

## **UC Merced**

### **Proceedings of the Annual Meeting of the Cognitive Science Society**

#### **Title**

Practice Doesn't Always Make Perfect: Goal Induced Decrements in the Accuracy of Action- and Observation-based Problem Solving

#### **Permalink**

<https://escholarship.org/uc/item/86r4s0rr>

#### **Journal**

Proceedings of the Annual Meeting of the Cognitive Science Society, 27(27)

#### **ISSN**

1069-7977

#### **Authors**

Heyes, Cecilia  
Osman, Magda

#### **Publication Date**

2005

Peer reviewed

# Practice Doesn't Always Make Perfect: Goal Induced Decrements in the Accuracy of Action- and Observation-based Problem Solving

**Magda Osman (m.osman@ucl.ac.uk)**

Department of Psychology, University College London, Gower Street  
London, WC1E 6BT England

**Cecilia Heyes (c.heyes@ucl.ac.uk)**

Department of Psychology, University College London, Gower Street  
London, WC1E 6BT England

## Abstract

Recently, studies of problem solving within a dynamic environment (e.g. Burns & Vollmeyer, 2002; Vollmeyer, Burns, & Holyoak, 1996) show that performance is more accurate when people are engaged in a non-specific goal compared with a specific goal. We discuss findings from an experiment that was designed to replicate and extend the original work of Burns and Vollmeyer (2002). In Burns and Vollmeyer's study the learning phase of the problem solving task was action based, whereas in the present study participants either learnt to solve the task under observational or action conditions. Consistent with the original findings, this study shows that goal specificity affects the accuracy of problem solving in the same way when the learning phase of the task is observational rather action based. Additionally, the present study is an example in which there are no benefits of action over observational learning.

## Goal Specific and Non-specific Problem Solving

Early research showed that people typically solve problems by selective search through large problem spaces using heuristics (short cut strategies) such as means-ends analysis (Hammond, 1996; Newell & Simon, 1972). This is a method that decomposes the ultimate goal (end) of a problem into smaller intermediate goals, and at each stage sets out actions (means) which help to achieve these smaller goals.

Recent laboratory studies have found that people also find solutions to problems without directed conscious effort (Geddes & Stevenson, 1997; Burn & Vollmeyer, 2002). The limited research on this type of problem solving suggests that problems are also solved intuitively, by exploring the task without a specific goal in mind. This incidental approach to solving problems may be highly profitable, involving less effort and more efficiency than when the solver works toward a specific goal. The main difference between specific and non-specific problem solving is thought to lie in the type of information search that is carried out. In specific goal problem solving people only examine information that is relevant to the identified goal, at the cost of having a superficial understanding of the problem (Sweller, 1988). In non-specific goal problem solving the search is unrestricted, and so people gain a

broader and deeper understanding of the problem (Miller, Lehman, & Koedinger, 1999).

## Observational learning

In the real world, most problem solving occurs in a social context. Studies of observational learning have shown that problem solving is easier when tackled during or after we have observed others attempting the same problems, rather than in social isolation. This work has led to a distinction between two forms of observational learning: imitation and mimicry. Imitation is thought to be intrinsically goal directed (or intentional) (Bekkering, Wohlschlaeger, & Gattis, 2000), to give rise to explicit knowledge (Kelly & Burton, 2001), to be unique to humans (Tomasello, Kruger, & Ratner, 1993), and to be a highly efficient way of acquiring information. Mimicry, on the other hand, is thought to be automatic (or unintentional), to occur implicitly, to be an evolutionary older phenomenon (Roth & Leslie, 1998; Melzoff, & Moore, 1977; Tomasello, et al. 1993), and to be a relatively weak means of acquiring knowledge. However, the results of several studies are inconsistent with this picture. For example, there is now evidence that observational learning can be both highly efficient and implicit (Bird, Osman, Saggerson, & Heyes, in press), and that enhancing goal-directedness can have a detrimental effect on the efficiency of observational learning (Bird, Brindley & Heyes, in prep). These findings raise the possibility that goal specificity affects observation learning in much the same way as it does problem solving; observational learning is more efficient when the goal is undirected or non specific, compared to when learning is goal directed or oriented toward a specific goal.

## Present Study

Current empirical advances in the understanding of problem solving and observational learning suggest that the ways in which people store and access new knowledge depends on the specificity of the goal. These distinct research fields have found evidence that goal specificity affects the efficiency of problem solving and learning in the same way. That is, provided with a specific goal, learners who perform simple motor tasks or complex problem solving tasks are impaired relative to the performance of non-specific goal

learners. If this is correct, one would expect goal specificity to have similar effects on performance in a problem solving task when the learning stage is action-based and when it is observation-based. Thus far there has been no empirical investigation of the effects of observational learning in a complex problem solving task. We investigate this issue by using a complex problems solving task. We predict that problem solvers who learn to control a complex problem solving task without directly intervening, given sufficient motivation, should engage with the task to the extent that they will successfully solve it. Additionally, observational learners should demonstrate the exact same goal specificity effects that have been reported in problems solving environments in which the task is learnt through direct intervention. We propose that the reason for this is based on the supposition that the underlying properties of a problem solving task involve learning processes which are also required in simple learning tasks (Holland, Holyoak, Nisbett, & Thagard, 1989).

### Experimental Paradigm

The present study is based on the Water tank problem solving task devised by Burns and Vollmeyer's. In their task participants were required to control a linear system (Water tank system). This consisted of three inputs (substances: salt, carbon, lime) which were connected to three outputs (measures: oxygenation, chlorine concentration, temperature) (see Figure 1). For each input to output link there was added noise. For example, if, on trial 1 a participant had entered a value of 100 to the input lime then the value of oxygen would be 202, because the starting value of oxygen is 100 on trial 1, and the added noise further increases the value of this output.

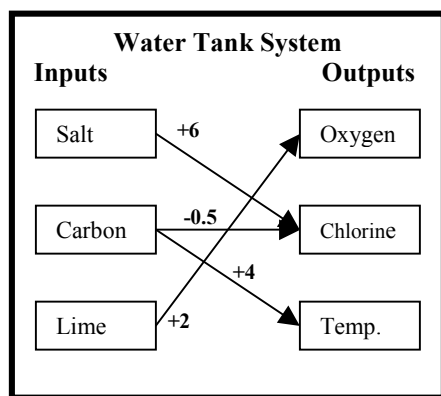


Figure 1: Water tank system

The task began with 12 learning trials described as the exploration phase. The starting values of the inputs were set to 0, and those of the outputs were: Oxygen = 100, Chlorine concentration = 500, Temperature = 1000. Participants would learn the causal structure of the system by changing the values of the inputs on each trial and learning which outputs were affected and by how much.

In the next phase (solution phase) participants demonstrated their skills at controlling the system by

attempting to attain and then maintain specific output values over the course of 6 trials.

The final phase (transfer phase) was exactly the same as the solution phase with the exception that participants were given a new set of output values to achieve.

The crucial manipulation in Burns and Vollmeyer's study was the instructions given to the two experimental groups. For the non-specific goal group, the task was presented as has just been described. Participants were simply told that they were to learn as well as they could the structure of the system in the exploration phase. In contrast, the specific goal group was told at the outset that they were to control the system in the exploration phase. Their goal was to reach a specific set of output values and to maintain them over the 12 trials. The values of the outputs were exactly the same as those that would later be examined in the solution phase.

Consistent with their predictions, Burns and Vollmeyer found that the non-specific group matched the performance of the specific group in the solution phase, but outperformed them in the transfer phase. Evidence from protocols confirmed that a non-specific goal improved learning because it increased hypothesis testing. The specific goal group used a limited form of hypothesis testing that ensured they gained enough knowledge of the task environment to attempt to reach their goal. They would switch between instance learning and minimal hypothesis testing which was a costly strategy for the reason that they only learned information relevant to performing a specific goal.

Burns and Vollmeyer proposed that without a specific goal, participants set an appropriate goal themselves by engaging in hypothesis testing. In turn, their understanding of the underlying causal structure is superior to the specific goal group because they search for rules and examine particular instances (input-outputs pairs). In contrast, the specific goal group are bounded by the goals they are trying to attain and this constrains their understanding of the task.

### Present Experiment

There were two main objectives of the present experiment. First, the aim was to replicate the effect reported by Burns and Vollmeyer that non-specific goal learning is superior to specific goal learning. Evidence from observational learning studies suggests that goal specificity effects performance in a similar way to that which has been reported in studies of problem solving. However, to our knowledge, no dedicated empirical work has examined the effects of goal specificity in an observation-based problem solving task. Therefore, the second objective was to investigate whether changing the exploration phase of Burn and Vollmeyer's problem solving task to an observational version yields the same effects of goal specificity detected in the action-based version.

In the following experiment there were four groups: *Non-specific goal* (action – participants would directly manipulate the values in all phases of the experiment), *Specific goal* (action), *Non-specific goal* (observation – participants would only observe values changing during the exploration phase, and could not directly manipulate the values themselves), *Specific goal* (observation).

We predicted that the groups given non-specific goal instructions would outperform groups that had been given specific goal instructions.

## Method

**Participants** Sixty-four students from University College London volunteered to take part in experiment and were paid £4 for their involvement. They were randomly allocated to one of the four conditions, with sixteen in each. Participants were tested individually.

**Design** The experiment was a 2x2 between subjects design with two levels of goal specificity (non-specific, specific) and two levels of learning (action, observation). All participants were presented with an exploration phase consisting of 12 trials; a solution phase consisting of 6 trials; and a transfer phase consisting of 6 trials. The main manipulation was the instructions that participants received before the exploration phase (non-specific, specific) and the effects of these were measured in the solution and transfer phases. The second manipulation was whether participants explored the problem solving environment by directly manipulating the input values themselves or by observing changes to inputs and outputs without setting the values themselves. In the latter case (non-specific observe, and specific-observe) because participants were not directly manipulating the input and output values on each trial during the exploration phase, the values and order in which they were changed was pre-determined. The actual values that were changed from trial to trial were based on a participant's responses from the *Specific goal* (action) group. The reason for this was twofold. The exact same input values in the exploration phase were presented to non-specific and specific observers in order to make direct comparisons, and to isolate the effects of goal specificity by reducing the differences between the groups to just the instructions. The input values of an above average specific action learner were used. It was thought that a high performer from this group would necessarily demonstrate a high level of rule and instance learning, and would give observers of both types a good opportunity to learn the problem solving environment.

**Procedure** Participants were told that they would be taking part in a problem solving task and they were to learn and later control a water tank system. They were also assured that throughout the task, were they to encounter any difficulties, they should ask the experimenter for assistance.

*Exploration phase:* All participants were presented with the same computer display which was similar to Figure 1 with the exception that the connections and added noise of each connection were not presented. At the beginning of the exploration phase the input values were set to 0. The starting values of the outputs were: Oxygen = 100, Chlorine concentration = 500, Temperature = 1000. In Burns and Vollmeyer's study participants were able to see the starting values of input and output values at the beginning of each

trial. In the present experiment participants were only shown the starting values of the inputs values and not the output values; the output values were only revealed on the first trial and not before.

In the action group participants would begin a trial by changing the value of an input using a slider on a scale ranging from -100 to 100. On each trial participants were free to change as many inputs as they liked, although everyone was advised that varying one input per trial was a useful strategy to adopt. Vollmeyer, et al. (1996) found that without this advice performance was highly impaired for specific and non-specific learners because many more sub-optimal strategies were used. Once they were satisfied with their changes to the inputs, participants would click on a button labeled 'output readings' and this would reveal the values of all three outputs. This procedure would complete a trial. On the next trial the input values entered would change the values of the outputs from the previous trial. This procedure continued for 6 trials. When it ended, participants were presented with a structure task; Burns and Vollmeyer devised the task to examine participants' knowledge of the causal structure of the environment. A diagram of the water system was shown on screen, and participants were asked to tick which of the options they took to be the causal links between a particular input and output (see Figure 2). They could also enter what they thought was the numerical relationship between inputs and outputs. When they had finished, they clicked a button to continue, and then completed the remaining six trials of the exploration phase, and another version of the structure task.

The non-specific group completed the exploration phase without having to achieve a particular goal. The specific goal group was told from the outset that they were to control the system by trying to reach and maintain the following output values: Oxygen = 50, Chlorine concentration = 700, Temperature = 900.

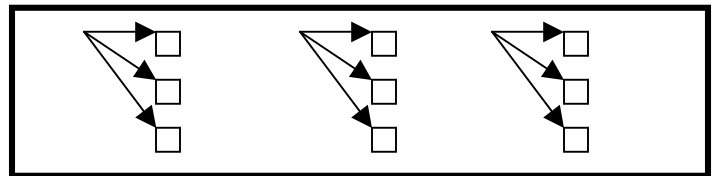


Figure 2: Structure Task

In the observation group the procedure was similar to the action group with two exceptions. First, instead of entering input values themselves, observers clicked a button to reveal the input values generated by a participant in one of the action groups. When they were ready they would then click a second button to reveal the output values for that trial. The second exception was that participants in the non-specific group were told to simply try and learn about the automated system that they were observing. The specific group were told to assess the effectiveness of the system's ability to achieve specific output values on each trials (i.e. Oxygen =

50, Chlorine concentration = 700, Temperature = 900). In all other respects, the action and observe groups did not differ, and the structure tasks were completed in the same way, as were the remaining phases of the experiment.

*Solution phase:* In this phase all participants were required to change the input values to achieve the following output values: Oxygen = 50, Chlorine concentration = 700, Temperature = 900. The specific groups had been previously exposed to these values, whereas the non-specific groups had not. Participants were allocated 6 trials in which they were to reach and then maintain the output values given.

*Transfer phase:* As in the solution phase, all participants were now required to change the input values to achieve given output values. However, the required output values were now : Oxygen = 250, Chlorine concentration = 350, Temperature = 1100. This was the first transfer test for the specific groups, because they had no prior experience of learning to set the values required in this phase.

### Scoring

Structure scores: A simple scoring scheme was used which computed the proportion of input-output links correctly identified. A correction for guessing was incorporated in the scoring and this was based on the same procedure used by Vollmeyer et al. (1996). The maximum value for each structure score was 1.

Solution error scores/ Transfer error scores: The procedure Burns and Vollmeyer used in scoring errors was adopted here. Success in achieving the goal states was computed as the sum of the absolute differences between each goal value and the value produced for each of the three outputs. All analyses of these scores are based on the mean solution error averaged over all six trials and averaged across all three output variables. A log transformation (base 10) was applied to the error scores of each individual participant for each round to minimize the skewness of the distribution of scores. Transfer error scores were calculated in the exact same way as the Solution error scores. In each case a lower score indicates better performance.

### Results

**Structure Scores:** Table 1 indicates that structure scores increased across all groups after the second round of 6 trials in the exploration phase. Additionally, the mean structure scores of both non-specific goal groups are higher than those of the specific goal groups, suggesting that the non-specific goal groups achieved better performance.

A 2x2x2 ANOVA was carried out with round (Structure Score 1, Structure Score 2) as the within subject factor, and goal type (non-specific, specific) and learning type (action, observation) as between subject factors. Confirming the trends indicated in Table 1, the analysis revealed significant main effects of round,  $F(1, 60) = 14.04, p < .0005$ , and of goal type  $F(1, 60) = 11.73, p < .001$ . Thus, the present experiment successfully replicated the findings originally

reported by Burns and Vollmeyer. Neither the main effect of learning type,  $F(1, 60) = .59, p = .45$ , nor any of the interactions, were significant. These results imply that regardless of the method by which participants learnt the problem solving environment, non-specific goal learners' knowledge of the structure was superior to specific goal learners.

Table 1: Overall Group Mean Structure Scores, Solution Error and Transfer Error Scores

Condition	Structure Score 1	Structure Score 2	Solution Error	Transfer Error
Non-specific Action	0.33	0.61	2.75	2.82
Specific Action	0.21	0.33	2.93	3.04
Non-specific Observe	0.38	0.48	2.77	2.85
Specific Observe	0.17	0.26	2.92	3.06

**Solution and Transfer Error Scores:** Table 1 also includes the overall mean solution and transfer error scores for each group. These figures indicate that the non-specific groups achieved lower solution and transfer error scores than the specific groups. In addition, they suggest that transfer error scores were higher than solution error scores. To analyze this a 2x2x2 ANOVA was conducted with phase (Solution, Transfer) as the within subject factor, and goal type (non-specific, specific) and learning type (action, observation) as between subject factors. The analysis revealed significant main effects of phase,  $F(1, 60) = 13.55, p < .0005$ , and of goal type,  $F(1, 60) = 40.33, p < .0005$ . As in the structure score analysis, neither the main effect of learning type, nor any of the interactions, were significant ( $F > 1$ ).

Burns and Vollmeyer reported a difference between the non-specific and the specific group in the transfer phase, but not in the solution phase. In contrast, the present study found superior performance by the non-specific group in both the solution and the transfer phases. Given that learning type was not significant, the error scores from both action and observation groups were combined. The tendency for non-specific groups to perform better than specific groups was confirmed using a one-way ANOVA with goal specificity as the independent factor, and Solution error scores as the dependent factor. This revealed a highly significant main effect,  $F(1, 63) = 17.38, p < .0005$ .

### Discussion

The evidence from this study can be summarized as follows: first, we successfully replicated Burn and Vollmeyer's findings which showed that goal specificity affects the accuracy of performance in a dynamic problem solving environment. Second, we demonstrated that the effects of goal specificity extend to a problem solving environment in which learning is performed under observational conditions. Third, both specific goal groups showed no advantage over

the non-specific goal groups in the solution phase, despite the fact that they had received 12 trials of prior training in the exploration phase. Rather, the non-specific goal groups outperformed the specific goal groups in both test phases.

The findings raise two questions: Why were there no differences between action and observation learning conditions? And, why was prior training relatively ineffective for specific goal groups?

Action and observation learning have been contrasted in simple SRT (serial reaction time) tasks (e.g., Bird & Heyes, in press; Kelly & Burton, 2001; Kelly, Burton, Riedel, & Lynch, 2003); in causal learning (e.g., Lagnado & Sloman, 2004); and problem solving (e.g., Berry, 1991; Lee, 1995). Although the evidence from these studies is mixed, the lack of difference between action and observation is still held as a controversial view. It is important to elaborate at this stage that although the action vs. observation contrast is often translated into a comparison between passive and active learning, we would like to propose that this translation is not relevant in the present experiment.

There are properties of the task that indicate why action and observation learning conditions did not differentially affect error scores in the later stages of the experiment. Observers were required to engage with the task throughout the exploration phase and this involved processes that were more similar than different to the action based learning condition. That is, they were motivated to examine the information on screen, particularly because during the first phase of the experiment changes to input values on a given trial were not accompanied by the subsequent output values until participants pressed a button to reveal them. The separation of input values from output values encouraged participants to actively process input information before output information on each trial during the exploration phase. For example, in the specific goal group participants were required to examine the outputs to check the extent to which they differed from the actual values that the system had to maintain on each trial. This process involved some mental calculation in order to track the differences between the values set and those achieved. The non-specific goal observers were required to learn the environment because they would later have to control it. As observers they would have had to identify which aspects of the task to focus on, i.e. which inputs were changed, was there a pattern to the change in inputs, what values were entered for each input and was there a pattern to this, which outputs changes, what were the changes to the actual output values. Thus, both types of observers were required to do more than passively watch the changes in input and output values.

It could be argued that observers may have responded poorly, but because the input values they received on each trial were from a high performing participant, this increased their performance. The rationale for yoking participants to the values of a member of the specific goal action group has been presented earlier in this article. Although we cannot eliminate the possibility that participants in the observe group would have performed poorly otherwise, the evidence from this study suggests that there was no difference between action and observation groups according to structure tests and error scores. Moreover, the main effect of

goal specificity still clearly emerged in the observer group even though the actual input values were from a participant from the specific action group, and despite both observers being presented exactly the same trial information. This is particularly noteworthy since it suggests that there is nothing particular to the actual input and output values themselves, or the number of input values changed on each trial that influences the goal specificity effect, only the instructions.

The second issue raised in this study was that both specific groups behaved in the same way, in that their performance was poorer than the non-specific group. This was most evident in the solution phase in which the specific group showed no advantage in performance in the solution phase compared to the non-specific group despite receiving 12 trials of prior learning in the exploration. In contrast, Burns and Vollmeyer found that there was no difference in error scores in the solution between specific and non-specific goal groups. Burn and Vollmeyer (Experiment 1) report that the mean solution error scores for the non-specific group was 2.64 (sd. 1.40) and for the specific group was 2.40 (sd. 2.10). In the present study the mean solution error scores for action and observers were higher than the mean values reported by Burns and Vollmeyer. Although nothing can be directly inferred from this comparison, it does give some indication that, overall, participants in Burns and Vollmeyer's study performed more accurately than in the present study. Of course further work would be needed to examine if the differences in performance between studies were the result of changes in difficulty or other factors.

One difference between Burns and Vollmeyer's study and the present experiment that can account for this difference is that of task difficulty. In order to make the task harder, in the present study participants were not given the starting values of the outputs at the beginning of each phase. Thus participants were forced to pay close attention to the change in output values from the first trial, than if they were given them to begin with. Furthermore, because they were not reminded of the starting values it would have been harder to track the changes in outputs across the trials unless they accurately recalled the change in input as well as output values. The non-specific group may have increased their knowledge of both instances and rules because they were required to work harder to recall and track changes in the system. In contrast the specific group was impaired by the added cognitive load. Although they were primarily engaged in instance learning, they were given a specific goal which concerned the output values, the effort required in recalling and tracking changes to these values detracted even more from any resources that otherwise would have been directed at learning the relationship between the inputs and outputs.

## Conclusions

The main objective of this study was to replicate and extend the findings reported by Burns and Vollmeyer. The present study was successful in providing further evidence that the specificity of the goal has a marked affect on the accuracy

of performance in a dynamic problem solving task. The innovation in this study was to investigate whether the effects would be found in the same problem solving task when the learning phase was observation based. We demonstrated that performance in action and observation conditions did not differ and that goal specificity was the most reliable factor in determining differences in accuracy of problem solving.

### Acknowledgments

Preparation for this article was supported by Economic and Social Research Council ESRC grant RES-000-27-0119. The support of the Economic and Social Research Council (ESRC) is gratefully acknowledged. The work was also part of the programme of the ESRC Research Centre For Economic Learning and Human Evolution.

### References

- Bekkering, H., Wohlschlagel, A., & Gattis, M. (2000). Imitation of gestures in children is goal-directed. *Quarterly Journal of Experimental Psychology*, *53*, 153-164.
- Berry, D. (1991). The role of action in implicit learning. *Quarterly Journal of Experimental Psychology*, *43*, 881-906.
- Bird, G., Brindley, R., & Heyes, C. (in preparation). The effects of goal specificity on observational sequence learning.
- Bird, G., & Heyes, C. (in press) Effector-dependent learning by observation of a finger movement sequence. *Journal of Experimental Psychology: Human Perception and Performance*.
- Bird, G., Osman, M., Saggerson, A., & Heyes, C. (in press). Implicit learning by observation of animate, but not inanimate, sequential stimuli.
- Burns, B. D., & Vollmeyer, R. (2002). Goal specificity effects on hypothesis testing in problem solving. *Quarterly Journal of Experimental Psychology*, *55*, 241-261.
- 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.
- Hammond, K. R. (1996). *Human judgment and social policy*. New York: Oxford Press.
- Kelly, S., & Burton, A. M. (2001). Learning complex sequences: No role for observation. *Psychological Research*, *65*, 15-23.
- Holland, J. H., Holyoak, K. J., Nisbett, R. E., & Thagard, P. R. (1989). *Induction: processes of inference, learning and discovery*. Cambridge, MA: MIT Press.
- Kelly, S., & Burton, A. (2001). Learning complex sequences: No role for observation? *Psychological Research*, *65*, 15-23.
- Kelly, S., Burton, A., Riedel, B., & Lynch, E. (2003) Sequence learning by action and observation: Evidence for separate mechanisms. *British Journal of Psychology*, *94*, 355-372.
- Lagnado, D., & Sloman, S. (2004). The advantage of timely intervention. *Journal of Experimental Psychology: Learning, Memory and Cognition*, *30*, 856-876.
- Lee, Y. (1995). Effects of learning contexts on implicit and explicit learning. *Memory and Cognition*, *23*, 723-734.
- Melzoff, A. N., & Moore, M. K. (1977). Imitation of facial and manual gestures by human neonates. *Science*, *198*, 75-78.
- Miller, C.S., Lehman, J. F., & Koedinger, K. R. (1999). Goals and learning in microworlds. *Cognitive Science*, *23*, 305-336.
- Newell, A., & Simon, H. A. (1972). *Human problem solving*. Englewood Cliffs, NJ: Prentice Hall.
- Roth, D., & Leslie, A. M. (1998). Solving belief problems: Towards a task analysis. *Cognition*, *66*, 1-31.
- Sweller, J. (1988). Cognitive load during problem solving: Effects of learning. *Cognitive Science*, *12*, 257-285.
- Tomasello, M., Kruger, A. C., & Ratner, H. H. (1993). Cultural learning. *Behavioral and Brain Sciences*, *16*, 495-511.
- Vollmeyer, R., Burns, B. D., & Holyoak, K. J. (1996). The impact of goal specificity and systematicity of strategies on the acquisition of problem structure. *Cognitive Science*, *20*, 75-100.