Enhancing Acquisition of Intuition versus Planning in Problem Solving

Dawn Chen (sdchen@ucla.edu)
Keith J. Holyoak (holyoak@lifesci.ucla.edu)
Department of Psychology
University of California, Los Angeles
Los Angeles, CA 90095 USA

Abstract

The acquisition of intuition, which guides problem solving by pruning unpromising strategies, is essential to the development of expertise in any domain. Problem-solving intuition may be viewed as analogous to search heuristics in artificial intelligence. One prediction inspired by this analogy is that practicing on subproblems and relaxed problems (versions of a problem with fewer constraints on the goal state and on the possible moves, respectively) may enhance the development of intuition for the full problem. Using the *n*-puzzle, we found that practice on relaxed problems did promote intuition compared to practice on the full problem, but impaired performance on solving the full problem. More detailed analyses suggest that practice on relaxed problems may discourage planning and encourage reliance on intuition. Planning is slower but more likely to produce optimal solutions if given enough time, whereas relying on intuition is faster but may lead to suboptimal solutions.

Keywords: Problem solving; intuition; planning; learning; heuristic search; *n*-puzzle.

Introduction

When encountering a problem in an unfamiliar domain for the first time, the novice may feel lost among what seems to be an indefinitely large number of possible actions that seem about equally promising, and end up pursuing some arbitrary path that leads nowhere. But after solving some number of problems from the same domain, the solver will eventually learn to consider only a limited number of approaches, those that are likely to prove effective. In commonsense terms, the learner has acquired *intuition* about the problem domain: an implicit sense of what to do in various types of situations that arise during problem solving (Gobet & Philippe, 2009). How is such intuition acquired through practice?

The standard account of general problem solving is Newell and Simon's (1972) proposal that the problem solver performs search within a problem space. A problem space can be visualized as a graph or tree in which the nodes represent possible states in the problem and each edge represents a legal move transforming one state into another. The legal moves in a problem are defined by its *operators*, or possible types of actions. The problem solver can search the problem space by starting at the node representing the initial state of the problem and moving to adjacent nodes by applying operators, until one of the nodes representing a goal state is reached. The solution to the problem is the successful path that the solver took through the problem-space graph.

Importantly, the problem solver may search the problem space not only by physically manipulating the external representation of the problem state (*external search*), but also by mentally transforming an internal representation (*internal search* or *planning*). During internal search, the problem solver need not always move from the current state to an adjacent node.

For most realistic problems, the problem-space tree is enormous, so that it is terribly inefficient even for a computer to solve the problem by using brute-force search algorithms that traverse the entire tree until a goal state is found. Heuristic search algorithms, on the other hand, are much more efficient because they use domain-specific knowledge to prune branches of the tree that never lead to the goal state or do not do so in an optimal way (i.e., in the minimum number of moves). A search heuristic may guide search by estimating the *distance* (minimum number of moves required) from any state to the goal so that, for example, a search algorithm can always choose to explore next the state that is closest to the goal (i.e., the greedy best-first search algorithm). This form of a search heuristic, commonly used in artificial intelligence, is called a *heuristic function*.

In many ways, the formal concept of a search heuristic is closely related to the commonsense concept of intuition in problem solving. Search heuristics prune branches in the problem-space tree that are unlikely to lead to the goal efficiently, just as problem-solving intuition focuses attention on just those paths that are likely to lead to a solution quickly. Search heuristics are usually fast to compute, but may lead to suboptimal solutions. Similarly, intuitive judgments arise quickly, but are fallible and may result in diminished accuracy or optimality compared to a solution strategy based on systematic analysis or careful planning. Furthermore, just as search heuristics rely on domain-specific knowledge, problem-solving intuition is restricted to a particular domain and is acquired only through multiple experiences with solving problems in that domain. Nonetheless, certain search heuristics are more general than others and apply to several domains with overlapping structure, just as the intuition gained from solving problems in one domain may apply to a related domain (see Hatano & Inagaki, 1986, for a discussion of routine vs. adaptive expertise). Finally, and most importantly for the present study, heuristic functions yield estimates analogous to the intuitive sense of closeness to the goal available to experienced problem solvers. The task we use to assess intuition will be based on subjective judgments of distance to the goal state.

The analogy between problem-solving intuition and search heuristics provides insights into how it might be possible to facilitate the acquisition of intuition in human problem solving. AI researchers have discovered that the solution lengths of *subproblems* and *relaxed problems* often provide good heuristic functions for the original problem (Prieditis, 1993). A subproblem removes one or more constraints on the goal state from the original problem, whereas a relaxed problem removes one or more constraints on the legal moves (i.e., it adds one or more operators). Thus, an instance of the original problem can be solved in fewer moves when translated into a corresponding subproblem or relaxed problem.

Applying the results from AI to the domain of human problem solving, solving subproblems and relaxed problems may facilitate the acquisition of intuition for the original problem. Therefore, learners who practice solving subproblems or relaxed problems may acquire better intuition for the original problem than those who receive the same amount of practice on only instances of the original problem. At the same time, planning may seem less necessary when solving subproblems and relaxed problems. Thus, the kind of learning experience that fosters development of intuition the most may also have a detrimental impact on planning. We will elaborate on these points in discussing our experimental findings.

Method

Participants

Seventy-two undergraduates from the University of California, Los Angeles participated for course credit. Participants were randomly assigned to either the control condition (n = 24), the subproblem condition (n = 24), or the relaxed problem condition (n = 24).

Materials

The *n*-puzzle Participants solved a computer version of the *n*-puzzle, which is illustrated in Figure 1. The *n*-puzzle consists of a square bounded space containing a smaller empty square and n initially misplaced square tiles numbered 1 to n. A legal move consists of sliding any tile into the empty square, and the goal state contains all the tiles in ascending order.

4	1	3
	2	5
7	8	6

initial state

1	2	3
4	5	6
7	8	

goal state

Figure 1: An 8-puzzle with a 5-step solution: Move 4 down, 1 left, 2 up, 5 left, and 6 up.

Subproblems and Relaxed Problems In the subproblems for the *n*-puzzle, participants were required to move only some of the tiles into their correct places. In the relaxed problems, participants could swap some of the tiles with adjacent tiles, in addition to sliding any tile into the empty square. These *swappable* tiles were displayed in a lighter color than the non-swappable tiles. Defined in this way, a subproblem that removes k goal constraints requires moving tiles 1 through $n - k$ into their correct places, and a relaxed problem that removes k move constraints contains one empty square and k tiles that can be swapped with neighboring tiles.

Generation of Puzzles All puzzles were generated randomly. The optimal A* search algorithm was used to ensure that each puzzle had the desired minimum solution length.

Procedure

All instructions and stimuli were presented on a computer, and participants responded using a mouse. In each condition, the participant was first given instructions on how to solve the type of puzzles (full, subproblem, or relaxed problem) in that condition. The participant then attempted to solve an initial 8-puzzle of the appropriate type, solvable in a minimum of three moves. An experimenter ensured that the participant understood the instructions and could solve the initial puzzle. In the subproblem condition, the initial puzzle required tiles 1-4 to be moved into place. In the relaxed problem condition, tiles 5-8 were swappable. That is, the number of constraints removed, k, was four for the initial puzzle in both the subproblem and relaxed problem conditions. After solving the initial puzzle, the participant took part in a training phase, a test phase, and finally an intuition assessment phase.

Training Phase The participant was told that more puzzles would now be given for practice, with a time limit of one minute and 30 seconds for each. The participant was told to solve each puzzle in as few moves as possible, and that there would be a penalty for every extra move made. These instructions were designed to discourage external search (the usual strategy for solving *n*-puzzles) and encourage internal search, which has been shown to enhance learning (O'Hara & Payne, 1998).

The participant then attempted to solve a sequence of 12 8-puzzles. In all conditions, the minimum solution lengths (a measure of difficulty) of the puzzles increased from 4 to 10 (i.e., the puzzles in the experimental conditions were not subproblem or relaxed versions of those in the control condition). In the experimental conditions, k also decreased from four to zero across the puzzles. During the presentation of each puzzle, the minimum solution length and the number of moves the participant had made so far were shown above the puzzle. After the participant solved each puzzle or the time limit expired for that puzzle, a dialog box informed the participant which event had occurred, the number of extra moves the participant made (if the puzzle was solved), and in the subproblem condition, the tiles to slide into place for the next puzzle. The participant could then take a break and click on a button to start the next puzzle when ready.

Test Phase After all 12 puzzles in the training phase had been presented, participants were told that there would now be a test, with the same instructions as for the practice puzzles. In the subproblem condition, participants were told to slide all tiles into place. Participants in all conditions then attempted to solve the same sequence of 12 full *n*-puzzles. The first six were 8-puzzles and the last six were 15-puzzles, and all puzzles could be solved in 12 moves. After each puzzle had been solved or had timed out, the next puzzle was presented without any feedback or time to rest. During both the training and test phases, the computer recorded for each puzzle whether it was solved, the solution time, the moves the participant made, the initial latency (the amount of time the participant took to make the first move), and the inter-move latencies (the time to make each subsequent move).

Intuition Assessment Phase After the test phase, participants made a series of 40 *pairwise distance comparisons*. In each comparison, they were presented with two different puzzle states and had to click on the one that they believed was closer to the goal within a short time limit. No feedback was given. The short time limit was designed to elicit a quick, intuitive judgment and prevent participants from solving the puzzles mentally and then counting the number of moves used. Because experts in a domain often have an intuitive sense of how close they are to solving a problem, and heuristic functions estimate the distance of any given state to the goal, this distance comparisons task serves to assess participants' intuition on the *n*-puzzle.

The first 20 pairs to be compared were 8-puzzles, with 10 seconds each, and the last 20 pairs were 15-puzzles, with 12 seconds each. The true distances of the puzzles ranged from 1 to 28, and the ratio of the shorter distance to the longer distance in each pair was between .2 and .91. For each comparison, which puzzle was chosen and the time taken to make that choice were recorded.

Results and Discussion

Dissociation of Performance on Solving Puzzles and Comparing Distances

The mean percentage of full *n*-puzzles solved during the test phase in each condition is shown in Figure 2. The relaxed problem group solved a significantly lower percentage of puzzles during the test phase (M = 57.99, SD = 23.25) than the control group (M = 69.79, SD = 14.08), $F(1, 69) = 5.18$, $p = .026$, and also the subproblem group (M = 68.75, SD = 15.20), $F(1, 69) = 4.30$, $p = .042$. The latter two groups did not differ reliably.

However, as shown in Figure 3, the relaxed problem group correctly solved the most problems on the distance comparisons task, which assesses intuition. The percentage

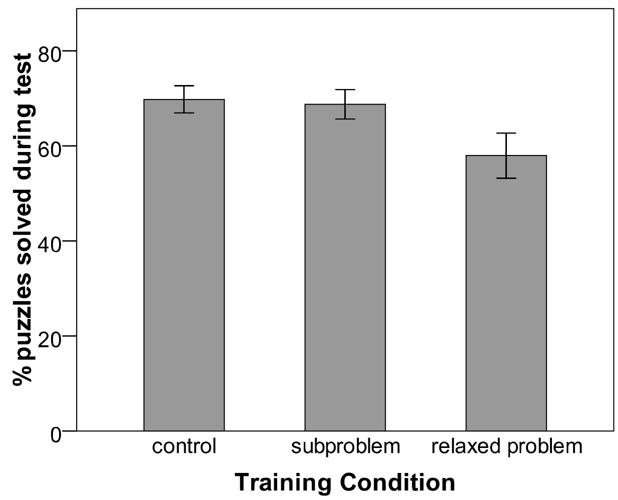

Figure 2: Mean percentage of *n*-puzzles solved by participants in each training condition during the test phase. Error bars in all data figures represent 1 standard error of the mean.

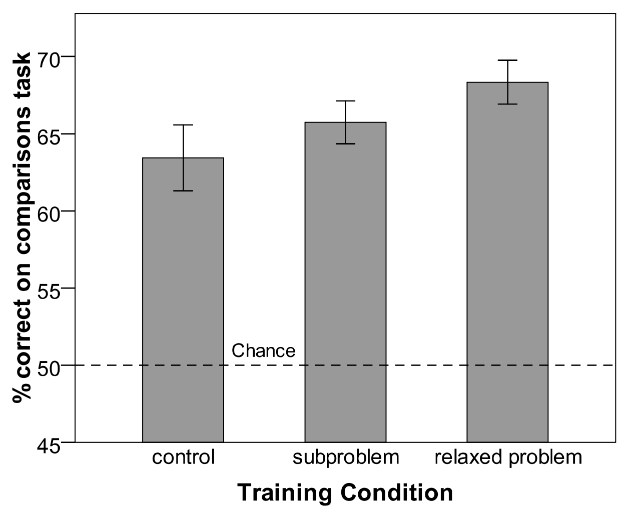

Figure 3: Mean percentage of comparisons solved correctly on the distance comparisons task in each condition.

of comparisons correct was significantly higher for the relaxed problem group (M = 68.33, SD = 6.94) than for the control group (M = 63.44, SD = 10.47), $F(1, 69) = 4.22$, $p = .044$. Performance of the subproblem group on the comparisons task fell between that of the other two groups, but did not differ significantly from either.

To further investigate the difference in performance on the distance comparisons task, we divided the pairwise distance comparisons into an "easy" set and a "hard" set based on the overall performance of the participants on each comparison. For each comparison problem, we calculated the proportion q of participants (over all three conditions) who solved that problem correctly. We then calculated the median value of q over all comparisons. A comparison that

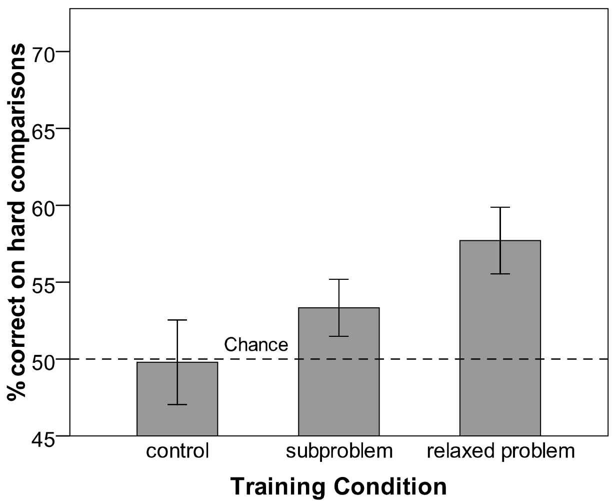

Figure 4: Mean percentage of hard comparisons solved correctly on the distance comparisons task in each condition.

had a *q*-value higher/lower than the median was assigned to the "easy"/"hard" set. All groups performed about the same on the easy comparisons, but as Figure 4 shows, the relaxed problem group performed the best on the hard comparisons. In particular, the relaxed problem group correctly solved a significantly higher percentage of the hard comparisons ($M = 57.71$, $SD = 10.63$) than the control group did ($M = 49.79$, $SD = 13.47$), $F(1, 69) = 6.00$, $p = .017$. Thus, the relaxed problem group performed very well on the intuition task, especially the harder problems, compared to the control group.

How could participants in the relaxed problem group have apparently acquired such good intuition on the full *n*-puzzle, and yet perform relatively poorly in actually solving it? A possible explanation is that because planning (internal search) is harder and seemingly less necessary when solving the relaxed problems, participants in the relaxed problem group learned to plan less and rely more on their intuition during the training phase. Thus, even though their intuition became more developed (as evidenced by their performance on the distance comparisons task), their decreased use of planning caused them to perform poorly on solving the puzzles in the test phase. Participants in the control group, on the other hand, learned to rely more on planning and less on their intuition during the training phase, because they were trying to minimize the number of moves they made and it was easier for them to plan. Increased planning led them to perform better on the test puzzles, but their intuition was less developed. We will now present evidence to support each of these claims.

The Relaxed Problem Training Condition Discourages Planning

Planning is Harder on Relaxed Problems This is true for two reasons. First, internally visualizing the move of swapping two tiles in the relaxed problem imposes a greater working memory load, because the participant must now keep track of the new locations of both tiles, rather than just one tile in the sliding move. Manipulating an internal representation of the puzzle state to reflect a swapping move might take longer as well. Second, the introduction of additional legal moves in the relaxed problem also makes planning harder because participants have to consider more moves at each state (that is, the *branching factor* is higher). In order to plan, participants must also remember more information about which paths they have already mentally explored to some depth and have determined to be unpromising.

The hypothesis that the swapping move consumes more working memory is supported by the finding that the average length of unbroken sequences of backtracking moves during the training phase was significantly lower in the relaxed problem group ($M = 1.34$, $SD = .36$) than in the control group ($M = 1.88$, $SD = 1.11$), $F(1, 61) = 4.66$, $p = .035$, and also the subproblem group ($M = 2.04$, $SD = .82$), $F(1, 61) = 8.93$, $p = .004$. In contrast, no reliable differences among conditions were observed in the test phase. Backtracking for a number of moves requires remembering all those previous moves, and participants solving relaxed problems may have backtracked for fewer moves because they could not remember as many past moves, since storing a single move requires more working memory capacity on average.

Planning Seems Unnecessary on Relaxed Problems Because relaxed problems have a higher branching factor, the problem-space graphs for relaxed problems are more connected and so there are more ways to reach the goal state. Thus, it may seem unnecessary to plan one's moves before executing them, since no matter how far away one wanders from the goal, there is always some way to get back onto the right track. In other words, local minima do not exist in the problem space, so a greedy (hill-climbing) search algorithm that always chooses the state with the shortest estimated distance to the goal to explore next cannot become trapped, and is thus sufficient. Accordingly, participants in the relaxed problem group probably learned to use a greedy search algorithm, which does not look ahead and thus requires little effort. Moreover, a greedy search algorithm relies heavily on the heuristic function, so its use would foster development of intuition for participants in this condition.

One piece of evidence that participants in the relaxed problem group planned less than those in the other conditions is that they made extra moves more often during the training phase. The percentage of solved puzzles in the training phase that were solved with extra moves was significantly higher in the relaxed problem group ($M = 49.99$, $SD = 19.76$) than in the control group ($M = 20.92$, $SD = 13.67$), $F(1, 69) = 34.53$, $p < .001$, and also the subproblem group ($M = 25.97$, $SD = 17.43$), $F(1, 69) = 23.58$, $p < .001$. Furthermore, the relaxed problem group had significantly higher average solution times during the training phase (M = 35.23s, SD = 9.07s) than did the control group (M = 27.68s, SD = 8.29s), $F(1, 69) = 9.86$, $p = .002$, and also the subproblem group (M = 24.64s, SD = 7.53s), $F(1, 69) = 19.41$, $p < .001$. Participants in the relaxed problem condition may have found planning harder and thus took longer on average to plan a single move (when they did plan); in addition, their longer, less optimal solutions took more time to execute. These differences indicate that the relaxed problem participants did not or could not plan as far ahead as did the participants in the other conditions, and tended to meander around the problem space for a while before reaching the goal.

The average initial latency on a puzzle, or the average amount of time a participant spent thinking before making the first move on a puzzle, is a clear indicator of how much a participant plans voluntarily. (While the average inter-move latency is also an indicator of planning, higher inter-move latencies could also indicate that the participant was stuck in the middle of solving a puzzle and was forced to think carefully about what to do next.) The average initial latency was not significantly lower for the relaxed problem group during the training phase, as might be expected if these participants were planning fewer moves ahead; however, the lack of a difference could reflect the offsetting effect of planning each move being harder for the relaxed problems and thus taking longer. During the test phase, when all participants were solving the full *n*-puzzles, the average initial latency was indeed significantly lower for the relaxed problem group (M = 10.37s, SD = 4.46s) than for the control group (M = 14.75s, SD = 6.02s), $F(1, 69) = 7.33$, $p = .009$, indicating that the relaxed problem group continued to plan fewer moves ahead during the test phase.

Increased Planning is Associated with Better Puzzle-Solving Performance

Not surprisingly, increased planning is associated with better puzzle-solving performance. The average initial latency was not correlated with the number of puzzles solved during the training or test phase, perhaps because some participants tended to get stuck at the very beginning and could not solve many puzzles, or were just too slow in general to solve many puzzles. However, average initial latency was negatively correlated with performance measures such as the average number of extra moves made on solved puzzles [$r(70) = –.37$, $p = .002$ for the training phase and $r(70) = –.46$, $p < .001$ for the test phase], and the percentage of backtracking moves [$r(70) = –.26$, $p = .026$ for the training phase and $r(70) = –.31$, $p = .007$ for the test phase]; and positively correlated with the percentage of moves that decreased the true distance of the problem state to the goal [$r(70) = .33$, $p = .005$ for the training phase and $r(70) = .47$, $p < .001$ for the test phase]. These results indicate that the more the participant planned before making the first move, the better the moves the participant made later on.

Recall that on relaxed problems, which do not have many local minima, a greedy search algorithm is sufficient. However, greedy search may get stuck in local minima on the full *n*-puzzle, for which the problem-space graph is not as well-connected. Accordingly, if participants in the relaxed problem group did indeed use a greedy search algorithm, they would perform poorly during the test phase. The control group, on the other hand, may have learned to use a more effective search algorithm involving greater look-ahead. Such a search algorithm could achieve an acceptable level of performance with a relatively poor heuristic function. Thus, participants in the control condition would not acquire intuition during the training phase to the degree that those in the relaxed problem group did.

Planning and Intuition are Dissociated

For every participant, we calculated a composite score on the intuition task by summing the values of $1 – q$ for all comparison problems that the participant solved correctly. Recall that for each comparison, q is the proportion of all participants who solved that comparison correctly. Thus, $1 – q$ is the estimated probability of choosing the incorrect response on a given comparison, an empirical measure of its difficulty. Therefore, the composite score on the intuition task gives greater weight to more difficult problems.

We calculated correlations between the composite intuition score and measures of planning for each training condition separately to test whether planning and intuition are dissociated within each group. The following correlations appeared for measures of planning during the training phase: The composite intuition score for the control group was negatively correlated with the average initial latency, $r(22) = –.41$, $p = .047$, as well as the average inter-move latency, $r(22) = –.47$, $p = .021$. For the subproblem group, the composite intuition score had a negative correlation with the average inter-move latency, $r(22) = –.50$, $p = .013$, and a near-significant positive correlation with the percentage of puzzles that were solved with extra moves, $r(22) = .40$, $p = .055$. Finally, for the relaxed problem group, there was a weak negative correlation between the composite intuition score and the percentage of moves that decreased the true distance of the problem state to the goal, $r(22) = –.35$, $p = .098$.

During the test phase, the composite intuition score for the control group had a near-significant negative correlation with the average initial latency, $r(22) = –.39$, $p = .061$, as well as a slight positive correlation with the average number of extra moves, $r(22) = .36$, $p = .082$.

These findings indicate that participants in our study mainly took one of two approaches to solving the puzzles and comparison problems. One was a more analytic or algorithmic approach based on planning, and the other was a more holistic or heuristic approach based on intuition. While the first approach was more effective for solving the full *n*-puzzles, the second approach was more effective on the task requiring speeded comparison of distances to the goal state. The control training condition encouraged the more analytic problem-solving style, and participants in this condition developed a more effective search algorithm. In contrast, the relaxed problem training condition encouraged the more intuitive problem-solving style, and participants in this condition developed a more accurate heuristic function.

Conclusions

The present study demonstrates a dissociation between two core mechanisms on which expertise in problem solving depends: internal search (planning) and use of a heuristic function to evaluate locally available moves (intuition). Training on problems with fewer possible moves at each choice point (full *n*-puzzles) encouraged a more analytic problem-solving style, whereas training on relaxed versions of the same problem type that allow more possible moves encouraged a more intuitive problem-solving style. In the present study, the analytic style led to better performance on actually solving the full *n*-puzzles, but the more intuitive style led to better performance on a task requiring fast evaluations of how close a problem is to being solved.

Our results should not be construed as evidence that the development of analytical thinking and intuition are mutually exclusive. In fact, true experts in solving problems in complex domains such as chess (Chase & Simon, 1973; Gobet & Charness, 2006) appear to rely heavily on both intuition and planning, with the relative importance of intuition increasing when performance is time-constrained (Gobet & Simon, 1996). The time frame of the present study was far shorter than the years required to develop true expertise (Ericcson, 1996). Even by the end of the experiment, our participants remained novices on the *n*-puzzle. An expert solver of the *n*-puzzle would no doubt plan ahead more as well as make better intuitive judgments relative to a novice. The ability to quickly evaluate problem states should allow the problem solver to plan more moves ahead, just as heuristic functions reduce the branching factor and thus allow the search algorithm to search to a greater depth within the same amount of time. In fact, Charness (1981) found that skilled chess players search more deeply than novice players do, indicating that good intuition aids planning in problem solving.

What our findings do indicate is that these two basic approaches to problem solving may not be acquired in lock-step fashion, and to some extent constitute competing problem-solving strategies. Moreover, the two different approaches may be maximally effective for different types of problems. The systematic, analytic approach is slower and places a greater burden on working memory, but is more likely to lead to optimal solutions, and thus may be preferable for problems that can be solved slowly with the assistance of external aids to memory. In contrast, the holistic, intuitive approach is faster and less dependent on working memory, and hence will often be preferable when the problem must be solved under severe time constraints.

One example of this dichotomy is battlefield versus hospital triage. In the hospital, medical personnel may take a more analytic approach, carefully considering the consequences of each possible action. On the battlefield, by contrast, the need for decisions may be so urgent that the only possible approach is to rely on intuition or “gut feelings.” An important direction for future research will be to determine whether the present findings using the toy example of the *n*-puzzle in fact generalize to real-world problem solving (cf. Gobet & Philippe, 2008).

Acknowledgments

This research was funded by a University Fellowship and a Chancellor’s Prize from the Graduate Division at the University of California, Los Angeles, and by ONR grant N000140810186.

References

Charness, N. (1981). Search in chess: Age and skill differences. *Journal of Experimental Psychology: Human Perception and Performance*, *7*, 467–476.

Chase, W. G., & Simon, H. A. (1973). The mind’s eye in chess. In W. G. Chase (Ed.), *Visual information processing*. New York: Academic Press.

Ericsson, K. A. (1996). The acquisition of expert performance: An introduction to some of the issues. In K. A. Ericsson (Ed.), *The road to excellence: The acquisition of expert performance in the arts and sciences, sports, and games*. Mahwah, NJ: Erlbaum.

Gobet, F., & Charness, N. (2006). Expertise in chess. In K. A. Ericsson, N. Charness, P. J. Feltovich, & R. R. Hoffman (Eds.), *The Cambridge handbook of expertise and expert performance.* New York: Cambridge University Press.

Gobet, F., & Philippe, C. (2008). Towards an alternative to Benner’s theory of expert intuition in nursing: A discussion paper. *International Journal of Nursing Studies*, *45*,129-139.

Gobet, F., & Philippe, C. (2009). Expertise and intuition: A tale of three theories. *Minds and Machines*, *19*, 151-180.

Gobet, F., & Simon, H. A. (1996). The roles of recognition processes and look-ahead search in time-constrained expert problem solving: Evidence from grandmaster-level chess. *Psychological Science*, *7*, 52-55.

Hatano, G., & Inagaki, K. (1986). Two courses of expertise. In H. Stevenson, H. Azuma, & K. Hakuta (Eds.), *Child development and education in Japan*. San Francisco: Freeman.

Newell, A., & Simon, H. A. (1972). *Human problem solving*. Englewood Cliffs, NJ: Prentice-Hall.

O’Hara, K. P., & Payne, S. J. (1998). The effects of operator implementation cost on planfulness of problem solving and learning. *Cognitive Psychology*, *35*, 34-70.

Prieditis, A. E. (1993). Machine discovery of effective admissible heuristics. *Machine Learning*, *12*, 117-141.