Memory & Cognition
2008, 36 (6), 1079-1086
doi: 10.3758/MC.36.6.1079

Is awareness necessary for true inference?

Peter D. Leo and Anthony J. Greene
University of Wisconsin, Milwaukee, Wisconsin

In transitive inference, participants learn a set of context-dependent discriminations that can be organized into a hierarchy that supports inference. Several studies show that inference occurs with or without task awareness. However, some studies assert that without awareness, performance is attributable to pseudoinference. By this account, inference-like performance is achieved by differential stimulus weighting according to the stimuli's proximity to the end items of the hierarchy. We implement an inference task that cannot be based on differential stimulus weighting. The design itself rules out pseudoinference strategies. Success on the task without evidence of deliberative strategies would therefore suggest that true inference can be achieved implicitly. We found that accurate performance on the inference task was not dependent on explicit awareness. The finding is consistent with a growing body of evidence that indicates that forms of learning and memory supporting inference and flexibility do not necessarily depend on task awareness.

The transitive inference (TI) task requires one to learn context-dependent relations and to make inferences about untrained relations. In a typical TI task, participants are trained on a set of overlapping two-item discriminations or premise pairs (e.g., A > B, B > C, C > D, D > E; "A > B" denotes a forced choice between A and B in which A is the correct choice). An efficient representation of these simultaneous context-dependent relations is A > B > C > D > E. At test and without reinforcement, the capacity for inference is tested with an untrained pair, B?D. The capacity for correct inference has been observed in numerous nonhuman animal species (Davis, 1992; Gillan, 1981; Weaver, Steirn, & Zentall, 1997) and in humans (Greene, Spellman, Dusek, Eichenbaum, & Levy, 2001; Heckers, Zalesak, Weiss, Ditman, & Titone, 2004; Nagode & Pardo, 2002; Preston, Shrager, Dudukovic, & Gabrieli, 2004).

In humans, inference is possible with or without explicit awareness of the contingencies (Frank, O'Reilly, & Curran, 2006; Greene, 2007; Greene, Gross, Elsinger, & Rao, 2006; Greene et al., 2001; van Elzakker, O'Reilly, & Rudy, 2003). If participants are not informed of a hierarchy at the outset, TI task awareness either is independent of task performance (Greene et al., 2006; Greene et al., 2001) or may occur in the absence of task awareness (Frank, Rudy, Levy, & O'Reilly, 2005). That learned contingencies may support inference without awareness is of particular interest because it suggests a level of flexibility previously believed to require conscious memory processes (Clark, Manns, & Squire, 2002; Reber, Knowlton, & Squire, 1996; Smith & Squire, 2005).

One possibility is that implicit versions of TI are not really demonstrating flexibility, but instead are solved using pseudoinference (Frank, Rudy, & O'Reilly, 2003; Siemann & Delius, 1994; van Elzakker et al., 2003; Zentall & Sherburne, 1994). Consider a five-item TI task (e.g., A > B > C > D > E). If the A item is present, it is always the correct choice; likewise, E is always incorrect. The middle items B–D are each correct only 50% of the time. Van Elzakker et al. argued that end items cause a blocking effect that is responsible for differential associative strength among items. Accordingly, the perfect predictive value of the end items causes them to be treated as atomic discriminations (select A, avoid E), and so blocking leads one to ignore B in the AB pairing and to ignore D in the DE pairing. This has the consequence that B is ignored when it would be an incorrect choice (A > B) and is attended to only when it is the correct choice (B > C). Similarly, D is ignored when it would be the correct choice (D > E) and is attended to only when it is the incorrect choice (C > D). According to this view, the novel pairing at test, B?D, is solved only because B has a more positive associative strength than does D, so true inference is not required (Frank et al., 2005; Frank et al., 2003; van Elzakker et al., 2003). Another proposed mechanism is value transfer theory (VTT; Siemann & Delius, 1994; Zentall & Sherburne, 1994), in which item B gains positive strength by repeated association with item A, and likewise item D receives some of E's negative value by repeated association. According to VTT, the selection of B > D at test merely reflects that B has been paired with a positive stimulus (A) and that D has been paired with a negative stimulus (E), so B has more positive associative strength than does D. VTT differs from the blocking approach only in the means by which differential associative strength is imparted to items B and D; both approaches argue that the differential associative values of B and D result in choices that resemble inference (pseudoinference) but are not

P. D. Leo, pdleo@uwm.edu

true inference (syllogism). Pseudoinference approaches would predict that true inference would be possible only with explicit strategies (e.g., Frank et al., 2005; Zentall & Sherburne, 1994). Conversely, true inference approaches would argue that the simultaneous representation of related contingencies can specify inferred relations as well (Howard, Fotedar, Datey, & Hasselmo, 2005), which does not necessarily depend on deliberative strategies (e.g., Ellenbogen, Hu, Payne, Titone, & Walker, 2007; Greene et al., 2006; Greene et al., 2001). Although pseudoinference approaches provide possible alternative explanations for implicit inference performance, they do not provide direct evidence that disproves the true-inference approach.

To assess whether implicit inference can occur without the possibility of pseudoinference, we employed a different implicit inference task in which no end items have an independent predictive value necessary for blocking or for value transfer. The task associations do not contain end pairs with either negative or positive associative value, and therefore the pseudoinference hypotheses do not pertain. The present nonlinear inference task is adapted from a task developed by Bunsey and Eichenbaum (1996) for testing rats with hippocampus lesions. The task is likewise similar to that employed by Preston et al. (2004), except that their task contingencies involved verbal strategies and were learned explicitly.

In the present task, participants learn context-dependent discriminations among faces that can be organized to make inferences about novel configurations. As can be seen in Figure 1A, the stimulus arrays constitute an inverted triangle. In the triangle, the bottom face serves as a discriminative stimulus (DS) that determines the correct choice between the top two items. Figure 1B shows the learned relations: Participants learn for faces A and B that M → A (given M as the DS at the bottom of the triangle, choose face A) and N → B. Then they learn for faces M and N the relations X → M and Y → N.

At test, participants are presented with two novel arrays in which the choice between A and B must be made with X or Y as the DS. Correct inference occurs when participants choose X → A and Y → B (see Figure 1B). Implicit performance on the present task would constitute evidence of inference not attributable to differential associative strength.

EXPERIMENT 1

Method

Participants. Participants were 10 male and 14 female undergraduates at the University of Wisconsin–Milwaukee between the ages of 18 and 27, participating for course extra credit.

Materials. The stimuli consisted of six high school yearbook head photographs from the 1950s. The era was chosen for the greater homogeneity in clothing and hairstyle, which would presumably encourage all participants to use facial features to discriminate among the faces. Each photo was of a young Caucasian male presented in black and white. The six photos were chosen out of several hundred because of their neutral facial expressions and their lack of unusual features. Faces were chosen as stimuli because they have been used in other TI tasks (Nagode & Pardo, 2002; Preston et al., 2004) and because, to humans, faces are environmentally salient and easily distinguishable, so that their use is comparable to that of odors in rat inference paradigms (Eichenbaum, Fagan, Mathews, & Cohen, 1988). As can be seen in Figure 1A, stimulus arrays were always presented in an inverted triangle formation, with two faces above (constituting the forced choice) and a single face below (constituting the DS). These computer-generated stimuli were presented on a 19-in. computer monitor. Participants responded to the stimuli by pressing either the left or right mouse button to indicate their choice between the two upper faces.

Procedure. For a given array, the position of the two upper faces alternated randomly throughout training and testing so that participants could not simply choose either right or left. Prior to training, participants were informed that they would be presented with three faces on a screen and that the face on the bottom of the array would determine their choice between the two faces at the top. The training phase consisted of four blocks, each containing 10 presentations of the arrays. In the first training block, each array was shown 10 times consecutively for a total of 40 trials. In only the first training block, participants were cued to the correct response; during this phase, the faces were presented with a green arrow that indicated the correct choice and a red arrow that indicated the incorrect choice. If the green arrow pointed to the face on the left, the participant learned to press the left mouse key; if it pointed to the face on the right, the participant learned to press the right arrow key. In all training blocks, choice was followed by feedback that was presented on the computer monitor for 2 sec (“correct” in white or “incorrect” in yellow, both against a black background). During the training phase, when the choice was incorrect, the presentation was repeated until the participant responded correctly.

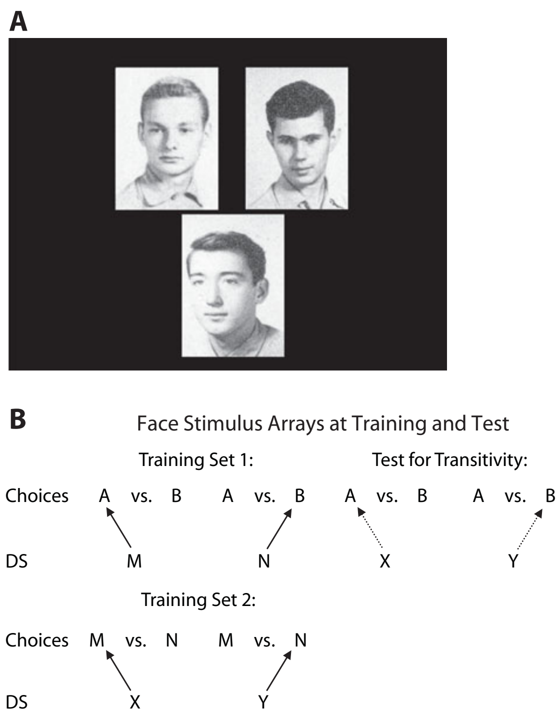


Figure 1. (A) Typical stimulus array used during training and test. (B) Training and test contingencies: The face at the bottom of the triangular array determines which of the two upper faces is the correct choice. Arrows in the training arrays indicate trained choices, and dashed arrows in the test condition indicate expected transitive choice without feedback.

In the second training phase, the arrays were shown 5 times consecutively. This sequence was repeated once for a total of 10 presentations of each array of each array type for a total of 40 trials total. In the third training phase, stimuli were shown 2 times consecutively and were then repeated for a total of 10 presentations of each array type, for a total of 40 trials. The fourth training phase consisted of a randomly generated sequence of 10 presentations of each array type for a total of 40 trials. As with the first training block, feedback was given and incorrect trials were repeated. However, the arrows were not used after the first training block unless the participant made an incorrect choice, in which case, the arrows were used on the repeated presentation to indicate the correct choice. A 20-sec break followed each training block. Participants repeated any training block for which they did not respond correctly on at least 80% of the trials. If a participant did not reach 80% by the third repetition of a given block, the experiment was terminated at that point. Three participants failed to reach training criteria and were replaced.

Following the completion of the four training blocks, participants were tested on each of the original arrays as well as on two novel arrays that required an inference. The test phase consisted of 20 randomized presentations of each array. Before starting the test phase, the participants were informed that feedback would no longer be given for the remainder of the task.

Upon completion of the session, participants were given a questionnaire (shown in Appendix A) to assess their level of awareness during the tasks. Six raters independently evaluated posttask questionnaires after training on uniform scoring criteria (detailed in Appendix B). The questionnaire and scoring process mirrored the procedure used in several other studies exploring awareness during inference tasks *in which a strong relationship between awareness and performance was found* (Frank et al., 2005; Greene et al., 2006; Greene et al., 2001; Smith & Squire, 2005); the questionnaire in the present experiment is identical, except that two questions were revised to reflect the different stimulus and hierarchy type (see Appendix A). The questionnaire probes the explicit strategies, if any, used by participants. Note that for an explicit strategy to be effective, it must first be discovered by the participant, and then it must be sufficient to organize all the task contingencies for correct performance. To determine whether an explicit strategy was used, the questionnaire asked participants what they did, beginning with general questions about the purpose of the task and continuing with increasingly leading questions about the global relationships among items. Criteria scores of awareness ranged from 1 to 5, with 1 corresponding to *no evidence of task awareness* and 5 indicating *definite awareness of the task*. To assess interrater consistency, a Cronbach's α test was performed, which determined high consistency among raters ($\alpha = .87$).

Results and Discussion

Figure 2 shows the mean accuracy scores and standard errors for the study arrays and inference arrays at test. Participants scored very close to perfect on the trained arrays and with very high accuracy on the inference arrays. Figure 3A shows the distribution of awareness scores versus test performance for inference arrays at test. To assess a relationship between the participant's awareness on the task and his/her performance, we computed a Pearson's correlation for the two variables and did not detect a significant relationship between performance and awareness on the inference arrays at test ($r = -.003$, n.s.). In the present inference task, only 3 participants did not score above chance (binomial $p < .01$ requires 70% correct) on the inference arrays at test. Of these 3 participants, 1 had an awareness rating of 3.4, indicating some awareness of task contingencies, and 2 had awareness ratings of 2, indicating no awareness of an inference. Overall, the participants who were rated as unaware (awareness rating ≤ 2.0) performed significantly above chance ($M = .916$, $SEM = .045$) [$t(12) = 9.147$, $p < .001$].

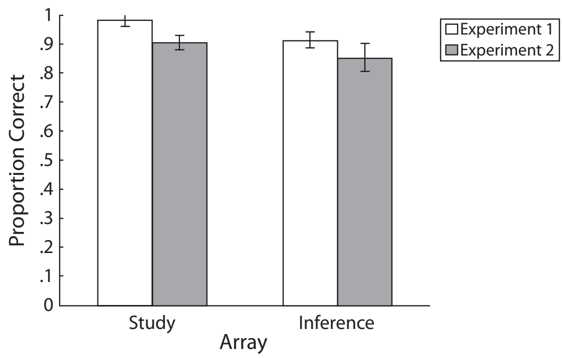


Figure 2. The mean accuracy and standard error during testing on the transitive inference test.

Analysis of reaction times at test revealed that correct decision times for studied arrays ($M = 1{,}498.8$ msec, $SEM = 20.7$) and inferred arrays ($M = 1{,}459.8$ msec, $SEM = 25.8$) did not differ [$t(23) = 0.384$, n.s.]. There was no significant relationship between postexperimental awareness and reaction time for inferred arrays ($r = .004$, n.s.).

EXPERIMENT 2

Experiment 1 demonstrated that both aware and unaware participants achieved high inference performance. However, ceiling effects may have concealed a relationship, if any exists, between awareness and performance. That is, it may be that awareness of the hierarchy could have a beneficial effect on inference performance, but the scores were too high to show such a differential effect. In order to detect any potential benefits of task awareness, we modified the training phase to make the task more difficult, which we expected would decrease inference performance, thereby avoiding ceiling effects.

Method

Participants were 7 male and 14 female undergraduates at the University of Wisconsin–Milwaukee between the ages of 18 and 28, participating for course extra credit. The materials and procedure were identical to those used in Experiment 1, except that the first block of training was eliminated (the block in which all trials received directed feedback) and the incorrect responses in training trials did not result in directed feedback (simply "correct" or "incorrect" feedback was given). Therefore, the participants in Experiment 2 were required to acquire the relationships between the faces by trial and error, consistent with other inference experiments (Greene et al., 2006; Greene et al., 2001; Smith & Squire, 2005).

The training phase consisted of three training blocks. In the first, training arrays were shown 5 times consecutively and then repeated once, for a total of 10 presentations of each array type for a total of 40 trials. In the second training block, stimuli were shown 2 times consecutively and repeated for a total of 10 presentations and 40 trials. The third training block consisted of a randomly generated sequence of 10 presentations of each array type for a total of 40 trials. As in Experiment 1, the participants were given feedback about

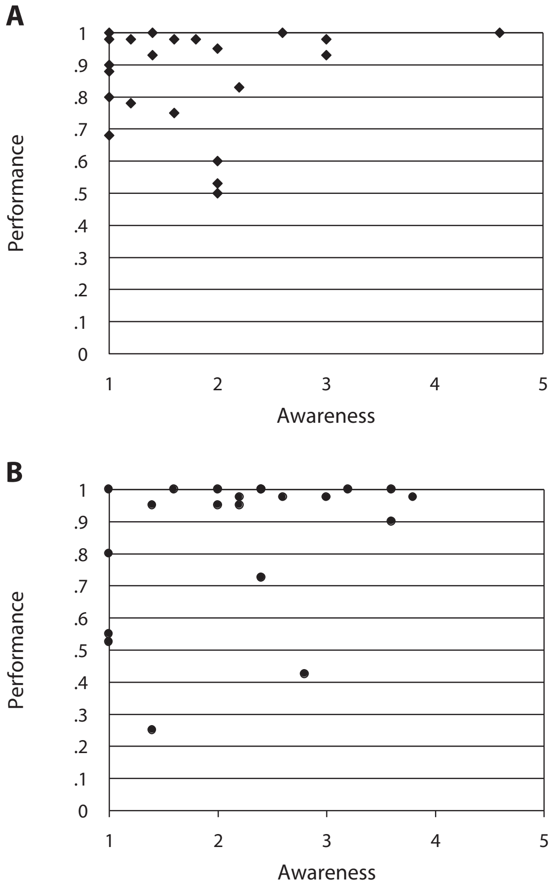


Figure 3. (A) Scatterplot for Experiment 1 of awareness scores versus performance for inference arrays at test. (B) Scatterplot for Experiment 2 of awareness scores versus performance for inference arrays at test.

the accuracy of their selection after each response. The test phase of Experiment 2 was identical to that of Experiment 1: To assess interrater consistency, a Cronbach's α test was performed, which determined high consistency among raters ($\alpha = .894$).

Results and Discussion

Figure 2A shows the mean accuracy scores and standard errors for the study arrays and for the inference arrays at test. Participants scored very close to perfect on the trained arrays and with very high accuracy on the inference arrays. Figure 3B shows the distribution of awareness scores versus test performance for Experiment 2 at test. To assess a relationship between the participants' awareness on the task and their performance, we computed a Pearson's correlation for the two variables and detected a trend relationship between performance and awareness ($r = .358$, one-tailed $p = .055$). To further assess a possible role of awareness, we dichotomized participants according to their response to Question 5 of the posttask questionnaire, which asked participants for the reason why they selected the face that they did on the novel inference array at test from among several options. The first option was that there was a logical solution to the problem; this was followed by other options that did not indicate awareness of the logical relationship (i.e., "One just seemed right, but I can't explain why."). It was possible that the awareness ratings assigned by the raters were biased by the participants' ability to communicate their awareness. Dichotomizing participants according to Question 5 removed the demand for them to explain the reason for their selection. When awareness was determined by Question 5, there was a significant relationship between awareness and performance on the inference items at test ($r = .404$, one-tailed $p = .035$). By eliminating the ceiling effect, a role for awareness in performance was detected, despite which, the participants who were rated as unaware (awareness rating ≤ 2.0) still performed significantly above chance ($M = .781$, $SEM = .091$) [$t(8) = 3.068$, $p < .05$].

Analysis of reaction times at test revealed that correct decision times for studied arrays ($M = 1{,}508.8$ msec, $SEM = 26.5$) and inferred arrays ($M = 1{,}625.5$ msec, $SEM = 37.9$) did not differ [$t(21) = -1.511$, n.s.]. There was no significant relationship between postexperimental awareness and reaction time for inferred arrays ($r = -.230$, n.s.).

EXPERIMENT 3

It was possible that successful inference required the use of deliberative strategies not detected by the questionnaire. To test for the use of deliberative strategies, we devised a recognition task. We noted that during the experimental debriefing, many participants expressed surprise that there were items at the test that had not been trained at study. We reasoned that the use of deliberative strategies for inference would require that participants know which arrays they had not learned because, at test, participants must recognize and apply a deliberative strategy to unlearned arrays. We therefore implemented a *yes–no* decision task following study in which participants simply determined whether an array was one of the studied arrays.

Method

Participants were 5 male and 7 female undergraduates at the University of Wisconsin–Milwaukee between the ages of 18 and 27, participating for course extra credit. The materials were identical to those used in Experiment 1. The study session was identical to that used in Experiment 1. At test, participants were instructed that arrays would appear on the screen and that, although they would all involve faces from the study session, the task was to indicate whether the specific array of faces had been learned at study. Test arrays were the four studied arrays, and eight novel arrays. The novel arrays consisted of two inference arrays, two arrays that had potential symmetry solutions (X and Y as the choice items and either M or N as the DS), and four foil arrays that had no possible solutions (e.g., A and X as the choice items and B as the DS; of all the possible foils, four were selected randomly for each participant). One criticism of the posttask questionnaire used in Experiments 1 and 2 was that raters might not have been scoring the participants' level of awareness, but rather that the scores indicated how well participants were able to communicate their strategies. The design of Experiment 3 allowed for three basic outcomes that would help us to identify the types of strategy that the participants used.

First, if participants correctly accepted studied arrays and correctly rejected inference arrays, symmetry arrays (i.e., when the DS stimuli in the array was switched with its correct choice stimuli at test), and foil arrays, it could be argued that they employed a deliberative strategy during the task. If participants correctly accepted the studied items and correctly rejected the foils but incorrectly accepted the transitive arrays and symmetry arrays, it was possible that participants were not initiating a deliberative strategy. Finally, if participants correctly accepted the studied arrays and correctly rejected both the inference arrays and the foil arrays, but incorrectly accepted the symmetry arrays, it was likely that participants simply responded "yes" when all items in the study array were present in the array at test. That is, replacing the two choice stimuli with two discriminative stimuli (X and Y) that had never been presented together eliminated the possibility that participants were simply responding "yes" to the two choice stimuli at the top of the array and were ignoring the discriminative stimuli at the bottom. Experiment 3 added converging data about the presence or absence of a deliberative explicit strategy.

Results and Discussion

The recognition task revealed that participants consistently mistook the inference and symmetry arrays for studied arrays, which is inconsistent with the use of deliberative strategies. In addition to *t* tests, Cohen's d' is included in order to assess potential response biases (Stretch & Wixted, 1998). Planned comparisons revealed that the rate of correct recognition for studied items (84.34%, $SE = 2.24\%$) did not differ from the rate of false alarms for either inference arrays (83.33%, $SE = 7.11\%$) [$t(11) < 1$; $d' = 0.04$] or symmetry arrays (79.17%, $SE = 7.43\%$) [$t(11) < 1$; $d' = 0.20$]. The rate of false alarms for foils (27.08%, $SE = 4.83\%$) was significantly different than those for inference arrays [$t(11) = 8.07$, $p < .01$; $d' = 1.57$] and symmetry arrays [$t(11) = 7.24$, $p < .01$; $d' = 1.42$] and also significantly different from the hit rate for studied items [$t(11) = 8.21$, $p < .01$; $d' = 1.62$].

GENERAL DISCUSSION

The present study provides evidence that humans are able to make correct inferences without employing pseudoinference and that this can be done without task awareness.

Previous studies have asserted that inference without awareness is either near chance (Smith & Squire, 2005) or that it relies on pseudoinference, as in the VTT or the blocking hypothesis, which depend on end-item weighting (Frank et al., 2006; Frank et al., 2005; van Elzakker et al., 2003). The present design of the inference task ruled out the possibility that pseudoinference strategies could be employed, although we observed that above-chance performance occurred in the absence of deliberative strategies.

The findings of Experiment 1 show performance accuracy at greater than 91% on inference arrays among participants with no evidence of awareness. That we did not detect a benefit of task awareness is unsurprising, given robust performance across levels of awareness. However, ceiling effects and restriction of range may have precluded detection of an added benefit of awareness in Experiment 1. In Experiment 2, when the task was more difficult, we still found high performance among unaware participants, but we also found that in the absence of ceiling effects, task awareness did facilitate somewhat higher performance.

The results of Experiment 3 provide additional evidence that inference at test is unlikely to depend on deliberative strategies. Because participants in Experiment 3 correctly accepted the studied arrays, but incorrectly accepted the inference arrays and symmetry arrays, it is unlikely that they used a deliberative strategy to organize the relationship between the arrays at study. If participants had used a deliberative method, they should have been able to easily reject inference arrays as ones that they had not seen previously. Therefore, we assert that deliberative strategies are not necessary for inference to occur, but that explicit strategies may enhance performance. Results from linear transitive inference tasks have converged on the same finding: Performance may benefit from task awareness, but above-chance performance still occurs in the absence of task awareness (Frank et al., 2005; Greene et al., 2001; Moses, Villate, & Ryan, 2006). Participants who become explicitly aware of task contingencies show modestly better performance because, when confronted with novel arrays, they are able to access the associational framework consciously. Previous studies have asserted that flexible encoding requires awareness (e.g., Clark et al., 2002; Smith & Squire, 2005). Regarding the present tasks, many participants had very little or no awareness and yet performed at levels well above chance. It is clear that high performance on these tasks is possible regardless of task awareness. This evidence is convergent with previous studies that have shown a logical relational process taking place during the TI task, irrespective of task awareness (Greene et al., 2006; Greene et al., 2001). The finding is part of a growing body of evidence demonstrating learning that is highly flexible under novelty, but that does not require awareness (e.g., Gross & Greene, 2007; Harrison, Duggins, & Friston, 2006; Myers et al., 2003; Ryan, Althoff, Whitlow, & Cohen, 2000).

The principle motivation for implementing the present inference task was to test whether implicit performance could be achieved on a task that did not have potential pseudoinference solutions. Pseudoinference hypotheses that pertain to linear TI tasks are not possible strategies for solving the present inference task. Although these findings do not rule out the possibility that linear TI (e.g., $A > B > C > D > E$) may be solved with pseudoinference solutions, they do demonstrate that such solutions are not required for implicit inference to occur. Indeed, pseudoinference approaches are predicated by the assumption that true inference solutions require deliberative strategies (e.g., Frank et al., 2005; Zentall & Sherburne, 1994).

AUTHOR NOTE

We thank Matthew Verber and Mark Oium from the Medical College of Wisconsin General Clinical Research Center for their technical assistance on this study. Address correspondence to P. D. Leo, Department of Psychology, Garland Hall Room 224, University of Wisconsin, Milwaukee, WI 53211 (e-mail: pdleo@uwm.edu).

REFERENCES

Bunsey, M., & Eichenbaum, H. (1996). Conservation of hippocampal memory function in rats and humans. *Nature*, **379**, 255-257.

Clark, R. E., Manns, J. R., & Squire, L. R. (2002). Classical condi-

tioning, awareness, and brain systems. *Trends in Cognitive Sciences*, **6**, 524-531.
Davis, H. (1992). Transitive inference in rats (*Rattus norvegicus*). *Journal of Comparative Psychology*, **106**, 342-349.
Eichenbaum, H., Fagan, A., Mathews, P., & Cohen, N. J. (1988). Hippocampal system dysfunction and odor discrimination learning in rats: Impairment or facilitation depending on representational demands. *Behavioral Neuroscience*, **102**, 331-339.
Ellenbogen, J. M., Hu, P. T., Payne, J. D., Titone, D., & Walker, M. P. (2007). Human relational memory requires time and sleep. *Proceedings of the National Academy of Sciences*, **104**, 7723-7728.
Frank, M. J., O'Reilly, R. C., & Curran, T. (2006). When memory fails, intuition reigns: Midazolam enhances implicit inference in humans. *Psychological Science*, **17**, 700-707.
Frank, M. J., Rudy, J. W., Levy, W. B., & O'Reilly, R. C. (2005). When logic fails: Implicit transitive inference in humans. *Memory & Cognition*, **33**, 742-750.
Frank, M. J., Rudy, J. W., & O'Reilly, R. C. (2003). Transitivity, flexibility, conjunctive representations, and the hippocampus: II. A computational analysis. *Hippocampus*, **13**, 341-354.
Gillan, D. J. (1981). Reasoning in the chimpanzee: II. Transitive inference. *Journal of Experimental Psychology: Animal Behavior Processes*, **7**, 150-164.
Greene, A. J. (2007). Human hippocampal-dependent tasks: Is awareness necessary or sufficient? *Hippocampus*, **17**, 429-433.
Greene, A. J., Gross, W. L., Elsinger, C. L., & Rao, S. M. (2006). An fMRI analysis of the human hippocampus: Inference, context, and task awareness. *Journal of Cognitive Neuroscience*, **18**, 1156-1173.
Greene, A. J., Spellman, B. A., Dusek, J. A., Eichenbaum, H. B., & Levy, W. B. (2001). Relational learning with and without awareness: Transitive inference using nonverbal stimuli in humans. *Memory & Cognition*, **29**, 893-902.
Gross, W. L., & Greene, A. J. (2007). Analogical inference: The role of awareness in abstract learning. *Memory*, **15**, 838-844.
Harrison, L. M., Duggins, A., & Friston, K. J. (2006). Encoding uncertainty in the hippocampus. *Neural Networks*, **19**, 535-546.
Heckers, S., Zalesak, M., Weiss, A. P., Ditman, T., & Titone, D. (2004). Hippocampal activation during transitive inference in humans. *Hippocampus*, **14**, 153-162.
Howard, M. W., Fotedar, M. S., Datey, A. V., & Hasselmo, M. E. (2005). The temporal context model in spatial navigation and relational learning: Toward a common explanation of medial temporal lobe function across domains. *Psychological Review*, **112**, 75-116.
Moses, S. N., Villate, C., & Ryan, J. D. (2006). An investigation of learning strategy supporting transitive inference performance in humans compared to other species. *Neuropsychologia*, **44**, 1370-1387.
Myers, C. E., Shohamy, D., Gluck, M. A., Grossman, S., Onlaor, S., & Kapur, N. (2003). Dissociating medial temporal and basal ganglia memory systems with a latent learning task. *Neuropsychologia*, **41**, 1919-1928.
Nagode, J. C., & Pardo, J. V. (2002). Human hippocampal activation during transitive inference. *NeuroReport*, **13**, 939-944.
Preston, A. R., Shrager, Y., Dudukovic, N. M., & Gabrieli, J. D. E. (2004). Hippocampal contribution to the novel use of relational information in declarative memory. *Hippocampus*, **14**, 148-152.
Reber, P. J., Knowlton, B. J., & Squire, L. R. (1996). Dissociable properties of memory systems: Differences in the flexibility of declarative and nondeclarative knowledge. *Behavioral Neuroscience*, **110**, 861-871.
Ryan, J. D., Althoff, R. R., Whitlow, S., & Cohen, N. J. (2000). Amnesia is a deficit in relational memory. *Psychological Science*, **11**, 454-461.
Siemann, M., & Delius, J. D. (1994). Processing of hierarchic stimulus structures has advantages in humans and animals. *Biological Cybernetics*, **71**, 531-536.
Smith, C., & Squire, L. R. (2005). Declarative memory, awareness, and transitive inference. *Journal of Neuroscience*, **25**, 10138-10146.
Stretch, V., & Wixted, J. T. (1998). Decision rules for recognition memory confidence judgments. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **24**, 1397-1410.
van Elzakker, M., O'Reilly, R. C., & Rudy, J. W. (2003). Transitivity, flexibility, conjunctive representations, and the hippocampus: I. An empirical analysis. *Hippocampus*, **13**, 334-340.
Weaver, J. E., Steirn, J. N., & Zentall, T. R. (1997). Transitive inference in pigeons: Control for differential value transfer. *Psychonomic Bulletin & Review*, **4**, 113-117.
Zentall, T. R., & Sherburne, L. M. (1994). Transfer of value from S+ to S− in a simultaneous discrimination. *Journal of Experimental Psychology: Animal Behavior Processes*, **20**, 176-183.

APPENDIX A
Posttask Questionnaire

(Page breaks and instructions not to return to the previous page after Questions 3, 6, and 8.)

Questions:

1. What did you think we were trying to find out with this experiment?

2. What did you think was the point of the trials where you were not told if you were right or wrong (the no feedback condition)?

3. Regarding Question 2, were all the face configurations in the no feedback condition the same as face configurations you had already learned when feedback was given? ___ Yes ___ Not Sure ___ No

   If no, do you think there was a correct answer? ___ Yes ___ Not Sure ___ No

   If you believe there was a correct answer, explain why:

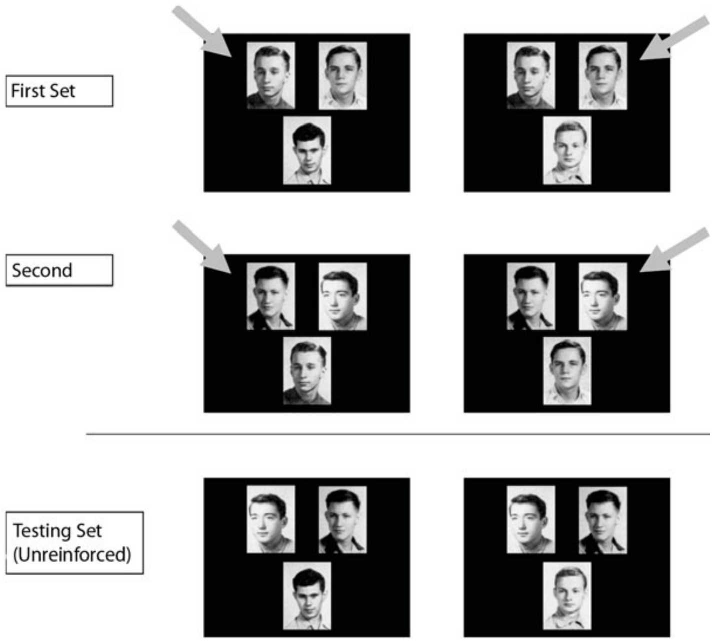
4. You were given the following combination of faces several times, but never told an answer. Circle the face you believe is correct (guess if necessary):

APPENDIX A (Continued)

5. Regarding Question 4, what reason (if any) did you have for your choice (check one):
   ___ There is a logically correct choice because (explain):
   ___ One just seemed right but I can't explain why.
   ___ I guessed: There may be a correct face, but I don't know what it is.
   ___ I made a random choice because there is no correct choice.
   ___ Other: Explain____

6. What strategy (if any) did you use to learn the faces (check one):
   ___ I already knew the faces: If so, from where? ______________________________
   ___ I gave them names. ___ I memorized each face. ___ I memorized part of each face.
   ___ I just watched and eventually got it. ___ I used their similarity to familiar faces.
   ___ No strategy. ___ Other strategy (please describe in the space below).

7. Below is a schematic of the organization of the training you received with an arrow pointing to the face of the correct choice that you learned.



   Based on your understanding of the task, circle the correct choice among the top two faces for both sets.

8. If applicable, please indicate how much knowledge you have of formal logic, syllogism, or transitive inference:

(*Continued on next page*)

APPENDIX B
Scoring Awareness

Raters were trained to assign a rating between 1–5 in the following manner: 5 = *an explicit statement of the transitive relationship*; 4 = *probable knowledge of the contingencies, although not directly stated*; 3 = *possible, but vague knowledge of task contingencies*; 2 = *uncertain reference to relationships among items, but no indication of task contingencies*; 1 = *no evidence of knowledge of contingencies or purpose of the task*.

Ratings on the basis of question responses were assigned as follows. Questions 1–3: some assertion that there was a hierarchy or an inference resulted in a score of 4 or 5, depending on the clarity of the assertion. Some participants clearly recognized that they have made an inference about novel configurations and made explicit statements to that effect when they were asked to speculate about the purpose of the experiment. Questions 4 and 5: Question 4 displays a transitive array and asks the participant which face they chose, whereas Question 5 asks the participant to explain the reason for their selection on Question 4. The first of several choices is "There is a logically correct choice because (explain)," and the last is "I made a random choice because there is no correct choice." Those selecting that there was a logical choice and then providing a rationale similar to the logic of inference (regardless of whether it was correct) were given a score of 3 or 4, depending upon the clarity. Those who provided no logical rationale but provided vague statements that they knew there was a relationship among items were given a score of 2 or 3 depending on the clarity. If participants did not choose the first option (a logical choice was made) they were assigned a score of 2 or 1. Those who chose that they had guessed were given a score of 1. During the debriefing, when the hierarchy was explained explicitly, many participants were quite surprised to learn of its existence and that they had answered correctly.

(Manuscript received May 17, 2007;
revision accepted for publication April 2, 2008).