Journal of Comparative Psychology
1995, Vol. 109, No. 2, 107–114

0735-7036/95/$3.00

Seeing and Knowing: Knowledge Attribution Versus Stimulus Control in Adult Humans (*Homo sapiens*)

Joseph L. Gagliardi, Kim K. Kirkpatrick-Steger, Jennifer Thomas, Gregory J. Allen, and Mark S. Blumberg
University of Iowa

Interest in cognition in nonhuman animals has inspired new approaches to discovering animals' ability to attribute knowledge to others (e.g., D. J. Povinelli, K. E. Nelson, & S. T. Boysen, 1990). The assumptions of such experiments were tested in this study by training a group of humans (*Homo sapiens*) to use accurate information provided by a confederate who was watching as 1 container among 4 was baited; a 2nd group was similarly trained to use accurate information provided by a confederate whose back was turned during baiting. On a single reversal trial, the roles of the 2 confederates were switched. Subjects were able to learn their respective tasks but attended to different aspects of the confederates, as revealed by the reversal trial. Although attributional interpretations can be applied to such data, many of the choices in this experiment can be explained more readily with the basic principles of contingency-based learning.

We are currently witnessing an explosion of interest in the desires, beliefs, and attributions of humans and how such mentalistic processes develop and come to affect human behavior (Fodor, 1992; Gopnik & Astington, 1988; Wellman, 1990). Although such research topics have been viewed in a largely unfavorable light within psychology during most of this century, they are now a focus of great excitement among philosophers and experimentalists alike. This excitement is not, however, limited to the mental life of humans; on the contrary, interest in the desires, beliefs, and attributions of nonhuman animals is flourishing as well (Cheney & Seyfarth, 1990; Griffin, 1992; Povinelli, 1993).

In a recent critique of experimental efforts to determine whether animals are capable of attributing knowledge to others, Heyes (1993) evaluated the usefulness of four different methods for determining mental state attribution in animals: collecting anecdotes, conditional discrimination training, trapping, and triangulating. The first three methods were dismissed by Heyes because they fail to differentiate between observed behaviors and their purported underlying mental states. Triangulation, however, was proposed to hold the greatest promise of success for demonstrating mental state attribution in animals.

Triangulation consists of conditional discrimination training followed by a transfer test that is designed to reveal the underlying factors that guided the original learning. This method was used by Povinelli, Nelson, and Boysen (1990) to examine mental state attribution in chimpanzees (*Pan troglodytes*). In their procedure, a chimpanzee was placed in a room with two human confederates and four containers. At the beginning of the experiment, a barrier was raised to prevent the chimp from seeing the containers. At that time, one of the confederates (the guesser) left the room while the other confederate (the knower) baited one of the containers with food. The guesser then returned to the room, the barrier was removed, and each confederate pointed to one of the containers; the chimp then made a selection and received food if it chose the baited container. The two confederates randomly alternated between roles as guesser and knower on each trial, so that the chimps learned to discriminate the behavior of the confederates, not their identity. The knower always pointed to the baited container, whereas the guesser always pointed to one of the three unbaited containers. After extensive training, 3 of the 4 chimps had learned to choose the correct container, that is, the one indicated by the knower.

In the critical transfer phase of the experiment, three features were changed: (a) Both confederates now stayed in the room while the containers were baited; (b) a third confederate, rather than the knower, now baited the containers; and (c) the guesser placed a bag over his head during the baiting process. It was reasoned that if the chimps were attributing knowledge to the knower during the training phase, then they ought to continue to perform well in the transfer phase despite the changes in the discriminative aspects of the confederates. In fact, the chimpanzees did perform well, which indicated successful transfer. Povinelli et al. (1990) then argued that because the discriminative stimuli that were correlated with knowing during training were significantly altered in the transfer phase, mental state attribution is left as the most probable interpretation.

Joseph L. Gagliardi, Kim K. Kirkpatrick-Steger, Jennifer Thomas, Gregory J. Allen, and Mark S. Blumberg, Department of Psychology, University of Iowa.

Preparation of this article was supported in part by National Institute of Mental Health Grant MH50701.

We thank Jie Xu, Edison Perdomo, and David Wiseman for their help in the running of subjects and Linda Van Hamme, Susan Schalk, and Yang-Ho Moon for technical assistance. Lisa Oakes and Ed Wasserman generously provided helpful criticisms and comments at various stages of the project.

Correspondence concerning this article should be addressed to Mark S. Blumberg, Department of Psychology, University of Iowa, Iowa City, Iowa, 52242.

Heyes (1993) saw significant potential in the triangulation method to unveil attributional processes in animals:

> In our human theory of mind, the link between seeing and believing is so strong that if [Povinelli and his colleagues'] interpretation were fully supported then I would regard the experiment as providing convincing evidence of mental state attribution in chimpanzees. (p. 183)

As Heyes pointed out, however, the power of the triangulation method relies on immediate transfer. In Povinelli et al.'s (1990) experiment, however, they only reported the chimps' mean performance on 30 transfer trials (10 trials each on 3 consecutive days) on which the subjects received differential reinforcement for choosing the knower. If the chimpanzees were using mental state attribution to gain reward, then performance during the transfer phase ought to have remained above chance, even on the first trial. Beyond the first trial there was the potential for further learning to occur, particularly because of the differential reinforcement procedures, and it was therefore impossible to determine whether successful performance was due to retraining or to true transfer from the original discrimination.

Moreover, even if the presentation of data from the first trial of the transfer phase were to reveal positive transfer, another problem exists. Specifically, although the transfer phase involved changes in various aspects of the experimental procedure, there were still observable cues that remained invariant and that continued to correlate with knowledge of the baited container. For example, although the knower was no longer baiting the cups, the knower's head continued to face the containers during the baiting process. The physical attributes of the knower were not altered; only the attributes of the guesser were substantially changed. Thus, even if single-trial transfer were successful, the methodological changes implemented during the transfer phase would not be sufficient to rule out a simple discrimination process.

Finally, the paradigm suffered from one other weakness. Namely, although it seems clear that chimps can learn to point to a baited container when the knower correctly identifies it, it has not been demonstrated whether chimps are similarly capable of learning to point to a baited container when the guesser (i.e., the confederate who is out of the room during the baiting process) correctly identifies it. If "seeing is believing" is as fundamental to human behavior as is commonly assumed, then it must be very difficult for a subject to learn to rely on a confederate who does not observe the baiting process. Successful learning of such a falsification task will cast doubt on some critical assumptions underlying the experimental search for evidence of mental state attribution.

In the present study, we ask, "If a confederate reliably provides correct information to a subject, what attributes of the confederate (e.g., individual identity, behavior, knowledge) are most important in controlling the subjects' choices?" To answer this question we used a paradigm of conditional discrimination training followed by a single reversal trial. The subjects were divided into two groups and were trained and tested with two different versions of video stimuli. The stimuli for both groups were identical in that they consisted of multiple depictions of two confederates and a baiter. In the facing group, subjects were trained to rely on truthful information provided by the confederate who was facing a row of containers as one of the containers was baited. The subject's view of these containers was obscured during the baiting process. The second confederate was also in sight of the subject, but her back was turned as the container was baited. The physical attributes of the confederate (i.e., her identity) and the direction she was facing during the baiting process (i.e., her orientation) were perfectly correlated throughout the training trials.[1] After subjects had learned this task, they were tested on a single reversal trial in which the confederate who had been observing the baiting process was now turned around and the confederate who had been turned around during the baiting process was now facing forward. Because the facing condition entailed a situation that is congruent with our real-world experiences about the connection between seeing and knowing, it was expected that subjects would find this task very easy to learn.

The subjects in the turned group were exposed to the reverse situation. That is, subjects were trained to rely on truthful information provided by the confederate who was turned away from the containers during the baiting process. All other aspects of the experiment were the same as for the facing group. The reversal trial consisted of the previously turned confederate now facing the baiting process and the previously facing confederate now turned away from the baiting process. It was expected that this second task would be more difficult for subjects to learn because of its novelty and ambiguity.

Thus, the aim of this experiment was to (a) determine the properties of confederates that guide subjects' behavior, (b) determine if subjects behave differently when reliable information is provided by confederates whose apparent source of information is not direct visual observation, and (c) eliminate additional learning during the transfer phase that might preclude the measurement of the subjects' initial response tendencies.

Method

Subjects

The subjects were 37 undergraduate students (*Homo sapiens*) enrolled in introductory psychology courses, 18 men and 19 women, aged 17 to 25 (Mdn = 19). All subjects were informed that they would receive monetary compensation for their involvement in the study. Also, the subjects were randomly assigned to view the two video conditions.

[1] This perfect correlation between identity and orientation among the confederates differs from the methodology of Povinelli, Nelson, and Boysen (1990), in which the identities of the guesser and knower changed from trial to trial. Our methodology made possible the use of a reversal trial on which the previously correlated stimulus dimensions of the confederates were uncoupled to reveal the controlling stimulus dimension.

Apparatus

The subjects were tested in a 10 × 15 ft (3.0 × 4.6 m) room. A table and two chairs were situated along one wall of the room. On the table was a 19-in. (48.2-cm) color monitor and a VCR for showing the experimental video. Four 2 × 7 in. (5.1 × 17.8 cm) cards were placed in a row in front of the monitor; each card was labeled with a color and a number (yellow–1, blue–2, green–3, and red–4). To the right of the cards sat a cup in which a quarter was deposited by the experimenter after each correct choice.

Video Stimuli

Two videos were prepared, each with 12 training trials and 1 reversal trial. Each trial began with a fade-in that revealed two confederates. The confederates stood approximately 1 m apart and were separated by a wooden divider that blocked their view of one another. Both confederates were women of similar physical build and age and wore distinguishably different clothing. One confederate (the facing confederate) was facing the camera, whereas the other confederate (the turned confederate) was turned away from the camera (Figure 1, Panel A).

Approximately 1 m in front of the confederates stood a table on which four cylindrical containers (8 in. [20.3 cm] high and 4.5 in. [11.4 cm] in diameter) were placed in a row. Each container was identifiable by two attributes, a color and a number, corresponding to the cards (i.e., yellow–1, blue–2, green–3, and red–4). After 2–3 s, a third confederate (the baiter) entered the camera's view from the right (from the subject's perspective), stood in front of the table with her back to the camera, and proceeded to bait one of the containers. During baiting, the baiter's body served to block the camera's view of the containers, but both confederates could still be seen in the background over the baiter's shoulders (Figure 1, Panel B). The baiter lifted and placed her hand under each container in succession from left to right, placing a small black object under one randomly chosen container. The baiter tried to make the same baiting movements toward all four containers to minimize cues about the placement of the object. The facing confederate oriented toward the containers with head and eye movements, whereas the turned confederate continued to face away from the containers. The baiter then exited to the right, out of the camera's view.

After the baiter exited, the turned confederate turned around to face the containers. Both confederates then shuffled a set of four 7 × 9 in. (17.8 × 22.9 cm) cards that they had been holding; each card corresponded in color and number with one of the containers. At the end of a 3-s period, each confederate held up one of the four cards (Figure 1, Panel C). After 5 s, the baiter returned, walked behind the table, and lifted the baited container (in view of the camera and the two confederates) to reveal the small black object (Figure 1, Panels D1 and D2). The camera then faded out for a 5-s intertrial interval. The left–right position of the two confederates was randomized.

Two different videos were made, one for each experimental condition. Both videos were identically scripted with regard to the baiting of the containers and the confederates' selections. The predetermined baiting and selection sequences occurred in a random order; the only deviation from randomness was in making sure that the two confederates made different choices from each other on the last two training trials (as is explained later). The two videos differed only with respect to which confederate provided correct information to the subject. In one video (the facing condition), the facing confederate always provided the correct information as to which container was baited (Figure 1, Panel D1); in the second video (the turned condition), the turned confederate was always correct (Figure 1, Panel D2). On all training trials and for both videos, the facing and turned confederates were played by the same persons.

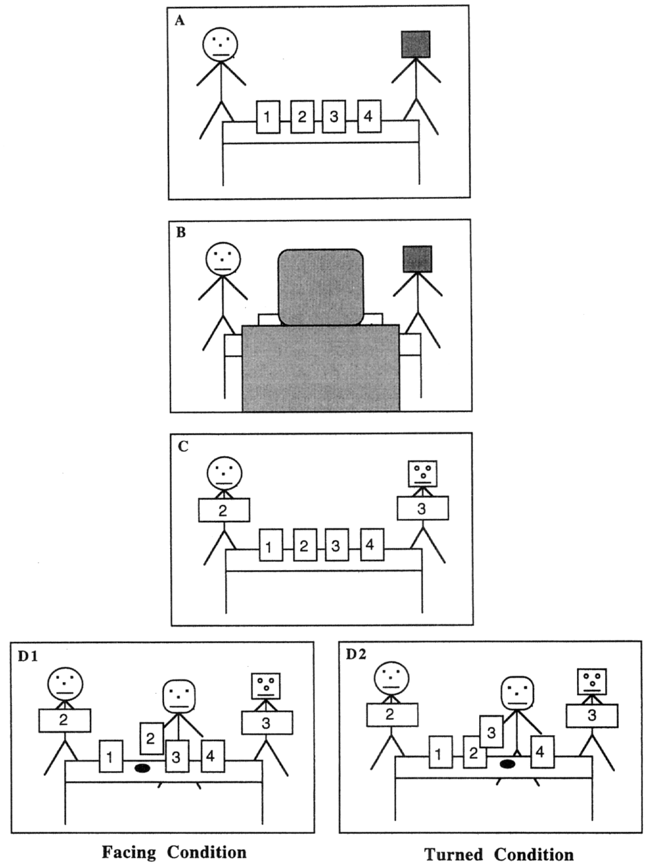

Figure 1. Schematic depiction of a sample video trial for each experimental condition. A: Two confederates are in view: One faces the containers on the table, and one is turned away. B: The baiter enters the scene and baits one of the containers while blocking the subject's view. C: After the baiter leaves the scene, the turned confederate turns around, and each confederate selects a card with a number on it corresponding to one of the four containers. D1 and D2: The baiter reenters the scene and, depending on the experimental condition (i.e., facing or turned), lifts one of the containers (revealing the baited object) corresponding to the selection made by the facing confederate (D1) or the turned confederate (D2). On the reversal trial (not shown), the facing confederate initially turns away from the containers and the turned confederate faces forward.

Because a true guesser would be expected on occasion to provide correct information by chance (specifically, with four containers, a true guesser ought to be correct on one of every four trials), the guessers on each video were correct on 25% of the trials. Thus, in the video for the facing condition, the facing confederate was correct on 100% of the trials, and the turned confederate was correct on 25% of the trials, whereas in the video for the turned condition, the turned confederate was correct on 100% of the trials, and the facing confederate was correct on 25% of the trials. As a result of these probabilities, both confederates chose the same container on 25% of the training trials.

The final, 13th trial of the video was a reversal trial and was procedurally identical to the training trials except that the roles of the confederates were reversed. That is, the facing confederate was now turned away (designated the *turned*$_R$ confederate; the subscript *R* refers to the reversal trial); similarly, the turned confederate was now facing forward (designated the *facing*$_R$ confederate). Thus, the facing confederate is the same person as the turned$_R$ confederate, and the turned confederate is the same person as the facing$_R$ confederate.

As during training, the baiter entered from the right and appeared to bait a container. However, none of the containers was actually baited. The baiter then left the camera's view, the turned$_R$ confederate turned around to face the containers, and each confederate displayed a card as on previous trials. Unlike the training trials, the reversal trial ended without revealing which container was baited.

Subject Testing

Each subject was escorted into the experimental room and was seated in front of the video monitor; an experimenter sat in an adjacent chair. The subject was asked to read the following set of instructions:

> The video screen in front of you will portray a series of events. In the foreground you will see four containers, the colors of which are represented by the cards in front of you. A person will enter the room and place an object under one of the containers. Shortly thereafter, you will be given the opportunity to choose one of the cards. The correct choice will appear on the monitor. If your choice is correct, you will receive a token which can be redeemed for money at the end of the experiment. The maximum amount of money that can be earned over the series of trials is \$3. All questions should be withheld until the completion of the experiment. If, for any reason, you wish to discontinue your participation in the experiment, you are free to leave at any time.

After the subject finished reading the instructions, the experimenter started one of the two videos, to begin the experiment. The subject watched the video on the television monitor until the confederates displayed their cards on the first trial. At this time, the experimenter paused the video and prompted the subject to select which container was baited. The subject could choose by either pointing at one of the corresponding cards or by verbally indicating either the color or number of one of the cards on the table in front of him or her. The experimenter recorded the subject's choice and then restarted the video to reveal which container held the small object. If the subject was correct, the experimenter dropped a quarter into the cup. If the subject was incorrect, no reinforcement was received. This procedure was followed for each of the 11 remaining training trials.

The subjects were not informed that the reversal trial was different from the previous 12 training trials. The trial proceeded in the same manner as the training trials. When the two confederates held up their choice card, the video was paused with the two confederates indicating their selections. The subjects were prompted for a choice as on the training trials. Their choice was recorded, and then subjects were asked, "Why did you make that choice?" The subjects gave verbal answers to this question, and the experimenter recorded their answers on a data sheet. The subjects were not reinforced on this final trial.

At the end of the experiment, the subjects were instructed to refrain from discussing the experiment with classmates. They were then allowed to take the quarters accumulated during the experiment (range, \$1.50–\$3.00).

Results

Figure 2 presents the learning curves for subjects in the two experimental groups. Both functions increased monotonically and plateaued by Trial 5.

As we discuss earlier, in order to simulate a situation in which one confederate was a knower and the other was a guesser, it was necessary that both confederates indicate the same container on 25% of the training trials. Such identical selections were made by confederates on Trials 4, 6, and 10 and are indicated by the arrows in Figure 2. These trials, however, were excluded from analyses because when both confederates indicated the same container on these trials, we could not be sure which confederate (if not both) influenced a subject's choice.

On the first trial, only 63% (12 of 19) of the facing subjects and 56% (10 of 18) of the turned subjects selected either of the confederates' choices. Thus, many subjects selected from the two containers that were not indicated by the confederates. In contrast, on the second trial, nearly all (36 of 37) subjects now selected from the confederates' choices. With small deviations, this remained true for the remainder of the training trials. Moreover, as expected, the choices of subjects in both groups were evenly distributed between the two confederates on Trial 1 but became increasingly differentiated as training progressed.

A repeated measures analysis of variance (ANOVA) was used to test the effects of experimental group and trial number on the subjects' learning of the task. To do this, each subject's choice on each trial was coded in the following way: If a subject chose correctly on a given trial (and thus was reinforced), that trial was coded as a 1; if a subject chose incorrectly on a given trial (and thus was not rein-

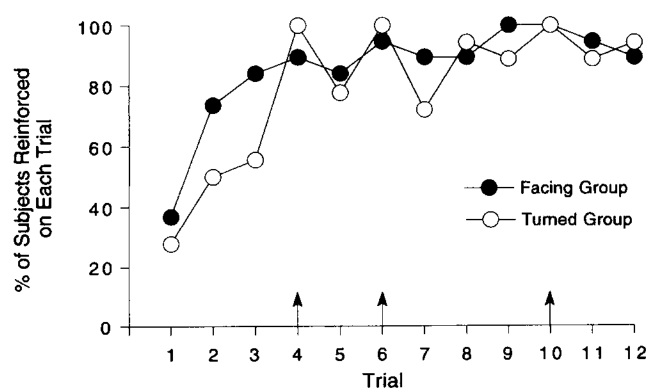

Figure 2. The percentage of subjects who received reinforcement on the 12 training trials for the two experimental groups. In the facing group, subjects were reinforced when they chose the container indicated by the facing confederate. In the turned group, the subjects were reinforced when they chose the container indicated by the turned confederate. Arrows indicate trials in which both confederates selected the same container.

forced), that trial was coded as a 0. Therefore, subjects' choices in the facing group were coded as 1s on those trials when the facing confederate's selection was chosen, and subjects' choices in the turned group were coded as 1s on those trials when the turned confederate's selection was chosen. Zeros were assigned on all other trials. The ANOVA was then computed with the 1s and 0s as parametric values.

There were significant main effects of group, $F(1, 35) = 5.76$, $p < .05$, and trial, $F(8, 280) = 11.56$, $p < .0001$, but the Group × Trial interaction was not significant, $F(8, 280) = 0.81$, $p > .5$. Thus, across all training trials, more subjects in the facing group ($M \pm SE_M = 83.0\% \pm 3.0\%$) made correct choices than did subjects in the turned group ($M \pm SE_M = 72.2\% \pm 3.5\%$), and subjects showed significant improvement throughout the training trials. Nonetheless, there were no differences in rate of learning between the groups, as evidenced by the nonsignificant interaction.

Table 1 presents the data for the reversal trials. The data indicate a significant difference between the facing and turned groups, $\chi^2(1, N = 37) = 6.06$, $p < .05$. Specifically, 74% (14 of 19) of the subjects who were trained to choose the facing confederate chose the facing$_R$ confederate on the reversal trial; in other words, these subjects appear to have made their choices on the basis of the orientation of the confederate on the reversal trial, not her identity. Interestingly, 67% (12 of 18) of the subjects trained to choose the turned confederate also chose the facing$_R$ confederate on the reversal trial, which suggests that these subjects made their choices on the basis of the identity of the confederate, not her orientation.

A handful of subjects may not have learned the task very well by the end of the training trials. To determine whether the data from these subjects influenced the effects observed on the reversal trial, similar analyses were conducted without these subjects. Therefore, in Table 1, we also present reversal trial data for only those subjects who made correct choices (and were thus reinforced) on the last 2 training trials. Restriction of the data to these high-performing subjects still revealed significant differences between the facing and turned subjects, $\chi^2(1, N = 31) = 3.90$, $p < .05$.

Table 1
Contingency Table for the Subjects' Choices for the Reversal Trial

	Training trials					
	Facing			Turned		
Reversal trial	Choice	Total	End	Choice	Total	End
Orientation	Facing$_R$	14	12	Turned$_R$	6	6
Identity	Turned$_R$	5	4	Facing$_R$	12	9

Note. For training trials, the headings indicate the confederate the subject was reinforced for relying on. For reversal trials, the headings indicate the confederate's attribute that the subject relied on. *End* indicates the subjects who on Trials 11 and 12 chose correctly (and were reinforced). *Facing* $_R$ indicates the choice of the confederate who was facing the baiting process during the reversal trial; *turned*$_R$ indicates the choice of the confederate who was turned away from the baiting process during the reversal trial.

Discussion

This experiment demonstrates that adult human subjects can learn to rely on confederates for information regardless of how the confederates appear to have come to possess that information. Moreover, there were no significant differences in the rates at which the subjects learned the two tasks, a finding that is surprising given the oddity of the turned condition and its incongruence with our typical real-world experiences. Finally, the subjects in both groups tended to rely on the confederate who faced the baiting process on the reversal trial (i.e., the facing$_R$ confederate).

Because these data are amenable to interpretations based on either attributional or stimulus control perspectives, they cannot definitively distinguish between them. Nonetheless, at the least, an attributional perspective demands that the subjects in the facing group overwhelmingly choose the facing$_R$ confederate on the reversal trial because that confederate now had access to the knowledge of which container had been baited. In fact, as shown in Table 1, 74% (14 of 19) of the subjects on the reversal trial did rely on the facing$_R$ confederate. On the other hand, 26% (5 of 19) of the subjects relied on the turned$_R$ confederate, a rather high percentage given the ease of the task and the age of the subjects.

The investigation of knowledge attribution typically relies on a strong link between seeing and knowing. This link helps to explain the difficulty in providing explanations for the behavior of subjects in the turned group. Indeed, it is clear that the attributional perspective is mute with respect to these subjects. How can it explain, for example, that subjects quickly learned to rely on a confederate who was turned away as the containers were being baited? Furthermore, how can the attributional perspective explain that 33% (6 of 18) of the subjects in the turned group relied on the turned$_R$ confederate and, by doing so, shifted from a confederate who, on the reversal trial, was facing the baiting process (and who also was correct on all previous trials) to a confederate who was turned away from the baiting process?

If the subjects were attributing knowledge to the confederates on the reversal trial and these attributions were guiding their choices, then such processes may be revealed by their verbal explanations; these explanations are presented in Table 2. We stress that these statements were placed in their respective cells based on the subject's choice on the reversal trial, not on the nature of the explanations. In response to the question, "Why did you make that choice?," subjects who were trained to rely on the facing confederate and who on the reversal trial relied on the facing$_R$ confederate stated that they relied on the confederate who was "facing forward," "looking," "watching," or "facing" (Table 2, top left cell). Similarly, subjects who were trained to rely on the turned confederate and who on the reversal trial relied on the turned$_R$ confederate stated they they relied on the confederate who was "turned around," "facing the wall," or "not looking" (Table 2, top right cell). Thus, when these subjects explained their choices on the reversal trial, they provided analogous explanations, which is consistent with

Table 2
Comments (Reconstructed From Written Notes) Elicited From Subjects After They Made Their Choice on the Final, Reversal Trial

	Training trials	
Reversal trial	Facing	Turned
Orientation	**$Facing_R$** She was facing forward. I also used cues from the baiter. (2) She was looking. (2) She was facing the cans. She saw the baiter place the object. She was the one who could have seen. She was watching the baiter. She was facing—the one who is facing is always right. She was the one watching—the one who is watching is always right. I went with the person who was looking. She was facing the boxes and observing the baiting process. I began to trust the person looking at the containers. I also tried to watch the baiter's arm. I used cues from the baiter.	**$Turned_R$** She was turned around—the person who is turned around is always right. (2) She was turned around—the one who is not looking is always correct. The one whose back is turned is right. She was facing the wall—there was a pattern. The person who was turned around is correct.
Identity	**$Turned_R$** She was right all the time. She should be right this time even though she was turned around. She was correct each time. I followed who the person was instead of what she did. She was right all the time. I watched the baiter. I was choosing the confederate who was facing forward through most of the experiment but went against the odds on the final trial on a hunch.	**$Facing_R$** I chose her because she got the majority right. I chose her because she's been right so far. She had them all right before. I chose her because she always had the right answer. She always had the right answer. She was told the right answer. It wasn't ESP. She was the one who got them all right. She was correct the rest of the time. I used cues from the baiter. (2) She has ESP so she is always right. The baiter's arm moved faster when she baited the container. She blinked after the baiter lifted the correct container.

Note. For training trials, the headings indicate the confederate the subject was reinforced for relying on. For reversal trial, the headings indicate the confederate's attribute that the subject relied on. $Facing_R$ indicates the choice of the confederate who was facing the baiting process during the reversal trial; $turned_R$ indicates the choice of the confederate who was turned away from the baiting process during the reversal trial. Numbers in parentheses indicate more than one occurrence.

their choices being guided by the orientation of the confederate, not her identity, regardless of whether the chosen confederate was looking or not.

In contrast, responses in the bottom row of Table 2 were consistent with the interpretation that the subjects' behavior *was guided by the identity* of the confederate, not her orientation. Thus, subjects who were trained to rely on the facing confederate and who on the reversal trial relied on the $turned_R$ confederate stated that they relied on the confederate who was "right all the time" or "correct each time" (Table 2, bottom left cell). Similarly, subjects who were trained to rely on the turned confederate and who on the reversal trial relied on the $facing_R$ confederate, stated that they relied on that confederate who "got the majority right," "had them all right before," or "always had the right answer" (Table 2, bottom right cell).

As Table 1 indicates, the majority of subjects in both the facing (74%) and turned (67%) groups relied on the facing$_R$ confederate on the reversal trial. Thus, one may argue that these subjects from both groups were particularly sensitive to the confederate who was facing the baiting process on the reversal trial and behaved similarly in relying on that same confederate. On the other hand, the subjects' explanations in Table 2 indicate that subjects made their choices based on the stimulus properties of the confederates, that is, their orientation and identity. For example, as indicated in the top row of Table 2, subjects in the facing group and the turned group relied on the facing$_R$ confederate and the turned$_R$ confederate, respectively. However, despite the subjects' reliance on different confederates on the trial, their responses are clearly analogous to each other. Of course, we must be cautious when interpreting the verbal reports of subjects (Nisbett & Wilson, 1977); nonetheless, the subjects' explanations in Table 2 are consistent with their choice behavior and suggest that the subjects were sensitive to the stimulus features of the confederates. Therefore, it appears that subjects were influenced by their reinforcement histories during the training trials.

As we state earlier, it is not possible to definitively prove or disprove the truthfulness of either the attributional or stimulus control perspectives. From the combined results of the reversal trial data (Table 1) and the corresponding verbal accounts (Table 2), it is evident that the attributional perspective is limited in its ability to account for the present data. Can the stimulus control perspective take us any further? Indeed, the design of our experiment can be viewed as a classic stimulus control study, in that it involves training trials with compound stimuli (individual identity vs. orientation) followed by a reversal trial in which the elements of the compound stimuli are re-paired (facing confederate now turned and turned confederate now facing). Similar experimental designs have been used to investigate stimulus features that control discrimination and attention. For example, Reynolds (1961) trained two pigeons to discriminate two compound stimuli; specifically, the pigeons were reinforced when they pecked a white triangle on a red background and were not reinforced when they pecked a white circle on a green background. When, after training, the elements of these compound stimuli were presented individually but no reinforcement was provided, one of the pigeons pecked predominantly on the red key, but the other pigeon pecked predominantly on the key with the white triangle. In other words, the two pigeons attended to different aspects of the compound stimuli during training. In an extension of Reynolds's findings, Farthing and Hearst (1970) trained pigeons to discriminate two compound stimuli; specifically, a vertical white line on a blue background or a horizontal white line on a green background. When, after training, the elements of these compound stimuli were presented individually, the background colors exerted greater control over the pigeons' behavior than did line orientation, which indicates that color was a more salient dimension than line orientation.

Such experiments provide an empirical path to assessing the stimulus features that guided discrimination learning in this experiment. As with the pigeons in the earlier experiments, our human subjects may have displayed a preference for the facing orientation because such an orientation has been associated with reliable information-giving in the past. This appeal to prior reinforcement history is, in this context, as untestable as any appeal to cognitive predispositions. Nonetheless, experimental manipulations (e.g., pretraining) can be used to alter the relative weight of these predispositions and thus test the hypothesis that prior learning affects performance on these kinds of tasks. In fact, Farthing and Hearst (1970) demonstrated that pretraining with line orientation alone before compound stimulus training increased the ability of this less salient element of the compound stimulus to control behavior, even to a greater extent than the initially more salient dimension of color.

Although many experimental paradigms exist to study the relative contributions of stimulus features to discrimination learning, the methods for studying knowledge attribution are not as well developed. As we detail in the introduction, Heyes (1993) dismissed three of the four methods used to study mental state attribution in nonhuman animals because they fail to separate the overt behavior from their purported associated mental states. Therefore, the critical question is whether the triangulation procedure overcomes the recognized flaws of the other three methods. Unlike Heyes, we question whether it is theoretically and methodologically possible to design an experiment in which all discriminative stimuli correlated with the state of knowing are expunged from the experiment. Even Povinelli et al. (1990) discussed the possibility that their chimpanzees may have "learned to use discriminative cues ... to solve the problem" (p. 207), but they simply argued that such a possibility was unlikely.[2] They also cautioned that determining the presence of mental state attribution in nonhuman animals requires converging evidence from multiple experiments. Our experiment, however, shows that even adult humans, whose propensity to make attributions ought to be maximal, often respond in a manner inconsistent with an attributional framework. Given this result, we must be especially cautious when interpreting data from similar experiments aimed at unveiling attributional processes in animals.

References

Cheney, D. L., & Seyfarth, R. M. (1990). *How monkeys see the world: Inside the mind of another species.* Chicago: University of Chicago Press.

Farthing, G. W., & Hearst, E. (1970). Attention in the pigeon:

[2] In Povinelli, Nelson, and Boysen's (1990) initial article, they dismissed discrimination learning as an explanation for their results based on the claim that "three of the 4 subjects showed *immediate transfer* [italics added] in the novel test situation" (p. 207). In fact, in a very recent reply to Heyes's (1993) critique, Povinelli (1994) disclosed that the chimpanzees in the earlier experiment actually performed at chance levels for at least the first 2 trials of the transfer phase.

Testing with compounds or elements. *Learning and Motivation, 1,* 65–78.
Fodor, J. A. (1992). A theory of the child's theory of mind. *Cognition, 44,* 283–296.
Gopnik, A., & Astington, J. W. (1988). Children's understanding of representational change and its relation to the understanding of false belief and the appearance-reality distinction. *Child Development, 59,* 26–37.
Griffin, D. R. (1992). *Animal minds.* Chicago: University of Chicago Press.
Heyes, C. M. (1993). Anecdotes, training, trapping and triangulating: Do animals attribute mental states? *Animal Behaviour, 46,* 177–188.
Nisbett, R. E., & Wilson, T. D. (1977). Telling more than we can know: Verbal reports on mental processes. *Psychological Review, 84,* 231–259.
Povinelli, D. J. (1993). Reconstructing the evolution of mind. *American Psychologist, 48,* 493–509.
Povinelli, D. J. (1994). Comparative studies of animal mental state attribution: A reply to Heyes. *Animal Behaviour, 48,* 239–241.
Povinelli, D. J., Nelson, K. E., & Boysen, S. T. (1990). Inferences about guessing and knowing by chimpanzees (*Pan troglodytes*). *Journal of Comparative Psychology, 104,* 203–210.
Reynolds, G. S. (1961). Attention in the pigeon. *Journal of the Experimental Analysis of Behavior, 4,* 203–208.
Wellman, H. M. (1990). *The child's theory of mind.* Cambridge, MA: MIT Press.

Received June 21, 1994
Revision received August 23, 1994
Accepted August 23, 1994 ■