, with the restriction that no two components could occur on successive pages in the same sequence as displayed in the modeled action pattern. The spatial arrangement of components and distractors on each page was also randomly determined. Subjects recorded their choice of correct components by writing the alphabetic letter printed next to each photograph. Ten seconds were allowed for each of these choices. The accuracy of cognitive representation was scored by awarding 1 point for each correct choice. The maximum score was 9.

In the pictorial-arrangement test, which immediately followed the test for recognition, subjects were shown photographs of the nine response components, depicted in a scrambled order. They were instructed to arrange the photographs from left to right in the order which accurately reflected the sequence of component responses exhibited by the model. The scrambled order was randomly selected, with the restriction that no two components could occur in the same order as depicted in the demonstration. Subjects were allowed a maximum of 2 min to complete this task. The accuracy of cognitive representation was scored by awarding 1 point for any two response components correctly sequenced. The maximum score was 8.

Following the tests for cognitive representation and after being told they would no longer see the modeled action pattern, all subjects were tested for reproduction accuracy.

Scoring of Reproduction Accuracy

Each response component and preceding transition movement were scored for reproduction accuracy. These segments were played back and viewed separately by freezing the frame or frames at which subjects completed the action component. Subjects were awarded 2 points for a perfect match to the modeled component in form and sequence. One point was awarded if the reproduction contained a minor, but discernible, error in wrist, arm, or paddle position, on component or transition movement, or if the component was correct but produced out of sequence. Subjects received no points if their component reproduction differed markedly from the modeled pattern in one or more features. The more errors subjects made in form and sequence, the lower was their reproduction score. The maximum score possible was 18 points.

Reproductions of a sample of pilot subjects (n = 16) were independently rated by two judges to ensure proficiency in using the scoring criteria. To increase inter scorer reliability, photographs of each response component were displayed while subjects’ reproductions were being scored. The same judges then independently rated the performance of all the subjects in the main experiment. The judges had no knowledge of the hypotheses or the conditions to which subjects had been assigned. The interscorer reliability was r = .91.

The mean ratings of the two judges were used for all subsequent analyses involving reproduction accuracy.

Results

The effects of number of model presentations (two or eight) and verbal coding (with or without) were analyzed by a 2 × 2 factorial ANOVA. Prior to performing an ANOVA on each of the dependent variables, a multivariate analysis of variance (MANOVA) was performed as an omnibus test of significance.

Since the gender factor did not approach significance on any of the three response measures, the data were pooled for subsequent analyses.

The multivariate analysis (MANOVA) yielded significant main effects for number of model presentations, F(3, 50) = 41.84, p < .0001, for verbal coding, F(3, 50) = 5.54, p < .005, and for the interaction between number of model presentations and verbal coding, F(3, 50) = 2.87, p < .05. Univariate ANOVAs were then performed on each of the dependent variables.

Reproduction Accuracy

As the right panel of Figure 1 shows, eight model presentations produced higher reproduction accuracy than did two, F(1, 52) = 49.36, p < .0001. Although the main effect of verbal coding did not approach significance, the interaction between number of model presentations and verbal coding was significant, F(1, 52) = 6.22, p < .025. Analysis of the simple effects of this interaction indicated that subjects did not profit from verbal coding of two model presentations, F(1, 52) = 1.81, ns, but they did when they had the benefit of eight exposures to the modeled actions, F(1, 52) = 4.75, p < .05.

Cognitive Representation

On the recognition test, subjects who had been exposed to the modeled pattern eight times were significantly more accurate at distinguishing correct response components from incorrect ones, F(1, 52) = 43.15, p < .0001. Neither verbal coding nor its interaction with number of model presentations approached significance.

On the pictorial-arrangement test, subjects who had received eight presentations were significantly better at sequencing the depicted response components than those who saw the modeled actions only twice, F(1, 52) = 44.55, p < .0001. Provision of verbal coding also enhanced the accuracy of the cognitive representation, as measured by the pictorial-arrangement test. F(1, 52) = 6.53, p < .025. Finally, the interaction between number of model presentations and verbal coding was also found to be significant, F(1, 52) = 4.95, p < .05. Analysis of the simple effects revealed that subjects who had eight
model presentations with verbal coding sequenced the action components more accurately than those who had the multiple exposures without having the modeled actions verbally coded, \( F(1, 52) = 11.42, p < .005 \). In contrast, verbal coding had no effect on those who had only two model presentations.

As a comparison of the left and right panels of Figure 1 reveals, the interaction between number of model presentations and verbal coding produced highly similar effects on the pictorial-arrangement measure of cognitive representation and on reproduction accuracy.

Examination of the zero-order correlations between response measures showed that representational accuracy as measured by the recognition test, \( r = .47, p < .005 \), or the pictorial-arrangement test, \( r = .73, p = .001 \), was significantly correlated with ability to reproduce the modeled actions. The two measures of cognitive representation were also positively related, \( r = .38, p < .005 \).

**Evaluation of the Causal Model**

To test the causal model, that the effect of model presentations and verbal coding on reproduction accuracy is mediated by cognitive representation, procedures described earlier were applied to the data. First, as has been reported, the univariate ANOVAs indicate that the main effect of number of model presentations and its interaction with verbal coding significantly affect both accuracy of the pictorial-arrangement measure of cognitive representation and accuracy of reproduction the modeled actions. Using orthogonally coded vectors to represent the main effects and interaction (Pedhazur, 1982), each of the response measures were regressed on the treatment variables. The standardized beta weights or path coefficients for the effect of number of model presentations and its interaction with verbal coding on the pictorial-arrangement measure of cognitive representation are, in order \( b = .68, p < .0001 \), and \( b = .20, p < .05 \). The effect of the same variables on reproduction accuracy yielded the following standardized beta weights, \( b = .68, p < .0001 \), and \( b = .24, p < .025 \). Although number of model presentations had a significant effect on the recognition measure, \( b = .67, p < .0001 \), the beta weight for the interaction between coding and number of model presentations did not approach significance, \( b = .05 \).

Second, reproduction scores were regressed on the measures of cognitive representation, after the coded vectors representing treatment effects were first entered in the multiple-regression equation. Testing the significance of the increment in explained variance due to the measures of cognitive representation indicated that the latter had a significant effect on reproduction accuracy, \( F(2, 50) = 17.67, p < .0001 \). The standardized beta weights for the recognition and pictorial-arrangement measures are, respectively, \( b = .14, p = .22 \), and \( b = .56, p < .001 \), indicating that only the latter measure has a significant effect on reproduction scores when the treatment effects are controlled. Examination of the standardized beta weights with treatment vectors and the measures of cognitive representation entered in the regression equation indicated that neither number of model presentations, \( b = .20, p = .20 \), coding, \( b = .13, p = .18 \), nor the interaction, \( b = .13, p = .13 \), had a significant effect on the reproduction measure.

In examining the variance explained in the above analyses, it was found that when the measures of cognitive representation were not statistically controlled, the main effect of number of model presentations accounted for 46% of the variance in reproduction accuracy, and the interaction between this main effect and verbal coding accounted for 6%. In marked contrast, when the contribution of the cognitive representation was controlled, number of model presentations accounted for only 2%, and the interaction accounted for only 3% of the variance in reproduction accuracy. Even when treatment effects were statistically controlled, cognitive representation explained 13% of the variance in reproduction accuracy.

In sum, the results of these analyses provide strong support for the validity of the causal model, which postulates that effects of number of model presentations and verbal coding on reproduction accuracy are mediated through development of a cognitive representation. The same causal analysis, using only the pictorial-arrangement measure of cognitive representation, yielded the same pattern of results as those reported using both measures of cognitive representation.
finding, that neither verbal coding nor its interaction with number of model presentations affected the recognition measure of cognitive representation, may have resulted from the failure of verbal coding to capture the more subtle, quantitative features of the movement pattern. Additionally, recognition was tested by presenting the correct components of the movement pattern in a random order rather than in the same order that they appeared in model presentations. Because temporal order appears to be an inherent part of the structure of long sequences of movements (Shaffer, 1980), the use of a random order in the present study may have attenuated the sensitivity of the recognition measure.

Discussion

In accordance with predictions, the findings of the present experiment reveal that multiple exposures to modeled actions increase the accuracy of both the cognitive representation and behavioral production of the actions. Under conditions of multiple exposure, verbal coding that highlighted the structural features of the modeled actions also enhanced the accuracy of both the cognitive representations and behavioral reproductions. Under limited exposure to the modeled actions, however, provision of verbal coding had no significant effect.

Correlational analyses yielded a pattern of results that was similar to previous research using the same action pattern and the same measures of cognitive representation and performance (Carroll & Bandura, 1987). The more accurate the cognitive representation, the more accurate were reproductions of the modeled actions.

The proposition, that the effect of number of model presentations and verbal coding on behavioral production is mediated by changes in the accuracy of cognitive representation, was supported by the results of the causal analysis. Even when the contribution of the treatment variables was statistically controlled, cognitive representation still had a pronounced effect on reproduction of the modeled actions. The strongest support for the validity of the causal model, however, comes from the finding that the effect on behavioral production of multiple model exposures and verbal coding disappears when the effect of the cognitive representation is partialled out.

The conclusion, that an accurate cognitive representation is necessary for proficient action, is supported by converging evidence from other lines of research showing that subjects display little or no observational learning even after repeated exposures to the model if opportunities for symbolic coding are blocked (e.g., Bandura & Jeffery, 1973, Bandura et al., 1974).

The size of the effect of verbal coding on reproduction accuracy merits some discussion. The efficacy of verbal coding can be affected by its form, modality, and locus. Observational learning is enhanced by pretraining in reductive verbal codes symbolizing the modeled actions and by elaborative linguistic coding that imposes a meaningful cognitive structure on movement patterns (Bandura et al., 1974). In the present study, descriptive codes were externally provided in order to standardize the form and duration of coding operation across experimental conditions. Personally generated meaningful codes, however, have greater impact on observational learning than do concrete descriptions of action patterns (Gerst, 1971).

The present study contributes to the growing body of literature that emphasizes the influential role played by cognition in the acquisition and refinement of motor skills (Adams, 1984, 1987; Annett, 1985; Newell et al., 1985). Both verbal coding (Bandura & Jeffery, 1973) and imaginal coding (Gerst, 1971) of movement features have been shown to facilitate reproduction of observationally learned actions. Previous research (Carroll & Bandura, 1987) as well as the present study show that structural features of modeled actions, such as the pattern, magnitude, direction, and sequence of those actions can be cognitively represented. Even such subtle information as the timing requirements of modeled rapid movements has recently been shown to be capable of cognitive representation (Adams, 1986). Further research is needed, however, to determine whether other motor aspects of movement patterns, such as static force and muscle tension, can be symbolically represented (Adams, 1984). To the extent that these features can be portrayed in visual, auditory, or kinesthetic form, they too are amenable to symbolic representation. Persons who are deaf and blind rely on kinesthetic matching of mouth and laryngeal muscular responses of verbalizing models (Young & Hawk, 1955). Laryngographic visual displays of larynx vibrations corresponding to speech patterns are replacing tactile modeling for hearing-impaired persons who have the benefit of sight (Fourcin & Abberton, 1972).

An important issue for future research is the contribution of motor rehearsal to the development of cognitive representations of modeled actions and to reproductive accuracy. Social cognitive theory (Bandura, 1986) has argued that motor rehearsal of modeled actions during learning provides an opportunity to practice transforming cognitive representation into action. Such behavioral enactments serve to refine the more motor aspects of modeled actions as well as to reduce discrepancies in sequence and structure between cognitive representation and performance. Detection and correction of such discrepancies, in turn, depend on the informativeness of the feedback accompanying performance. Motor rehearsal of modeled actions with impoverished feedback leads to markedly flawed performance, as compared to performance with the opportunity to monitor one's actions visually (Carroll & Bandura, 1982, 1985). Monitored enactments, however, have not been found to be effective until the cognitive representation of modeled actions has achieved at least a moderate degree of accuracy (Carroll & Bandura, 1982).

Although research bearing on the issue of motor rehearsal is sparse, motor rehearsal of modeled actions during learning can facilitate their subsequent reproduction when subjects have to create their own sym-
bolic guides to action (Bandura & Jeffery, 1973). There is also evidence that motor rehearsal can enhance cognitive representation as well as increase reproduction accuracy when conditions are not optimal for representational coding (Carroll & Bandura, 1985). Even when experimenter-provided symbolic codes are rehearsed and retained, subsequent performance is sometimes flawed because subjects are unable to accurately translate cognitive representation into action (Bandura & Jeffery, 1973). Similar findings have been reported using assembly tasks which varied in complexity. Symbolic coding of the modeled constructions facilitated their later reproduction. On the more intricate task, symbolic coding followed by motor rehearsal was found to result in the highest level of reproduction accuracy, whereas motor rehearsal alone produced the lowest level (Jeffery, 1976). Since cognitive representation was not measured, however, it remains an open question whether the effect of motor rehearsal on subsequent reproduction accuracy was mediated by cognitive representation.

Conclusive evidence bearing on the role of motor rehearsal in observational learning will require both independent assessment of cognitive representation and a comparison of observational learning with and without the opportunity for behavioral enactments during the acquisition phase. The time at which enactments are introduced during the acquisition phase may determine whether they affect representational processes. Enactments before one has at least some notion of the structure of modeled actions will produce only faulty performances. Flawed enactments are not only uninformative, but may be misleading. After a number of structural features have been cognitively represented, however, attempts to reproduce the behavior help to identify the insufficiently articulated aspects of the representation. To the extent that this heightens and channels attentiveness to the problematic aspects in subsequent modeling, it can help to refine cognitive representation of the activity.

REFERENCES


Carroll, W. R., & Bandura, A. (1982). The role of visual monitoring in observational learn-