Role of Timing of Visual Monitoring and Motor Rehearsal in Observational Learning of Action Patterns

Wayne R. Carroll
University of Arizona

Albert Bandura
Stanford University

ABSTRACT. This experiment tested the hypothesis that observational learning is enhanced by visual monitoring of enactments that is optimally timed for conception-action matching and by motor rehearsal that serves to refine the cognitive representation. Subjects observed a modeled action pattern, after which they enacted it with either concurrent, delayed, or no visual monitoring. They then engaged in motor rehearsal or did not rehearse the action pattern. Development of the cognitive representation of the modeled action was also measured. Concurrent visual monitoring of enactments greatly facilitated observational learning, whereas delayed visual monitoring did not affect the acquisition process. Rehearsal aided cognitive representation and behavioral reproduction. The more accurate the cognitive representation of the modeled action pattern, the more skilled were the subsequent reproductions of it. After gaining proficiency in converting conception to action, subjects showed no decline in reproduction accuracy when modeling and visual monitoring were withdrawn.

PSYCHOLOGICAL THEORIES have traditionally emphasized learning that occurs by performing responses and experiencing their effects. In everyday life, virtually all learning phenomena resulting from direct experience can occur vicariously by observing other people's behavior and its consequences for them (Bandura, 1984; Rosenthal & Zimmerman, 1978). By conveying the structure and underlying rules of behavior, modeling can accelerate the acquisition process.

In observational learning, the information conveyed by modeled activities is transformed into cognitive representations that function as internal models for action. Psychological research tends to focus on issues...

This research was supported by Public Health Research Grant MH-5162 from the National Institute of Mental Health. We gratefully acknowledge the assistance of James Carbone, Lisa Glow, Laurie Hanchett, Greg Kunz, and Richard Zraick in the conduct of the experiment. Requests for reprints should be sent to Wayne R. Carroll, Department of Psychology, University of Arizona, Tucson, Arizona 85721.
concerning either acquisition of knowledge or development of performance skills. The processes governing the interrelationship between knowledge and action have been largely neglected (Newell, 1978). Observational learning of motor skills provides a fruitful paradigm for elucidating the mechanisms by which knowledge is transformed into appropriate action. The present program of research is designed to clarify some aspects of this transformation process.

In the multiprocess theory of observational learning proposed by Bandura (1977, 1984), acquisition of modeled behavior is governed by four constituent processes. Attentional processes determine what is selectively observed and what information is extracted from ongoing modeled events. Models usually present more information than observers can process effectively (Newell, 1981). Observational learning is, therefore, enhanced by factors that channel attention to critical features of modeled performances (Lumsdaine, 1962; Minas, 1980; Yussen, 1974).

A second major subfunction in observational learning concerns retention processes. If modeling influences are to have enduring effects, the modeled information must be transformed and restructured into appropriate symbolic representations, which function as internal models for action. Cognitive activities that aid representational development, such as symbolic coding and cognitive rehearsal, markedly affect the rate and level of observational learning of such diverse activities as complex movement sequences (Bandura & Jeffery, 1973; Bandura, Jeffery, & Bachicha, 1974), intricate manual patterns from sign language (Gerst, 1971), and abstract three-dimensional constructions (Jeffery, 1976).

In the third subfunction of modeling—the behavioral production process—conceptions of modeled activities are translated into appropriate actions. Behavioral production involves a conception matching process in which feedback from action is compared against an internal model. The behavior is then modified by the comparative information to achieve close correspondence between the internal conception and action. If the spatial and temporal features of modeled behavior can be easily discerned, encoded, and cognitively rehearsed, accurate reproduction can be achieved with little or no overt practice or response feedback (Bandura & Jeffery, 1973; Gerst, 1971). However, if features of the modeled activity are more intricate, difficulties arise in translating cognitive representation into action (Martens, Burwitz, & Zuckerman, 1976). Overt performance is needed to detect mismatches between performance feedback and conception. On subsequent opportunities to observe the model, learners will be more apt to encode features of the modeled activity that had been poorly articulated in the initial representation. With representational development and iterative transformation of conception to corresponding action, reliance on performance feedback should decrease (Carroll & Bandura, 1982).

Finally, motivational processes are concerned with incentive factors that govern observational learning and performance. Incentive factors facilitate acquisition through their impact on attentional and retention processes, and affect performance by motivating observers to execute
what they have learned observationally. The motivators may be extrinsic, vicarious, or self-generated by means of internal standards (Bandura, 1984).

A common problem in learning is that in many activities, performers do not fully observe their own behavior because it lies outside their field of vision. Lack of opportunity to monitor response production visually increases the likelihood that mismatches between cognitive representation and performance will go undetected and that corrective adjustments in performance will not be made. Indeed, learners may erroneously assume that they are faithfully reproducing a modeled action pattern, whereas they are actually committing numerous errors.

That behavioral production is greatly aided by conception acquisition and concurrent observation of enactments, has been revealed in a recent study of observational learning of a complex action pattern that normally lies outside the field of vision (Carroll & Bandura, 1982). Learners achieved little accuracy in producing modeled patterns if they could not see what they were doing or if they had observed their enactments on a video monitor before they had formed an adequate cognitive representation of the activity. Without a conception to serve as a standard, they could not use the visual feedback correctly. However, after learners had conceptualized the action pattern, being able to observe their enactments markedly facilitated accurate performance.

The present experiment was designed to test the hypothesis that the timing of visually guided enactments, rather than mere self-observation of enactments, is the critical factor in facilitating the conception-matching process. Concurrent visual monitoring of enactments, immediately upon observing the modeled pattern, makes the correspondence between conception and action salient. In contrast, delayed monitoring of enactments makes it difficult to detect and correct mismatches between conception and action, especially if subjects get absorbed simply in the replay of a past performance without constantly comparing it to the memory of how it should be done. In many instances, of course, the cognitive representation may be too sketchy to provide an adequate internal model for correcting enactments.

Another objective of the present study was to examine the effect of motor rehearsal on accurate reproduction of modeled activity. Previous research (Bandura & Jeffery, 1973) found a weak effect for motor rehearsal on observational learning, but only if the symbolic coding system that observers used was not optimal for representing the modeled performance information. However, there is reason to expect that rehearsal may have a greater impact on observational learning when observers form the cognitive representation through their own coding devices and when modeling and rehearsal of the activity alternate, as they commonly do in everyday life. In efforts to rehearse a previously modeled activity, performers become acutely aware that aspects of their conception of the activity may be lacking or poorly delineated. Recognition of representational deficiencies heightens attentiveness to subsequent modeling to fill in missing elements. By enhancing and channel-
ing attention, motor rehearsal can contribute to refining of the symbolic representation of the modeled activity.

To test these notions, subjects observed a model perform a novel action pattern that would normally be outside their field of vision. After observing the behavior modeled, they were tested for their ability to reproduce it with either concurrent, delayed, or no visual monitoring of their enactments. Following each enactment, subjects in each of the three monitoring conditions either engaged in motor rehearsal or did not rehearse the action pattern. At selected points in the series, the accuracy of subjects' cognitive representation of the modeled activity was measured. In the final phase of the experiment, subjects reproduced the activity without the aid of modeling or visual feedback.

It was predicted that concurrent visual monitoring, which is ideally suited for matching conception to action, would produce the most skilled reproduction of the modeled pattern, whereas delay or lack of visual monitoring would retard the acquisition process. A further hypothesis was that rehearsal would similarly facilitate the acquisition process. Accuracy of cognitive representation was expected to be positively related to skillfulness of subsequent behavioral reproduction. Additionally, it was hypothesized that in the final phase, by which time subjects would have gained facility in translating conception to action, performance could be guided by the internal model of the action in the absence of external modeling or visual monitoring.

Method

Subjects

Thirty male and 30 female right-handed, paid volunteers were recruited from among undergraduate students at the University of Arizona. Subjects of each sex were randomly assigned in equal numbers to each of six treatment conditions.

Modeling Stimuli and Apparatus

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

Each subject watched a video monitor showing a male model performing a complex action pattern containing nine different response components, which varied as to the spatial attributes 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.

The modeled performance presented only the extreme right portion of the body, as recorded from behind the model. A 19-in. (48.26 cm)
video monitor was used to play back the action pattern. The camera angles for recording the model's and subjects' performances 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 knot-tying activities (Roshal, 1961) and avoidance responses (Greenwald & Albert, 1968).

A second videocassette recorder (connected to the same video monitor and concealed by a screen), was used to record subjects' reproductions of the modeled action pattern. It was also used to play back subjects' reproductions for subsequent scoring. A manual switching device connected to the two videocassette recorders allowed subjects in the visual monitoring conditions to observe their enactments on the video monitor, whereas those not scheduled to receive visual feedback saw a neutral gray, imageless raster on the monitor. Connected to the second videocassette recorder was a smaller, 12-in. (30.48 cm) monitor that was used to observe subjects' performances when no image appeared in the larger monitor.

Procedure

Subjects sat before the large video monitor. The angle on 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.

Subjects were informed by a male experimenter that the study dealt with the learning of movement patterns. They were told to put on a pair of plastic safety goggles, which were painted black and which had their lenses removed. Although subjects were unable to view their movements, they wore the goggles to insure the unobservability of the action pattern. Subjects were instructed to watch the television monitor at all times.

After demonstrating the correct grip for holding the paddle handle, the experimenter moved the subject's arm to correspond to the correct starting position twice. Subjects then twice practiced the correct grip and the designated starting position. They were instructed to pay attention to the position of the model's arm, wrist, and paddle in order to reproduce the action pattern accurately. They were also told to report when they had completed their attempt to reproduce the modeled action pattern.

All subjects observed the modeled action pattern 12 times and were verbally cued to reproduce it after every two presentations, each of which was separated by a 10 s interval. Thus, all subjects reproduced the modeled pattern 6 times, followed by a final set of three reproductions during which both modeling and visual monitoring were omitted. In the present study, timing of visual monitoring refers to the degree of
temporal separation between enactment and self-observation of it rather than to the temporal regulation of action sequences.

Subjects who were scheduled to receive concurrent visual monitoring were informed that they would be able to see their actions on the video monitor as they performed them. Subjects who were to receive delayed visual monitoring were told that they would be able to see their actions sometime after they completed them. Subjects who performed without visual monitoring at any time were merely told to focus their eyes on the imageless monitor while performing.

After each attempt to reproduce the modeled action pattern, half of the subjects in each of the above monitoring conditions were told to count backwards by threes from a three-digit number supplied by the experimenter and changed every 20 s. Backward counting was used as a device to prevent subjects' cognitive rehearsal of the modeled information. The remaining half of the subjects assigned to the motor rehearsal condition were verbally cued to continue to reproduce the action pattern. Within each monitoring condition, the amount of time allowed for motor rehearsal was the same as that allowed for counting backwards. Subjects in the conditions involving concurrent and delayed visual monitoring were told that their motor rehearsals would not be shown on the video monitor. The delayed monitoring involved a 100 s interval, with the enactment being shown during the last 25 s.

To control for the fact that delaying monitoring necessarily delays the next modeled presentation, the time interval between the reproduction attempt following each pair of demonstrations and the presentation of the next pair of demonstrations was held constant at 100 s for all conditions. During this interval, subjects in the delayed visual monitoring conditions either counted backwards or engaged in motor rehearsal for 75 s, depending on the rehearsal condition to which they were assigned. They then viewed a replay of their enactment of the modeled behavior. The mean duration for replays was 25 s in a pilot study \( (n = 18) \), and 26.4 s in the main study, which do not differ significantly.

After Trials 3, 6, and 9, subjects were administered the test of cognitive representation. They were shown nine photographs that depicted, in a scrambled order, the various components of the action pattern. They were instructed to arrange the photographs from left to right in the order that accurately reflected the sequence of component responses exhibited by the model. The scrambled orders, which were different for each of the three tests, were 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 one point for any two response components correctly sequenced. The maximum score was 8.

Extensive preliminary research on different ways of measuring cognitive representation included a recognition measure in which subjects had to choose between photographs of each of the nine response components of the movement pattern and highly similar distractors. A multi-
ple regression analysis revealed that recognition scores did not significantly contribute to the percentage of variance explained in accuracy of subjects' enactments beyond that provided by the markedly significant effect of the picture sequencing measure. The recognition measure was, therefore, not included in the present study for this reason and because of the possibility that exposure to many incorrect component responses may create extensive interference effects in the process of measurement itself.

After completing the test for cognitive representation on Trial 6, all subjects were told they would no longer see the modeled action pattern, but they should continue to perform it. Subjects who had been performing with concurrent or delayed visual monitoring were also informed that they would no longer be able to observe their reproductions. Thus, the final block of three trials was devoid of both modeling and visual monitoring.

Experimental Design

The effects of timing of visual monitoring (concurrent, delayed, or none) and motor rehearsal (presence or absence) were analyzed by a $3 \times 2$ mixed-design ANOVA; trial blocks (blocks of three trials) constituted the within-subjects factor. Prior to performing an ANOVA on each of the dependent variables, a Multivariate Analysis of Variance (MANOVA) was used as an omnibus test of significance.

Reproduction Score

Each response component and the 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 paused in their performance of the action sequence. Subjects were awarded two 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 for each of the nine trials.

Reproductions of a sample of pilot subjects ($n = 18$) were independently rated by two judges to ensure proficiency in using the scoring criteria. To increase interscorer reliability, photographs of each response component were present while subjects' reproductions were being scored. These same judges then independently rated the reproductions 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. Reliability, based on the total sum of scores achieved on the three trial blocks, was $r = .96$. The reliability coefficients computed for each trial block were $r = .95$; $r = .97$; $r = .93$, respectively.
Results

Since the gender factor did not approach significance on any of the response measures, data for males and females were pooled. Figure 1 presents the mean reproduction accuracy as a function of timing of visual monitoring and rehearsal across blocks of trials.

The multivariate analysis (MANOVA) yielded significant main effects for timing of visual monitoring, $F(4, 106) = 12.07, p < .001$, rehearsal, $F(2, 53) = 9.55, p < .001$, and for trial blocks, $F(4, 214) = 53.69, p < .001$. The Monitoring $\times$ Blocks interaction, $F(8, 214) = 2.57, p < .025$, and the Rehearsal $\times$ Blocks interaction, $F(4, 214) = 3.05, p < .05$, were also significant. Univariate ANOVAs were then performed on each of the dependent variables.

Reproduction Accuracy

Results of analysis of variance revealed significant main effects for timing of visual monitoring, $F(2, 54) = 26.86, p < .001$, and for rehearsal,
$F(1, \ 54) = 18.58, \ p < .001$. Reproduction accuracy improved significantly over trial blocks, $F(2, \ 108) = 133.17, \ p < .001$, which primarily reflected the substantial improvement in the second block of trials. The interactions between Monitoring $\times$ Blocks, $F(4, \ 108) = 3.05, \ p < .05$, and Rehearsal $\times$ Blocks, $F(2, \ 108) = 3.40, \ p < .05$, were also significant. These interactions are shown in Figure 2.

Planned comparisons using the Dunn-Bonferroni $t$ test (Kirk, 1968), revealed that the reproductions of subjects performing with concurrent visual monitoring were superior to those performing with delayed, or without any visual monitoring on each of the trial blocks ($p < .01$). However, the benefits of concurrent visual monitoring were larger in the second than in the first block of trials. The reproductions of subjects for whom visual monitoring was delayed were no better than those who had no opportunity to observe their enactments.

Analysis of the simple effects of the Rehearsal $\times$ Blocks interaction revealed that subjects who engaged in motor rehearsal produced significantly superior reproductions on all trial blocks ($p < .01$) as compared

![Graphs showing mean reproduction accuracy scores](image)

**Fig. 2**—The right panel shows the mean reproduction accuracy scores as a function of timing of visual monitoring across blocks of trials. The left panel shows the mean reproduction accuracy scores as a function of motor rehearsal across blocks of trials. $T =$ training trials; $W =$ trials in which modeling and visual monitoring of reproductions are completely withdrawn.
to those who did not. The disparity was larger in the second than in the first block of trials (Figure 2).

As shown in Figure 2, subjects in all conditions maintained their reproduction accuracy on the last block of trials, even though they performed without the aid of the modeling or visual monitoring. There is no significant decline in reproduction accuracy.

**Cognitive Representation**

The mean scores for cognitive representation at each of the three tests were 4.47, 6.85, and 6.68, respectively. The increasing ability of subjects to identify the correct sequential patterning of the modeled action was highly significant, $F(2, 108) = 49.80$, $p < .001$. Timing of visual monitoring had no effect on the development of the cognitive representation, but motor rehearsal enhanced it, $F(1, 54) = 6.65$; $p < .025$. Subjects who engaged in motor rehearsal formed a more accurate conception of the modeled pattern than those who did not.

In order to determine the relationship between conception and action, level of cognitive representation was correlated with the accuracy of reproduction of the action pattern. The better the conception of the modeled pattern, the greater was the reproduction accuracy, $r = .42$; $p < .001$.

**Discussion**

The findings of the present experiment corroborate the influential role played by timing of visual monitoring and rehearsal in mastering a novel action pattern that normally lies outside the visual field. Concurrent visual monitoring of one's enactments greatly enhances observational learning. The ineffectiveness of delayed visual monitoring is consistent with studies of other motor activities (Henderson, 1977). Additionally, Noble and Noble (1958) found that delaying informative visual feedback until the end of a series of component responses produced an inferior performance when compared to presenting such feedback after the completion of each component. As discussed earlier, delayed monitoring of enactments makes it difficult to detect the discrepancies between conception and action necessary to institute corrective adjustments. Delayed visual monitoring may also increase the difficulty of integrating visual and kinesthetic sources of information (Connolly & Jones, 1970; Henderson, 1977).

As hypothesized, motor rehearsal substantially improved reproduction accuracy. Bandura and Jeffery (1973) also found that motor rehearsal enhanced observational learning when subjects were not instructed to encode the modeled activity, as in the present experiment. In contrast, they found that providing a symbolic code raised the level of observational learning substantially so that motor rehearsal produced no further benefits. The code not only represented critical features of the action pattern, but it also specified the operations for translating symbols to actions. Since coding provided a precise representation of the activity, there was less need for information conveyed by motor rehearsal. In the present
study, subjects had to generate their own symbolic guides for reproduction, and motor rehearsal contributed to the development of a cognitive representation.

Repeated exposure to the modeled activity increased the accuracy of the cognitive representation, thus providing a standard for error detection and correction. The accuracy of reproductions then improved. Previous research showed a similar correspondence between representation and reproduction (Carroll & Bandura, 1982). The obtained positive correlation between accuracy of cognitive representation and adequateness of behavioral reproduction is in accord with a conception-matching process. Results of studies in which representational mediators are manipulated experimentally further attest to the causal role of cognitive representation in complex performances (Bandura & Jeffery, 1973; Bandura, Jeffery, & Bachicha, 1974; Gerst, 1971).

It is clear from both the present and the previous study (Carroll & Bandura, 1982) that visual monitoring does not affect the development of cognitive representation. Adams, Gopher, and Lintern (1977) have noted that visual feedback is particularly suited to specifying the details of movement production. Bernstein (1967) has argued that movements are determined by abstract structures that represent the movements in terms of topological or qualitative features. These abstract structures are then differentiated into more precise or quantitative information about the movement pattern. In a similar vein, Turvey (1977) states "... that the detailed contents of vision must be interjected into the act during its evolution" (p. 257). To continue this line of reasoning, it may be that the information conveyed by visual feedback is represented at a lower level of abstraction than the cognitive representation, and it is, therefore, not directly assessed by the measure of representation employed in the present study. One possibility is that visual information is used to construct a system of spatial coordinates that helps to guide overt performance (Lashley, 1951; Russell, 1976; Stelmach & Larish, 1980).

An alternative explanation for why visual monitoring of enactments has little effect on representational development is in terms of attentional processes. Visual monitoring is helpful in correcting erroneous enactments by comparing them with what one already knows, but it does not necessarily highlight what one does not know. In contrast, repeated rehearsal involves constructing and sequencing behavior patterns. Having to construct behavior patterns makes salient what one does not know. Such awareness can heighten attentiveness to subsequent modeling aimed at mastering problematic features of the representation. Indeed, motor rehearsal facilitated representational development.

After subjects achieved some cognitive representation of the modeled activity, they were able to maintain their level of performance without visual or external modeling aids. In studies of simple movements, such as linear positioning, removal of visual feedback usually impairs performance (Adams, Gopher, & Lintern, 1977). However, observational learning encourages the transformation of visual performance information into symbolic codes that later guide response production. Once
people know what to do, have gained proficiency in translating it into corresponding action through visually monitored enactments, and have routinized it through repeated execution, they no longer need modeling guides or visual aids.

A question that has received relatively little attention in the literature concerns similarities and differences between processes governing learning motor skills by observation and by repeated performance. Viewed from the social learning perspective (Bandura, 1977; 1984), learning involves a similar conception-matching process regardless of whether the conception is derived from observing the structure of the behavior being modeled or constructing it from observing the effects of one’s own actions. However, the rate of learning from observing skilled models and from observing performance feedback differs substantially. The structure of complex behavior can be symbolically constructed more readily from observing the behavior displayed in an already integrated form than in trying to construct it gradually from examining the effects of one’s exploratory actions (Bandura, 1984). Indeed, observers often learn faster even from trial-and-error performers than the performers do themselves, especially if the activities require extensive cognitive processing, as do complex performances (Berger, 1961; Hillix & Marx, 1960; Rosenbaum & Hewitt, 1966). Similarly, Klein and Posner (1974) found that observing a complex tracking pattern produced better performance than did physically tracking it when the action had to be reproduced later from memory. It is easier for observers to construct a cognitive representation of complex action patterns when unhampered by the attentional demands of simultaneously planning and executing actions and interpreting their effects.

It would, thus, seem that for activities requiring extensive cognitive processing, such as those involving both acquisition and sequencing of intricate response components, modeling is better suited for conveying the structure of the behavior than relying exclusively on performance. However, physical enactment provides the vehicle for translating conception to skilled action. The issue of the relationship between cognition and action can be fruitfully studied in terms of a conception-matching process. The observational learning paradigm provides a means for clarifying the interaction between the two.

REFERENCES


*Submitted May 23, 1984*

*Revision submitted December 12, 1984*