

# Learning by imitation, by reinforcement and by verbal rules in problem solving

Frédéric Dandurand, Melissa Bowen, Thomas Shultz  
McGill University, Department of Psychology

## Abstract

Learning by imitation is a powerful process for acquiring new knowledge, but there has been little research exploring imitation's potential service to the problem-solving domain. Classical problem-solving techniques tend to center around reinforcement learning, which requires significant trial-and-error learning to reach successful goals and problem solutions. Heuristics, hints, and reasoning by analogy have been favoured as improvements over reinforcement learning, whereas imitation learning has been regarded as rote memorizing. However, research on imitation learning in animals and infants suggests that what is being learned is the overall arrangement of actions (sequencing and planning) (Byrne, Russon, 1998). Applied to problem solving, this suggests that imitation learning might enable a problem solver to infer a complex hierarchical problem representation from observation alone.

We compared three types of learning in problem solving tasks: imitation learning (a group that viewed successful problem solving demonstrations), reinforcement learning (a group that got feedback indicating whether their answer was correct or not) and explicit learning (a group that was presented specific instructions to solve the problem). On a task consisting in finding, with 3 uses of a scale, the one ball, which is either heavier or lighter in the set of 12 balls, we found that subjects in the imitation learning and explicit learning groups outperformed those in the reinforcement learning group. We conclude that learning by imitation in problem solving tasks is worthwhile, efficient and even superior to explicit learning because of the minimal time and energy investment required from the mentor.

## 1. Introduction

People are constantly testing an assortment of actions and reactions to problems and novel situations in their lives. This process of trial-and-error, which provides feedback by rewarding successful tactics and punishing failed strategies, may not be paramount. With problem-solving being such a central aspect of life, researchers continue to search for superior methods. Psychologists and biologists alike have been attracted to the concept of social learning for generations. Imitation in particular, as a mechanism for learning, has generated a significant amount of research. It is widely accepted that humans are innately imitative creatures in some sense of the word, but what is more controversial is the definition of imitation and what is accepted as "true imitation". Many definitions that researchers have offered for the concept of imitation are vague, and at present, no one interpretation has been agreed upon. To this point few researchers have explored both learning by imitation and problem solving; thus this project aims to bridge the two disciplines. Classical problem-solving methods discussed in the literature rely on learning through reinforcement and feedback or explicit verbal instructions, but imitation learning has not been addressed in the problem-solving domain. We intend to show the value of learning by imitation in problem-solving.

## 2. Problem-Solving

"Diverse cognitive abilities such as perception, language, sequencing of actions, memory, categorization, judgment, and choice all play important roles in human problem solving." (Holyoak, 1995) Problems arise when the path to a goal is not immediately obvious. Newell and Simon (1972) first introduced the idea of a problem as a spatial conception and a search through space for a solution (Holyoak, 1995). These theorists described problems in terms of initial states, goal states, and operators or actions to take to move the current state closer to the goal state. Moreover, path constraints are thought to further limit the route to a successful

solution (Holyoak, 1995). A problem solution is defined by a sequence of operators that transforms the initial state into the goal state. Problem-solving methods tend to involve problem decomposition, i.e. breaking the problem down into separate sub-goals, and then planning action sequences to reduce error (Holyoak, 1995). Information processing theorists believe that cognitive development is continuous, and that human problem-solving strategies are a result of their experiences. Prior problem exposure and feedback from each problem-solving experience are fundamental to strategy development.

Nevertheless it is important to note that different problems necessitate different strategies. One important way to categorize problems is to decide whether they are well-defined or ill-defined. Well-defined problems must have clearly defined start states and clearly defined goals, as well as unambiguous and explicit actions and constraints (Best, 1995). An example of a well-defined problem with a hierarchy of sub-goals is the Tower of Hanoi problem. The aim, with three pegs available, is to move a tower of disks from one peg to another, moving only one disk at a time (Medin and Ross, 1992). To solve this problem, one must keep track of the current state of the problem, and perform a series of transformations to satisfy each sub-goal in order to finish in the final goal state. The challenge posed by the problem is how to select the optimal moves between states from the available options. This way of learning how to solve the problem demands a great deal of trial-and-error learning to achieve success.

Problem solvers approach each novel problem with a general strategy as well as any prior information they possess about similar problems. A common problem-solving strategy called means-end analysis attempts to reduce the discrepancy between the current state and the goal state by selecting the best possible next step (Glass and Holyoak, 1986). When the desired action cannot yet be employed, the problem solver must first set a new sub-goal to reach the state where he can take action (Holyoak, 1995). Another basic problem-solving approach is the generate-test method, which involves the creation of a solution strategy and then implementation of the strategy to determine its effectiveness (Newell and Simon, 1972). For these sorts of classical problem-solving methods, assessment of the strategy relies solely on feedback from the problem in terms of success or failure to solve it. Problem solvers also tend to use prior knowledge when approaching new problems. This knowledge is often helpful, but it can become a hindrance to problem solvers especially if the solution path to the current problem is counterintuitive. These practical strategies, or heuristics, can block the generation of new strategies if one becomes “functionally fixed”, or stuck in one strategy (Benjafield, 1997).

### **3. Types of Learning**

#### *Reinforcement Learning*

The basic concept of learning by reinforcement dates back to Edward Thorndike and subsequently B.F. Skinner and their ideas of learning associations and contingencies between stimuli and responses. While these behaviorists thought about learning in terms of responses to certain stimuli, cognitive scientists represent reinforcement learning as responses generated by information processing procedures that examine the given inputs (Leahey and Harris, 1997). Correct responses are rewarded with successful problem solutions and incorrect responses are punished with solution failures. What is agreed upon is that reinforcement learning occurs through

an interaction with one's environment. Each action the learner performs will lead to a unique learning experience (Dautenhahn, Nehaniv, & Alissandrakis, 2003).

Instrumental conditioning, introduced by Thorndike, began the investigation of trial-and-error learning (Leahey and Harris, 1997). By observing cats placed into what he called a "puzzle box", Thorndike observed that the animals would try a variety of different escape methods. After struggling in repeated test sessions in the "puzzle box", the animals began to learn the successful escape behaviors and the unsuccessful approaches ceased (Leahey and Harris, 1997). Intuitively, this method of gaining knowledge is in line with human behavior in novel situations: we test an assortment of distinct reactions, find one that works, and then apply that one first to the next set of similar circumstances (Leahey and Harris, 1997).

This classical learning process, still thought of by many as the dominant learning approach, provides a restricted amount of information to a learner especially in a problem-solving domain. Such learning relies solely on feedback from the problem in terms of success or failure to solve it without offering further information on how to succeed. This way of learning how to solve a problem demands a considerable amount of trial-and-error learning to achieve success. Formulation of and assessment of problem-solving strategies are based completely on the reinforcement of the success of finding a solution. The learner attempts to devise certain rules or heuristics about managing certain stimuli or solving particular problems, but the details of their strategies are not evaluated, only their end results. Consequently, their rules tend to be sub-optimal. This is a direct result of the limited information their successes and failures afford.

### *Explicit Learning*

Often in IF/THEN format, conditional rules provide clear direction for specific situations. Rules aid the problem solver in the understanding and the concept formation of the problem that needs to be solved (Solso, 1995). In contrast to learning by imitation, which provides implicit instruction, or reinforcement learning, which offers feedback only, rules provide explicit instruction on how to act.

Explicit learning is limited in scope. First, it assumes the availability of a skilled teacher who has the time, energy and ability to express problem solving reasoning explicitly, concisely, completely and coherently. Second, broad classes of problems, such as information-integration category-learning tasks (Ashby & Ell, 2001), are highly problematic for explicit instructions. Even skilled problem solvers are unable to express their problem solving strategies explicitly because solution strategies for these problems are learned and accessed implicitly.

### *Learning by Imitation*

Many different researchers have described imitation, but a single definition has yet to be selected. In 1898, Thorndike defined imitation as any situation in which animals "from an act witnessed learn to do an act". Learning is central to his definition. Later in 1963, Thorpe defined "true imitation as the copying of a novel or otherwise improbable act or utterance, or some act for which there is clearly no instinctive tendency". The controversy remains over whether learning, which implies something new, can be equated to novelty. Thorpe would not deem learning a new sequence of previously established behaviors to be novel, despite the fact that the arrangement would be newly learned. For this reason, Thorpe's definition is

frequently considered too restrictive since it excludes any copied behaviors or movements, which were known (either innate or learned)( Dautenhahn, Nehaniv, & Alissandrakis, 2003).

More recently, Call and Carpenter (2002) developed a broader explanation, in order to differentiate between the many terms found in the imitation literature, by classifying imitation in terms of the result of the action, the goal of the action and the action itself. This organization clarifies terms such as ‘emulation’ (replicating the result but with a different action), ‘goal emulation’ (awareness of the goal and attempting to duplicate the result with any action), ‘mimicry’ (reproducing the action regardless of result), ‘imitation’ (comprehension of the goal with correct recreation of the action and result). With these distinctions it is clear that these authors believe that judgments made on whether or not imitation is deemed efficacious vary depending on the criteria.

Much data has been gathered to support the belief that human babies are born with the distinct capacity to imitate many behaviors they observe executed by adults in their surroundings. Infants even engage in self-correction to move closer to the model or target behavior. Such imitation of the physical and social world holds developmental value: “According to Meltzoff and Moore (Meltzoff and Moore 1992) imitative reciprocal interaction games between infants and adults bootstrap social cognition and provide the foundation of mature folk psychology, i.e. our understanding of other people.” (Dautenhahn, Nehaniv, & Alissandrakis, 2003)

Jacqueline Nadel et al.(1999) also note the interactive format of imitation and its contribution to communicative development, as it fosters pragmatic communication .

Michael Tomasello (1999) points to the importance of understanding the intentions of the demonstrator to produce “true imitation”. In a study by Tomasello, Carpenter and Akhtar (1998) infants witnessed adults performing several intentional and several unintentional actions vocally marked by “There!” and “Whoops!” respectively. Results showed that the intentional acts were two times more likely to be reproduced, or imitated, by the infants, signifying that they chose to copy not merely the superficial actions of the adults, but their deliberate acts. This sense of impersonation emphasizes the replication of each detail of behavior (Tomasello, 1990).

As an alternative, Byrne and Russon (1998) introduced ideas of hierarchical levels of imitation. They suggested that it is not necessary for humans and animals to mimic the fine details of the displayed actions, although such behavior would be considered imitation at the action level. Instead they proposed a specific yet more flexible approach called program level imitation, which focuses on reproduction of the overall arrangement of actions, particularly the planning of and sequencing of actions. According to these authors the learning of a new pattern of behavioral units qualifies as novelty. For example, they proposed that gorillas can learn a specific eating pattern by watching and imitating steps throughout a hierarchy, where each step corresponds to a specific goal. These researchers explain that a hierarchical problem conceptualization allows for imitation to occur on many levels. Although they concede that the details of how to meet each sub-goal may be attained by individual learning, they declare that reproduction of the overall structure is imitation. This definition, like Tomasello’s, implies that the imitators understand the goals of the observed actions and the intentions of the demonstrator.

The notable difference between learning by imitation and learning by reinforcement is the quality of information supplied. Imitation learning offers visible target behaviors in the form of demonstrated steps to reach various goals. This information need merely to be understood and transferred into motor behavior. Clearly this “how to” information is significantly more useful than the binary information provided by reinforcement learning (correct/incorrect) (Shultz, 2003).

#### **4. Combining Imitation and Problem-Solving**

We believe that imitation learning has much value that has not been sufficiently explored in the problem-solving domain. Considering the previous discussion of Byron and Russon’s theory of hierarchical goals involved in imitation, a well-defined problem is an appealing prototype to introduce the union of the fields of problem-solving and learning by imitation. In this study a computer based variant of the well known class of ball-weighing problems was created to investigate learning by imitation. Another variant of this problem that has been used in psychological research is Simmel’s Coin Problem (1953), described by Benjafield (1997) as follows: “Suppose you have eight coins and a balance. One of the coins is counterfeit, and therefore is lighter than the others. How can you find the counterfeit coin by using the balance only twice?” This class of problems consists of well-defined problems that require a hierarchy of transition states to reach the final goal state.

In this study, three groups of human participants were asked to solve a well-defined computer based ball-weighing problem. One group watched five successful demonstrations of solving the problem and had no feedback in terms of whether their answers were correct or not (imitation learning group). Another group was presented with a set of abstract verbal rules on how to solve the problem and also had no feedback (explicit learning group). All participants in these two groups were explicitly told that the demonstrations and verbal rules would help them solve the problems. Thus although during the problem-solving session the participants did not get correct/incorrect feedback, we can assume that they expected to find correct solutions by following the demonstrated steps, or verbal rules. The final group was asked to solve the problem with no demonstrations or rules, but feedback was provided (reinforcement learning group). Since the quality of information provided by each learning mechanism varies, we hypothesized that the group with access to the demonstrations (imitation learning group) should outperform the explicit learning group, which should in turn outperform the reinforcement learning group. We expected to show that compared to the richness of learning by imitation and the explicit instruction of verbal rules, learning by reinforcement is limited. In light of the debate concerning the underlying mechanism of imitation, the hypothesis was sufficiently flexible to accommodate any possible underlying mechanism of imitation.

#### **5. Methods**

##### *Participants*

McGill undergraduate and graduate students were recruited to participate in the computer based problem-solving task. The 70 participants tested yielded 63 sets of usable data (17 males and 46 females). Participants were discarded for being statistical outliers and for not completing the warm-up task. An incentive of a \$50 prize to one of the top five performers in each group was offered to encourage

maximal performance of all participants. Participants were randomly assigned to one of the three groups, and written consent was obtained from each.

### *Materials*

A two-part (level 1 and 2) computer program was created to run the ball-weighting problem. The ball problem variant presented in this experiment had 3 identical looking balls in level 1 and 12 balls in level 2, one of which was heavier or lighter than the rest at a common weight that we called “normal”. Use of the scale was limited to twice in level 1 and three times in level 2. The computer screen showed a scale, a ball container with the balls to be weighed, and a color selecting tool for marking the balls with different colors to keep track of their suspected weights. Seven different colors were available for marking, representing the following seven categories: normal weight, light weight, heavy weight, heavy or normal weight, light or normal weight, heavy or light weight, and unknown. Displayed at the bottom of the screen were the weight button to activate and deactivate the scale, the answer button and two text areas showing the number of scale trials allowed and the number used. The participants were also able to see how much time was remaining as well as the number of problems they had completed (See *Figure 1*).

### *Procedure*

Each participant was tested individually in an experimental session that lasted approximately 1 hour. Prior to their introduction to the user interface, all participants were presented with a thorough problem description including a diagram of the screen. The experimenters reminded participants of a few key points in the directions and encouraged them to ask questions at any point during the experiment.

All participants began in level 1 with no demonstrations. Level 1, a simple 3-ball variation of the problem, helped familiarize the participants with the software and the task, but the data collected in level 1 were not used in the statistical analysis of the experiment. After completing three successful trials in level 1, participants moved on to level 2 where they were limited to three uses of the scale. Each participant in the imitation learning group saw a random sample of 5 out of 24 possible demonstrations of successful trials before entering level 2. A complete trial was defined by the time from presentation of a new problem until the participant clicked on the answer button. Each trial required the participant to find the new target ball and to identify whether it was heavier or lighter than the other 11. Each participant in the explicit learning group was presented with abstract verbal rules on how to solve the problem before level 2 and was allowed to study these rules for up to 10 minutes, which was the time it took for all 5 demonstrations to run. In contrast, each participant in the reinforcement learning group began level 2 immediately after level 1. Each participant, regardless of group assignment, had 30 minutes to solve different problems in level 2. The elapsed time (in seconds) as well as the number of trials, and the number of correct or incorrect answers were all recorded.

It is important to point out that there is no single way to solve each of the possible problems of the 24 (12 balls x 2 states {heavy or light}) different ball/weight combinations, because each solution depends on the use of the scale. The ultimate goal of this problem is to identify the target ball, which is either heavier or lighter. To do this, the problem solver must loop through a pair of sub-goals each time he uses the scale, namely: to select which balls to weigh in order to maximize the

information from the scale, and to appropriately update the state of the balls with the color markings using the information from the weighing. Each subsequent approach towards using the scale will vary depending on what the problem solver has learned about the ball states from the previous weighings.

In the initial state of the problem all 12 balls are unknown. The problem solver must then proceed to select which balls and how many balls to include in the first weighing. The information from the first weighing must be used to update the states of each ball. For example, if the scale moves down on the left side, then the problem solver concludes that the balls on the left side of the scale should be marked heavy or normal, the balls on the right side of the scale should be marked light or normal, and the balls remaining in the ball bank should be marked normal. The next step toward problem solution involves assessing the current state of the problem and repeating the sub-goals of ball selection and subsequent ball marking. These sub-goals are repeated until the problem solver finds a solution or has no further uses of the scale. An ideal problem solver would follow the steps shown in Figure 2 to reach each solution.

Successful solutions to the ball-weighing problem require the use of some strategies that may be counterintuitive. The demonstration viewed by the imitation group provided implicit information to the participants, not explicit strategy instruction. We do not know and cannot control the amount of information that the imitation learning group obtains from the demonstrations; it is this that we are interested in. In this experiment the participants in the explicit learning group were provided with a subset of the complete set of rules required to solve all of the 24 problems. That is, each participant in this group studied a randomly generated set of rules that was essentially equal to the amount of information presented to each participant in the imitation learning group through the random set of 5 demonstrations. (See Figure 2 for a complete set of rules.) The reinforcement group did not have an absence of information, but simply different information based on the feedback from the program (correct or incorrect answer).

### *Analysis & Results*

Performance of each group ( $n=21$  per group) was determined by the two dependent variables: the average correct answer rate for the trials, and average elapsed time for trials. To normalize the elapsed time data, we used a log transformation. Given that each participant completed a different number of trials, we were presented with the challenge of analyzing unbalanced data. To address this issue, we devised a strategy that allowed us to compare the data despite the disparate number of trials. By creating a bin system, determined by the participant with the fewest trials completed (4 trials), we were able to divide each participant's number of trials completed into quartiles with each holding 25% of the trials completed.

An ANOVA was performed on log of elapsed time per trial across the independent variables group (3 levels - imitation, explicit learning and reinforcement) and quartile of completed trials (4 levels). The significant interaction effect of group and quartile ( $F(6, 180) = 2.63, p = .018$ ) indicated that the groups differ on their decrease in elapsed time across quartile, which suggests different rates of learning. The main effect of quartile ( $F(3, 180) = 85.7, p < .001$ ) demonstrates that all groups sped up over time. There was no significant main effect of group on elapsed time per trial ( $F(2, 60) = 1.42, p = .25$ ), that is no one group was significantly faster than any of the others.

To address the non-normally distributed answer correct rate data, we used a Kruskal-Wallis nonparametric test, which yielded significant differences in answer correct rate across group ( $\chi^2 = 7.054$ ,  $p = .029$ ). Pair-wise Kruskal-Wallis tests revealed significant differences between the imitation learning group and the reinforcement learning group, as well as between the explicit learning group and the reinforcement learning group ( $p = .021$ ,  $p = .029$  respectively). No significant differences were found between the imitation and explicit learning groups ( $p = .512$ ). Furthermore, 2 Chi-square tests of contingency were performed on the answer correct data. The first compared the 3 groups on the number of perfect performances (100% correct rate), while the second compared the 3 groups on the overall number of correct and incorrect answers. Both showed significant effects of group ( $\chi^2 = 6.107$  and  $\chi^2 = 21.08$  respectively,  $p < .05$ ). These data indicate that both the imitation learning group and the explicit learning group outperformed the reinforcement learning group by answering more quickly and more accurately. No significant differences were uncovered between the imitation learning group and the verbal rule instructions group.

## 6. Discussion

Our project was designed to demonstrate the value of imitation learning in the problem-solving domain. We predicted that participants who had the opportunity to learn by imitation would perform more accurately and more quickly (i.e. highest correct answer rate and lowest elapsed time per trial) on the well-defined ball-weighting problem than the explicit learning group and reinforcement learning group. We anticipated that the quality of information provided by demonstrations would be more useful to the participants than the possibly confusing verbal rules and the restricted information from the feedback (correct/ incorrect). In the reinforcement group, participants were obliged to rely only on the feedback of the program to tailor their problem-solving techniques. We expected to illustrate the limitations of reinforcement learning despite its influential position as a learning theory.

The data showed that both the imitation learning group and the explicit learning group outperformed the reinforcement learning group even without feedback information. There were no significant differences between the imitation learning group and the explicit learning group. In light of these findings, it is important to note that learning by imitation is more valuable than some might think, because it “teaches” at a much lower cost to the mentor than learning by verbal rules. Whereas the production and presentation of abstract rules requires some work of the mentor, in imitation learning the mentor need not even be aware of being observed and imitated. We have shown that conditional rules (if/ then format) can help improve problem solving performance, but it is crucial to note that there is variability in rule writing. It is difficult to compose concise, coherent and complete instructions. Abstract rules which are ambiguous could be hard to absorb, making the information less useful and rendering the task more difficult. Additionally, a poor understanding of the abstract verbal rules could cause them to be incorrectly implemented. This experiment shows that learning by imitation in problem solving tasks is advantageous, efficient and even superior to learning through explicit verbal rule instruction because of the minimal time and energy investment required from the mentor.

Now that we have established the difference in performance between learning by imitation and classical problem-solving techniques, it may be interesting to carry out follow up experiments to discern the underlying mechanism of imitation, which

accounts for the increased performance. Imitation learning might enable a problem solver to infer a complex hierarchical problem representation from observation alone. To begin to separate the different imitation effects, a group that views successful demonstrations of a problem using a different number of balls than in the target problem (i.e. 9 vs. 12) could be added. It would show whether the subjects memorize steps such as weigh 3 balls on each side of the scale, or are primed to make generalizations like weigh 1/3 of the balls. A group who sees a distractor task between the demonstrations and the trials could be added to disrupt possible memorization.

To improve the real life applicability of this research, learning by imitation should be studied further on problems that are less well-defined. For now we can conclude that imitation learning improves problem-solving performance even without reinforcement feedback.

### ***References***

- Ashby, G. & Ell, S. (2001) The neurobiology of human category learning. *Trends in Cognitive Sciences*, 5, 204-210.
- Benjafield, J. (1997). *Cognition* 2nd Edition. Upper Saddle River, New Jersey: Prentice Hall.
- Best, J. (1995). *Cognitive Psychology* 4th Edition. St. Paul, Minnesota: West Publishing Company.
- Breazeal, C. & Scassellati, B. (2002). Robots that Imitate Humans. *TRENDS in Cognitive Sciences*, 6, 481-487.
- Byrne, R. W., & Russon, A. E. (1998). Learning by Imitation: A hierarchical approach. *Behavioral and Brain Sciences*, 21, 667-721.
- Call, J. & Carpenter, M. (2002). Three Sources of Information in Social Learning. In Dautenhahn, K. and Nhaniv, C. L. *Imitation in Animals and Artifacts*. Cambridge, Massachusetts: MIT Press.
- Dautenhahn, K, Nahaniv, C. L., & Alissandrakis, A. (2003). Learning from Others-Social Learning and Imitation in Animals and Robots. In Kuhn, R., Menzel, R., Menzel, W., Ratsch, U., Richter, M. M., & Stamatescu, I.-O., *Adaptivity and Learning - An interdisciplinary debate*. Berlin: Springer.
- Glass, A. & Holyoak, K. J. (1986). *Cognition* 2nd Edition. New York, New York: Random House.
- Holyoak, K. J. (1995). Problem Solving. In E. E. Smith and D. N. Osherson, *An Invitation to Cognitive Science: Thinking*. Vol. 3 2nd Edition. Cambridge, Massachusetts: The MIT Press.
- Leahey, T. H., & Harris, R. J. (1997). *Learning and Cognition* 4th Edition. Upper Saddle River, New Jersey: Prentice Hall.
- Medin, D. & Ross, B. (1992). *Cognitive Psychology*. Orlando, Florida: Harcourt Brace Jovanovich, Inc.

Nadel, J., Guerini, C., Peze, A. and Rivet, C. (1999). The Evolving Nature of Imitation as a Format of Communication. In Nadel, J. & Butterworth, G. Imitation in Infancy.

Cambridge, England: Cambridge University Press.

Newell, A. & Simon, H. A. (1972). Human Problem Solving. Englewood Cliffs, New Jersey: Prentice Hall.

Shultz, T. (2003). Computational Developmental Psychology. Cambridge, Massachusetts: MIT Press

Solso, R. (1995). Cognitive Psychology 4th Edition. Needham Heights, Massachusetts: A Simon & Schuster Company.

Tomasello, M. (1999). The Cultural Origins of Human Cognition. Cambridge, Massachusetts: Harvard University Press.

### Tables

Group	Correct Rate	Std Deviation
Reinforcement learning	0.59	0.49
Imitation learning	0.76	0.42
Verbal Rule Instruction	0.71	0.45

*Table 2*

Average elapsed time (ms per trial) per group

Group	Elapsed time	Std Deviation
Reinforcement learning	104 237	79 020
Imitation learning	87 990	51 068
Verbal Rule Instruction	94 545	61 180

*Table 3*

Contingency table of the number of perfect and imperfect performers in each group

Group	Participants	Perfect Score	Imperfect Score
Reinforcement learning	21	0	21
Imitation learning	21	5	16

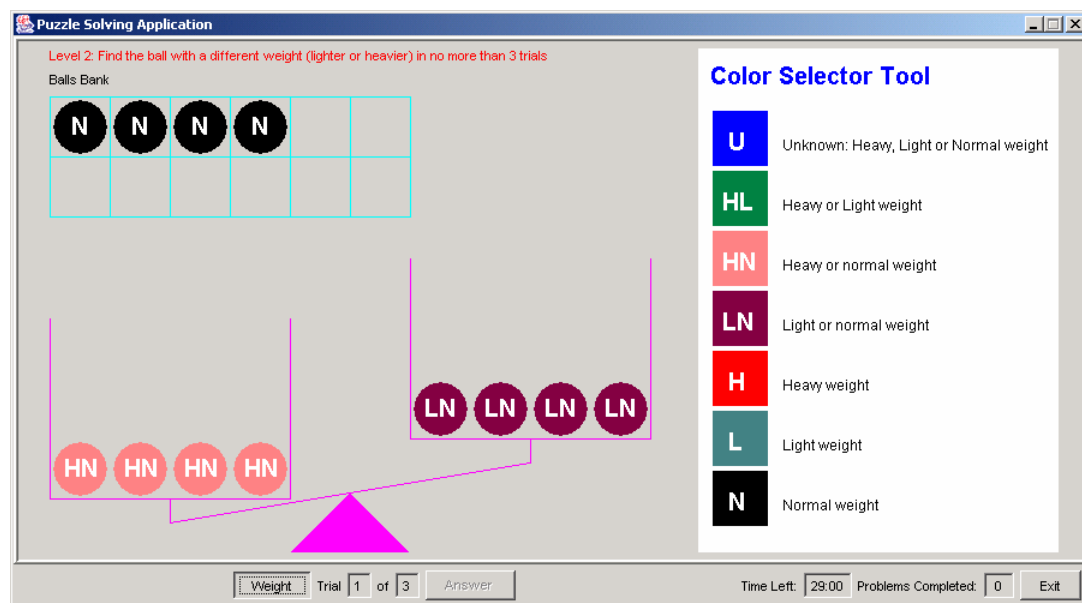
Verbal Rule Instruction	21	2	19
-------------------------	----	---	----

*Table 4*

Contingency table of the number of correct and incorrect answers in each group

Group	Trials	Correct	Incorrect
Reinforcement learning	350	207	143
Imitation learning	209	311	98
Verbal Rule Instruction	390	276	114

### Figures



Screen shot of the ball-weighing computer task in action in level 2.

**Symbolic/Verbal Instructions**

**Rules for Solving the 12 Balls / 3 Weighing Problem**

There are two important sub-goals to keep in mind while solving each problem.  
It will be necessary to alternate between selecting which balls to weigh and deciding which color markings to use.

**Selecting Balls**

IF this is the first weighing, THEN use 1/3 of the balls on each side of the scale.

IF the scale does not move in the first weighing, THEN use 3 unknown vs. 3 normal for the second weighing.

IF the scale moves in the first weighing, THEN use 1 potentially heavy ball + 2 potentially light balls vs. 1 normal ball + 1 potentially heavy ball + 1 potentially light ball for the second weighing.

IF the scale does not move in second weighing, THEN use 1 unknown ball from the ball bank vs. 1 normal ball for the third weighing.

IF the scale does not move in the second weighing, THEN use 1 potentially heavy ball from the ball bank vs. 1 potentially heavy ball from the ball bank for the third weighing.

IF the scale moves in the second weighing, THEN use 1 potentially light ball vs. 1 potentially light ball, OR 1 potentially heavy ball vs. 1 potentially heavy ball from the scale for the third weighing.

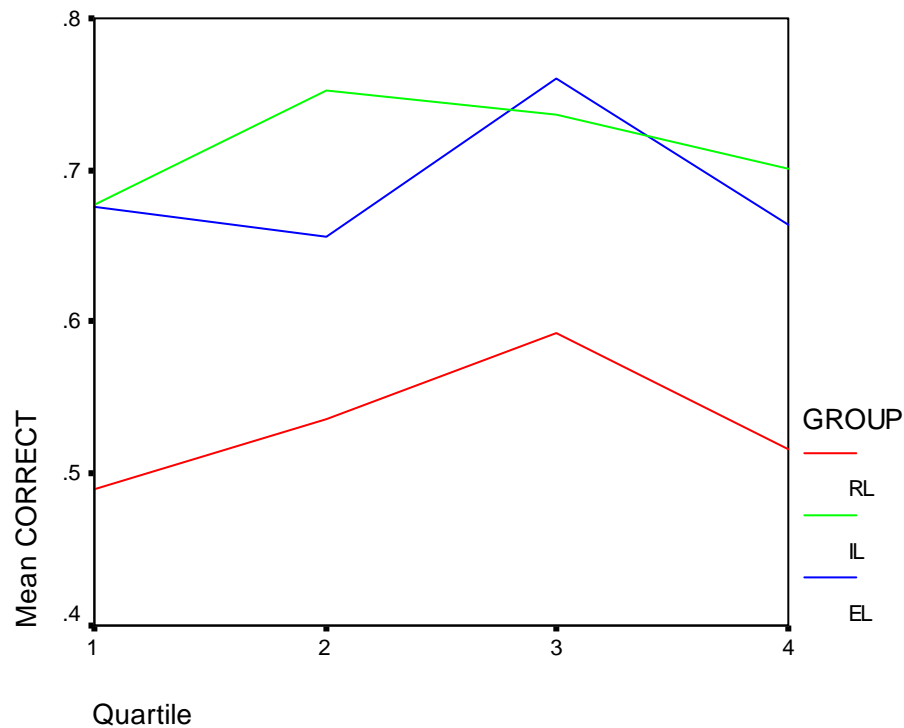
IF the scale moves in the second weighing, THEN use 1 normal ball vs. 1 potentially heavy ball from the scale OR use 1 normal ball vs. 1 potentially light ball from the scale OR use 1 potentially light ball from the scale vs. 1 potentially light ball from the scale for the third weighing.

**Marking Balls**

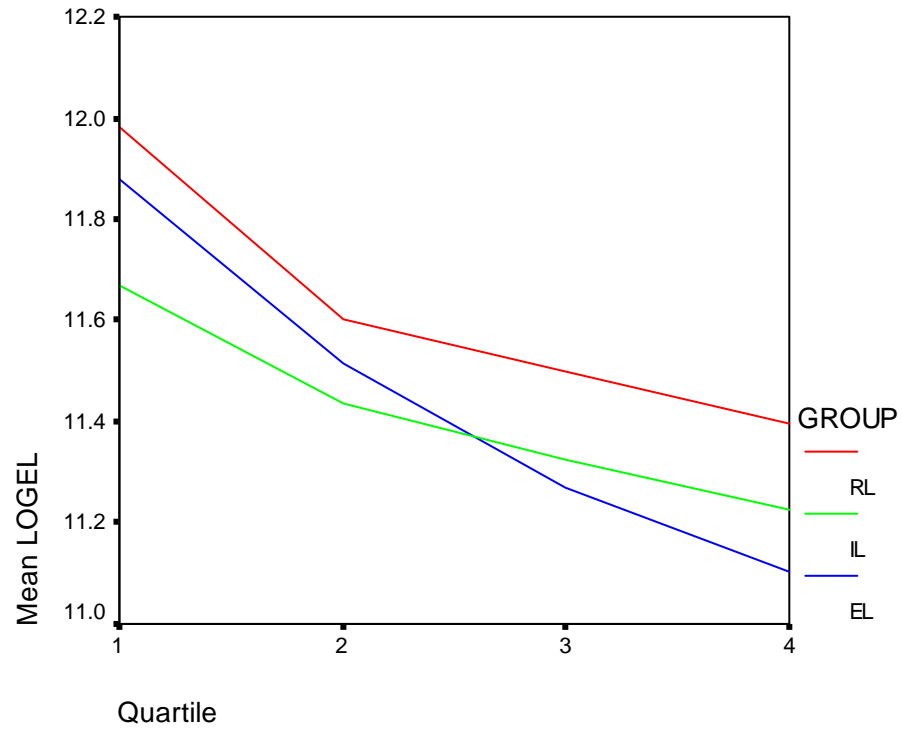
1. IF the scale does not move, THEN all balls on it are of normal weight.
2. IF the scale moves, THEN all balls left in the bank are of normal weight.
3. IF there are balls of unknown weight located on the side of the scale that moves up, THEN they are of "light or normal weight"
4. IF there are balls of unknown weight located on the side of the scale that moves down, THEN they are of "heavy or normal weight".
5. IF there are balls of "light or normal weight" located on the side of the scale that moves down, THEN they are of normal weight.
6. IF there are balls of "heavy or normal weight" located on the side of the scale that moves up, THEN they are of normal weight.

Done

Screen shot of the complete set of abstract rules. The abstract rule group will be allowed to study a subset of these for 10 minutes between level 1 and level 2.



Graph of average correct rate per trial quartile.



Graph of log of average elapsed time per trial quartile.