Look and Learn: Observational Learning of Rules and Instances

Rosemary J. Stevenson [rosemary.stevenson @durham.ac.uk]
Bruce W. Geddes, J. Beth Sumner & Bella M. K. Travis
Department of Psychology,
University of Durham,
Durham, DH1 3LE, U.K.

Abstract

We describe an experiment that examines observational learning of either rules or instances. Subjects were asked to learn a dynamic computer control task and were given either a specific goal, to make the computer produce a specific response, or a non-specific goal, to find the pattern underlying the computer's behaviour. Subjects either interacted directly with the computer (the 'models') or observed a model's learning trials (the 'observers'). Both the goal of the models and the goal of the observers were varied so that specific goal and non-specific goal models were crossed with specific goal and non-specific goal observers. We predicted that the goal of the observer and not the goal of the model would determine whether observers learned rules or instances and that learning through observation would hinder instance learning. These predictions were confirmed. Non-specific goal models learned rules whereas specific goal models learned instances. Non-specific goal observers also learned rules, irrespective of the goal of the model, but specific goal observers failed to learn at all. A subsequent test confirmed that the failure of the specific goal observers to learn was due to the lack of feedback about correct responses. When such feedback was provided, specific goal observers learned instances. However, the presence of feedback was detrimental to rule learning. When non-specific goal observers received feedback, they learned only instances. These results support the view that both goal specificity and the presence or absence of feedback guide learning by directing attention to either instance space or both instance space and rule space.

Introduction

Rule learning has been distinguished from instance learning in both research on concept learning (Erickson & Kruschke in press; Lebowitz, 1986; Wisniewski, & Medin, 1995) and research on implicit learning (Geddes, & Stevenson, 1997; Shanks & St. John, 1994). Geddes & Stevenson (1997) showed that when learning to control a dynamic control task, whether subjects learn rules or instances depends on their learning goal: If they have a specific goal (to control the system), they learn rules; if they have a non-specific goal (to understand the system), they learn rules. In this paper, we investigate the learning of rules and instances through observation so that the roles of action and of feedback can be examined.

Geddes and Stevenson (1997) used a dynamic control task in which subjects interact with a 'computer person' called Clegg and try to get him to become and stay *Very Friendly*. Clegg initiates the interaction by displaying one of twelve attitudes (e.g. *Polite, Very Friendly, Loving*) on the computer screen, after which the subject responds by typing in another attitude. The attitudes reflect an intimacy scale from low to high and Clegg's response to the subject's choice of attitude is retaliatory. If Clegg is *Polite*, and the subject responds with *Friendly*, then Clegg retaliates with the attitude *Loving*. Clegg's attitude on each trial is a simple numerical function of the subject's response on that trial and Clegg's previous output. Subjects successfully learn to carry out this task, but when questioned about the experiment afterwards, they are unable to describe how they did it or what the underlying rule is (Berry & Broadbent, 1984).

In Geddes and Stevenson's study, one group of subjects was given a specific learning goal, comparable to the learning goal used in Berry and Broadbent (1984). Subjects were instructed to make Clegg polite and stay polite. However, in contrast to Berry and Broadbent, Geddes and Stevenson gave a second group of subjects a non-specific learning goal. These subjects were instructed to find out the pattern that explained Clegg's behaviour.

All the subjects had 30 learning trials, after which they were tested on what they had learned. In the first test, subjects in both goal groups were given 30 trials to learn a novel specific goal - to make Clegg very friendly. The results showed that non-specific goal subjects performed better than specific goal subjects with the novel specific goal (52% correct responses vs. 41%). In a second test, all subjects predicted Clegg's response, given a sequence of three responses. For example, a subject might be told "You were *very cool*, then Clegg was *very rude*, You were then *polite*. What did Clegg do next?" Some of these prediction questions described 'old' situations, which the subject had encountered during learning. Others described 'new' situations, which the subject had not seen before. Non-specific goal subjects made correct predictions in both old and new situations while specific goal subjects only made correct predictions in old situations. In a third test, subjects were asked to describe the rule that governed Clegg's behaviour. Whereas 79% of the non-specific goal subjects gave either correct or partially correct rule descriptions, over 80% of the specific goal subjects gave wrong descriptions.

Thus, subjects given a non-specific goal learned the abstract rule underlying Clegg's behaviour while subjects given a specific goal remembered specific responses. These results are consistent with other evidence suggesting that the learning goal can have profound effects on learning, whether it be instance learning (Whittlesea & Dorken, 1993) or rule learning (Owen & Sweller, 1986; Vollmeyer, Burns & Holyoak, 1996).

"Dual space" models of learning explain rule learning and instance learning within a single framework (Klahr & Dunbar, 1988; Simon & Lea, 1974). Simon and Lea, for example, proposed that the problem space is separated into two spaces: a rule space and an instance space. People search instance space when seeking the solution to a specific goal. Geddes and Stevenson suggested that one way in which instance space is searched to reach a specific goal is through means-ends analysis, involving successive reductions of the difference between the learner's current state and the goal state until the goal is reached. Thus, what get learned are the specific states encountered on the route to the goal. On the other hand, people search both rule space and instance space when generating and testing hypotheses. Explicit hypotheses are generated in rule space, which are then tested by experiments that generate states in instance space. In these circumstances, subjects learn rules that explain the system being studied.

In order to explore the boundary conditions of instance learning and rule learning, the present study was conducted. Instead of having our learners interact with the computer, we asked them instead to observe other learners. In addition, we systematically varied the goal of the model and the goal of the observer so that the observer had either a control task goal or a pattern search goal and observed a model who also had either a control task or pattern search goal. Berry (1991) originally used this observation procedure, but both her models and observers were given a control task goal. She found that the observers' learning was very poor under these conditions, suggesting that action is the critical ingredient for instance learning. However, Berry did not explicitly check that her subjects were learning instances, hence in the present study we used the same tests of learning as Geddes and Stevenson (1997) so that we could examine the effectiveness of observational learning on the learning of instances compared to rules. When learning to control the system, and so learning instances, the decisions made by the learner concern which response to make on each trial. These decisions are less likely to be made if the learner observes someone else making and implementing those decisions. Consequently, in accordance with Berry's (1991) results, we hypothesize that instance learning will be inhibited during observational learning. When trying to understand the system, the decisions made by the learner concern generating and testing hypotheses. Since these decisions are purely cognitive, they should be unaffected by whether the learner interacts with the computer or observes someone else interacting. Consequently, we also hypothesize that rule learning will be successful during observational learning.

Method

Subjects

Seventy two student volunteers from Durham University served as subjects. Their ages ranged from 18 to 24 years. Twenty four were models, 12 in each goal group; 48 were observers, 12 in each of the four groups defined by the goal of the model and the goal of the observer.

Design

A two (goal of model) by two (goal of observer) independent groups design was used for the four observer groups. Both observers and models were given either a control task goal or a pattern search goal. Half the control task observers observed control task models and half observed pattern search models. Similarly, half the pattern search observers observed control task models and half observed pattern search models. See Table 1. The two groups of models were tested first after which the observers were tested.

Table 1: Design of the Experiment (**Spec.** = Specific)

Goal of Model	Spec. Goal		Non-Spec. Goal	
Goal of Observer	Spec. Goal	Non-Spec. Goal	Spec. Goal	Non-Spec. Goal

All subjects were required to complete 30 learning and 30 test trials. The goal groups were defined by the nature of the goal in the 30 learning trials, either specific ('Make Clegg polite') or non-specific ('Find the underlying pattern'). The models interacted with the computer during learning whereas the observers observed the models' learning trials. In the test trials, all subjects were given a new specific goal ('Make Clegg very friendly') and they all interacted with the computer. After the test trials, all subjects were given two further (unexpected) tests of learning: predicting Clegg's next response from a sequence of three responses and answering questions designed to elicit descriptions of the rule underlying Clegg's behaviour.

Learning and Test Trials Models were told that they would be meeting a computer person named Clegg and would communicate with Clegg through the screen and keyboard. Clegg would express his attitude towards them by displaying one of twelve descriptions (*Very Rude, Rude, Very Cool, Cool, Indifferent, Polite, Very Polite, Friendly, Very Friendly, Affectionate, Very Affectionate, Loving*). Following this, subjects responded to Clegg by choosing one of the above descriptions. This was done by typing in the first letter or letters of that description (e.g. VP for *Very Polite*). Once subjects had responded, Clegg would display his new attitude (produced by the equation described below). It would then be the subject's turn to enter their next attitude, and so on. The list of possible responses was displayed on a piece of paper attached to the bottom of the screen for permanent reference.

In addition to the above instructions, each group of models was given specific instructions concerning their learning goal and their secondary task. Models in the control task group were told "Your aim is to shift Clegg to the *Polite* level and maintain him at that level". Models in the pattern search group were told "Your aim is to establish under what pattern Clegg is reacting". To remind subjects of their respective goals, the goal of their task was permanently displayed on a piece of paper attached to the bottom of the screen.

On each trial Clegg's and the subject's responses were displayed on the screen. These scrolled up the screen so that it was possible to see the previous six trials on the screen at any one time. The equation relating Clegg's responses to those of the subject's was identical to the non-salient rule used by Berry and Broadbent (1984). The descriptions were given a value from 1 (*Very Rude*) to 12 (*Loving*) and Clegg's response was determined by the equation:

$$CNR = (2 \times SOR) - COR + Z,$$

where CNR = Clegg's new response, SOR = subject's old response, COR = Clegg's old response and Z = a random number with the value of -1, 0 or +1. The random element in the equation ensures that subjects must exercise continuous control over the computer person. It also means that there is no unique input associated with any one output. If subjects reached their target output then simply re-entering the same input is unlikely to keep them on target (Berry & Broadbent, 1984). To allow for the random element in the equation producing Clegg's response, the responses of subjects in the specific goal group were scored as correct if they were either on the target or one response either side of the target. That is, a response from Clegg of *Indifferent, Polite,* or *Very Polite* was scored as correct.

Each observer was randomly assigned to a model with the relevant goal. In addition to the general instructions about the task, the control task observers were told "For this section of the experiment however, you will not interact with Clegg, but, will view some interactions that have occurred. You should watch what the person has done on the earlier occasion as this should give you a feel for how Clegg responds. It is important you pay close attention to the interactions you shall be viewing as, later, you will have to control Clegg, making him produce a specific output and then maintaining his output at the specific level." For the pattern search observers, the final sentence was changed to "It is important you pay close attention to the interactions you shall be viewing as, later, you will have to establish under what pattern Clegg is reacting." The observers were not told the task that the model had been set. Observers pressed the space bar of the computer to display each trial of their model. With each key press, the trials scrolled up the screen in the same way as they did for the models.

The test trials were identical to the learning trials for the control task models except that the goal was changed. As was the case in the learning trials, a response either on the target or one step either side of the target was scored as correct, to allow for the random element in the equation.

Prediction Questions There were 15 prediction questions, 5 new, 5 old correct and 5 old wrong. For each question, a typical trial situation was presented. The subject's and Clegg's behaviour was displayed on the screen, below this the subject's new behaviour was displayed - e.g. You were *Very Cool,* Clegg was *Very Rude,* You were then *Polite.* Subjects then had to predict what Clegg's response would be. The five 'new' situations were generated randomly from a list of all possible trial situations that the subject had not encountered during either the learning trials or the testing trials. The five *'Old-wrong'* situations were randomly selected from all the trials the subject had got wrong during the test phase. The five *'Old-correct' situations* were randomly selected from all the trials the subject had got correct during the test phase. To produce five *Old-wrong* and five *Old-correct* questions meant that the subject must get at least five wrong or five correct respectively during the test trials. The program controlling the experiment allowed for the possibility of this not occurring and would have substituted any uncreated questions with *New* questions.

Rule descriptions. Two questions tested the subjects' ability to describe the rule underlying Clegg's behaviour. One was "How did you get Clegg to behave as you wanted him to?" This question was designed to be sensitive to any procedural knowledge that may have been acquired during learning. The other question was "Could you try to describe what sort of pattern you thought Clegg was using to respond to your behaviour?" This question was designed to be sensitive to declarative knowledge.

Procedure

Subjects were randomly allocated to one of the six experimental groups. The two model groups carried out the learning instructions while interacting with the computer, the four observer groups observed the learning trials of the models.

On completing the learning trials, all subjects were told the learning goal for the test trials and then the test trials started. Clegg initiated both learning and test trials by displaying one of the three adjectives centered on *Polite.* Following the test trials, subjects were instructed on the prediction questions. The instructions described the questions and gave an example of a prediction situation. The instructions also explained that each question was unrelated to the previous one. After completing the prediction questions subjects were given a pen and paper and were asked to answer the two general questions appearing on the paper.

Results and Discussion

Models

Test trials (Novel Specific Goal) Test trials were scored as correct for the control task subjects if they obtained a response from Clegg of *Indifferent, Polite* or *Very Polite.* This scoring takes into account the random element of the equation producing Clegg's behaviour. Both control task and pattern search models produced 43% correct test trials. These scores were significantly above the chance level of 24.7% ($p < .03$). Chance level was calculated by running 50,000 simulated sessions, each of 30 trials, in which the subjects chose any one of the 12 responses with equal probability.

Prediction questions Responses to the prediction questions were scored as correct if the response predicted by the subjects was one above, the same as, or one below the response expected from Clegg in each situation. The response expected from Clegg was calculated by using the equation from the learning phase of the experiment, but not including the random element of the equation, since the scoring process took it into account. All subjects produced sufficient correct and incorrect responses in the test trials to have 5 old correct and 5 old wrong prediction questions. The data for old-correct and old-wrong situations were combined in the results. Control task models produced 18.3% correct responses in new situations and 45.37% correct responses in old situations. Pattern search models produced 66.7% correct responses in new situations and 71.7% correct responses in old situations. A two (learning goal) by two (prediction situation) analysis of variance with repeated measures on the last factor revealed a significant main effect of goal ($F(1,22)=16.26$, $p<.001$), a significant main effect of type of situation ($F(1,22)=10.05$, $p<.001$), and a significant interaction ($F(1,22)=4.75$, $p<.04$). Pattern search models produce more correct responses than control task models; there were more correct responses in old than in new situations; and the difference between old and new situations was confined to control task models ($t(11)=3.05$, $p<.01$; pattern search models: $t(11)=1.03$).

Rule descriptions Subjects' answers to the two questions about the rule (asking how to control Clegg and asking what was Clegg's underlying pattern) were treated together as subjects generally answered only one of the questions and included information in that answer that was relevant to both questions. The answers were judged by two judges and placed into one of three categories; *No information or Wrong, Partially Correct, Correct.* Answers were categorized as *No information or Wrong* if subjects gave no relevant information about the pattern Clegg was following or about how they controlled Clegg, and if part of the answer gave wrong information. Answers were categorized as *Partially Correct* if subjects mentioned Clegg's tendency to move along the scale beyond the subject's response (away from his own); mentioned any other information that described this approximate characteristic of Clegg's behaviour; made one precise possible prediction of Clegg's behaviour: or mentioned how Clegg's behaviour clustered around a continuous behaviour of the subjects. Answers were categorized as *Correct* when subjects mentioned Clegg's tendency to move along the scale, beyond the subject's response (away from his own) AND described the distance along the scale that Clegg would move (i.e. roughly double the distance the subject was from Clegg). Answers that made 3 or more precise possible predictions of Clegg's behaviour were also classified as *Correct*.

Control task models produced 11 wrong rule descriptions and 1 partially correct description. Pattern search models produced 7 correct rule descriptions, 3 partially correct descriptions and 2 wrong descriptions. Fisher Exact Probability tests comparing the number of answers in the *Correct* or *Partially Correct* categories with those in the *Wrong* category showed that pattern search models performed significantly better than control task models($p<.01$).

All of the above results replicate the findings of Geddes and Stevenson (1997)

Observers

Test Trials (Novel Specific Goal) Table 2 shows the percent correct test trials for observers as a function of their learning goal and the goal of their model. A 2 (goal of model) by 2 (goal of observer) analysis of variance revealed a significant main effect of goal of observer ($F(1,44)=23.37,p<.001$). Pattern search observers produced more correct trials than control task observers. There were no other significant effects. Furthermore, the performance of both pattern search observer groups was significantly better than chance (both p values $<.02$), whereas neither of the control task observer groups was significantly different from chance.

Table 2: Percent correct responses on the test trials as a function of goal of model and goal of observer.

	Goal of Model	
	Control Task	Pattern Search
Goal of Observer		
Control Task	24	18
Pattern Search	46	54

Prediction Questions The percent correct responses for old and new situations are shown in Figure one. A two (goal of model) by 2 (goal of observer) by 2 (situation type) analysis of variance was performed on the data. The results showed a significant main effect of goal of observer ($F(1,44)=47.59$, $p<.001$). The pattern search observers performed better than control task observers. There were no other significant effects. In addition, both the new and

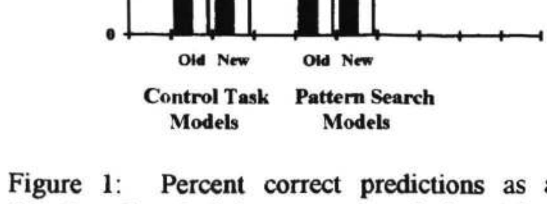

Figure 1: Percent correct predictions as a function of goal of observer and goal of model.

old correct responses of the two pattern search observer groups were significantly better than the chance level of 24% (all p's<.02), whereas none of the response of the two control task observer groups differed from chance.

Rule Descriptions The results are shown in Table 3. Fisher Exact Probability tests comparing the number of answers in the *Correct* or *Partially Correct* categories with those in the *Wrong* category showed that pattern search observers performed significantly better than control task observers (p<.01).

Table 3: Percentage of correct, partially correct and wrong rule descriptions as a function of model and observer goals. (**CT**=Control Task; **PS**=Pattern Search.)

Goal of Model	Goal of Observer	Correct	Partially Correct	Wrong
CT	CT	8	8	84
	PS	33	33	33
PS	CT	0	16	84
	PS	75	0	25

Overall, therefore, the results for the observers indicate that whereas pattern search observers learned the rule successfully, control task learners failed to learn at all, since their performance was consistently at chance level. This lack of any learning on the part of control task observers is consistent with the idea that the decisions made by instance learners must be tied to their actions for learning to occur. When the relevant actions are not performed, there is no learning.

However, there is an alternative possibility for the lack of any learning by the control task observers. In the experiment, none of the observers were told the goal of the model. Hence, when the model had a control task goal, observers were unable to tell when the model had made a correct response. Thus it may have been the lack of feedback about correct responses that was responsible for the failure of control task observers to learn. Only with feedback can control task observers decide what would be the correct response. To examine this possibility, two additional groups of observers were tested. One group was given a control task goal, the other was given a pattern search goal. Both groups observed control goal models and both groups were told the goal of the model so that they would receive feedback about the model's correct and incorrect trials. We wished to know whether or not such feedback would improve the performance of control task observers. We also wondered whether the presence of feedback would affect the performance of pattern search observers. It is possible that having feedback about correct responses might inhibit rule learning by directing the observer's attention to the correct response and away from a general exploration of the overall pattern of responses.

To examine these possibilities, we focused on the responses to the prediction questions and the results are shown in Figure two. Figure two also shows the data for the original observers of control goal models so that performance with and without feedback can be directly compared.

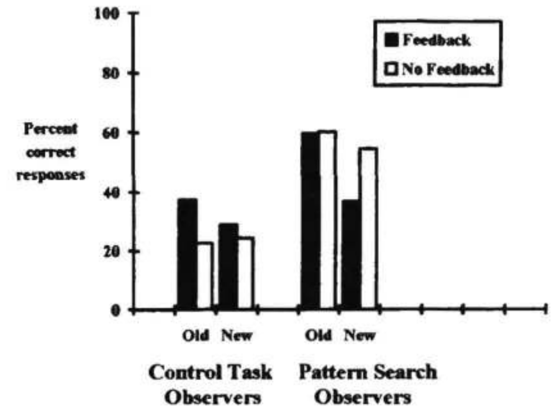

Figure 2: Percent correct predictions as a function of goal of observer and presence or absence of feedback. (All observers had control task models.)

A 2 (goal of observer) by 2 (feedback) by 2 (situation type) analysis of variance was carried out on the data in Figure 2. The results showed a significant main effect of goal of observer ($F(1,44)=12.87$, $p<.001$), a significant main effect of situation type ($F(1,44)=7.04$, $p<.02$) and a significant interaction between goal of observer and situation type ($F(1,44)=4.12$, $p<.05$). Pattern search observers produced more correct responses than control task observers; there were more correct responses in old than in new situations; and the difference between old and new situations was confined to the feedback conditions (old: 96%, new: 55%; no feedback conditions: old: 82%, new: 78%). There were no other significant effects.

As we saw above, in the absence of feedback, only pattern search observers are significantly better than chance with both new and old predictions. When feedback is present, control task observers are above chance with old predictions (p<.05) but not with new. By contrast, when pattern search observers are given feedback, they perform at chance level with new predictions, although they are still above chance with old predictions (p<.01).

General Discussion

We conclude from these results that feedback rather than action is the critical ingredient for instance learning. When feedback was provided control task observers were above chance level in old but not new situations. When there was no feedback, control task observers were at chance level in both old and new situations, indicating that learning had not taken place. Presumably, feedback affects the kinds of decisions that can be made while learning instances. We also conclude that feedback prevents rule learning. When informed of the model's goal, pattern search observers learned instances and not rules, as indicated by the difference between their performance in old and new situations. By contrast, when not informed of the model's goal, pattern search observers made correct predictions in both old and new situations, indicating successful rule learning.

These results suggest that learners may use either empirical learning or rule learning, depending on both their learning goal - whether to produce a specific response or discover the underlying pattern - and on whether or not they are given feedback about correct responses. In the concept learning literature, Erickson and Kruschke (in press) and Wisniewski and Medin (1995) have proposed models in which empirical learning and theory driven learning interact. Machine learning researchers have also developed systems that combine both empirical and explanation based learning (e.g. Lebowitz, 1986).

Our results pose a problem for these learning models: how to explain the influence of learning goal and feedback on the acquisition of instances on the one hand and rules on the other. In our experiment, the strongest evidence for a dissociation between instances and rules comes from the prediction questions. According to Wisniewski and Medin's interactive model, people learn instances when they have no prior knowledge to inform learning, otherwise they learn rules. However, in our study, we can assume that all subjects had roughly the same prior knowledge available to them. Subjects who had a pattern search goal and no feedback presumably used their prior knowledge of mathematics to help them generate and test hypotheses. But this prior knowledge was not used by subjects who had a specific goal or who had no feedback. In Erickson and Kruschke's (in press) model, every stimulus is processed simultaneously by the rule module and the exemplar module, with the final output being a combination of the outputs of these two modules. This model too, therefore, has difficulty accounting for the way attention is directed to either rule learning or instance learning as a function of learning goal and of the presence or absence of feedback.

The dual space models of Klahr and Dunbar (1988) and Simon and Lea (1974) give the best general framework for explaining our observations. In such models, learning can be directed to one or both problem spaces as a function of learning goal and feedback. A pattern search goal encourages learners to explore both rule space and problem space. By contrast, a control task goal to produce a specific output from the system and the presence of feedback direct the learner's attention to the goal state and how to reach it and so the learner focuses on instance space. In the absence of such direction, it is likely that relevant prior knowledge guides the learner to use rule space as well as instance space, as was observed Wisniewski and Medin (1995).

In conclusion, our findings suggest ways in which learners can be guided to learn more effectively, since we have shown that goal orientation and whether or not feedback is given can be tailored to the type of learning intended. The results also highlight the limitations of observational learning. Observation supports instance learning only when feedback is given, and it supports rule learning only when feedback is withheld.

Acknowledgments

Rosemary Stevenson acknowledges the Human Communication Research Centre, funded by the Economic and Social Research Council of Great Britain. Bruce Geddes was supported by a research studentship from the Engineering and Physical Sciences Research Council of Great Britain.

References

Berry, D. C. (1991). The role of action in implicit learning. *Quarterly Journal of Experimental Psychology, 43A*, 881-906.

Berry, D. C. & Broadbent, D. E. (1984). On the relationship between task performance and associated verbalisable knowledge. *The Quarterly Journal of Experimental Psychology, 36A*, 209-231.

Erickson, M.A. & Kruschke, J.K. (in press). Rules and exemplars in category learning. *Journal of Experimental Psychology: Learning, memory & Cognition.*

Geddes, B.W. & Stevenson, R.J. (1997). Explicit learning of a dynamic system with a non-salient pattern. *Quarterly Journal of Experimental Psychology,50A*, 742-765.

Klahr, D., & Dunbar, K. (1988). Dual space search during scientific reasoning. *Cognitive Science, 12*, 1-55.

Lebowitz, M. (1986). Integrated learning: Controlling explanation. *Cognitive Science, 10*, 219-240.

Owen E., & Sweller, J. (1985). What do students learn while solving mathematics problems? *Journal of Educational Psychology, 77*, 272-284.

Shanks, D. R., & St. John, M. F., (1994). Characteristics of dissociable human learning systems. *Behavioral and Brain Sciences*, 17, 367-447.

Vollmeyer, R., Burns, B.D., & Holyoak, K.J. (1996) The impact of goal specificity on strategy use and the acquisition of problem structure. *Cognitive Science, 20*, 75-100.

Whittlesea, B.W.A., & Dorken, M.D. (1993). Incidentally, things in general are particularly determined: An episodic-processing account of implicit learning. *Journal of Experimental Psychology: General, 122*, 227-248.

Wisniewski, E.J. & Medin, D.L. (1995). Harpoons and Long Sticks: The Interaction of Theory and Similarity in Rule Induction. In A. Ram & D.B. Leake (Eds.) *Goal-Driven Learning*. Cambridge, Mass., London England, MIT Press.