

ILLEGAL MOVES AS A SOURCE OF PROBLEM DIFFICULTY

By

MARTIN E. KNOWLES

A THESIS PRESENTED TO THE GRADUATE SCHOOL
OF THE UNIVERSITY OF FLORIDA IN PARTIAL FULFILLMENT
OF THE REQUIREMENTS FOR THE DEGREE OF
MASTER OF SCIENCE

UNIVERSITY OF FLORIDA

2004

Copyright 2004

by

Martin E. Knowles

ACKNOWLEDGMENTS

I would like to thank my parents, who always did everything they could to provide me with the best education possible, for their continued support of my education and career. I would like to thank my new wife Stephanie for her support and patience and for planning our wedding while I continued to work on my thesis until the day of our wedding. I thank my advisor Peter Delaney, who without his help and guidance I would not be where I am today. I thank Dr. Berg and Dr. Fischler who provided valuable feedback throughout the course of this project. I would also like to thank Dania Hadjez and Paul Niesen for helping to collect the data for this project.

TABLE OF CONTENTS

	<u>Page</u>
ACKNOWLEDGMENTS	iii
LIST OF TABLES	vi
LIST OF FIGURES	vii
ABSTRACT	viii
CHAPTER	
1 PROBLEM SOLVING	1
Sources of Difficulty: Legal Moves	5
Sources of Difficulty: Illegal Moves	8
Purpose	14
2 EXPERIMENT 1	16
Thinking Aloud	16
Improvement	17
Individual Differences	17
Methods	18
Participants	18
Problem and Interface	19
Procedure	20
Illegal Moves Made and Considered	22
Results and Discussion	23
Silent Versus Aloud Comparisons	23
Improvement	25
Individual Differences	26
State Versus State Comparisons	28
3 EXPERIMENT 2	32
Methods	33
Participants	33
Problem and Interface	34
Design	35
Aloud variable	35

Cost variable.....	35
Procedures	36
Results and Discussion	37
Silent Versus Aloud Comparisons.....	38
Cost Versus No-Cost Comparisons	39
Interactions	40
Improvement.....	40
Individual Differences	41
State Versus State Comparisons.....	42
 4 GENERAL DISCUSSION	 44
Thinking Aloud.....	44
Individual Differences	45
Illegal Move Selection.....	46
Legal Moves	49
Improvement.....	50
State Differences.....	52
Conclusion	53
 APPENDIX	
 A QUESTIONNAIRES.....	 55
Need For Cognition	55
Impulsivity.....	56
 B WORKING MEMORY	 59
OSPAN	59
 LIST OF REFERENCES.....	 62
 BIOGRAPHICAL SKETCH	 64

LIST OF TABLES

<u>Table</u>	<u>page</u>
2-1. Silent and Think Aloud Comparisons, Experiment 1 [*]	24
2-2. Correlations for Individual Differences, Experiment 1 [*]	27
2-3. State vs State Comparisons, Experiment 1 [*]	29
3-1. Silent and Think Aloud Comparisons, Experiment 2 [*]	38
3-2. No-Cost and Cost Comparisons, Experiment 2 [*]	40
3-3. Cost Group State vs State Comparisons, Experiment 2 [*]	42
3-4. No-Cost Group State vs State Comparisons, Experiment 2 [*]	43

LIST OF FIGURES

<u>Figure</u>	<u>page</u>
1-1. The circles represent nodes in the problem space.	3
1-2. This figure is a map of the legal and illegal problem space for the hobbits and orcs problem.	4
2-1. Above is the interface seen by participants in Experiment 1.	20
2-2. The figure shows a comparison of illegal moves made in Experiment 1, by State and the illegal moves made by participants in Jeffries et al. (1977).....	29
3-1. This figure shows the interface that participants saw in Experiment 2.	35
3-2. The graph shows a comparison between the cost and no-cost groups for illegal moves committed in Experiment 2.	39

Abstract of Thesis Presented to the Graduate School
of the University of Florida in Partial Fulfillment of the
Requirements for the Degree of Master of Science

ILLEGAL MOVES AS A SOURCE OF PROBLEM DIFFICULTY

By

Martin E. Knowles

May 2004

Chair: Peter F. Delaney
Major Department: Psychology

I present two experiments exploring the role of illegal moves in problem solving and the contribution of illegal alternatives to problem difficulty. In both experiments, I attempted to find out if having participants think aloud while working on a problem would affect problem solving ability. I also attempted to assess which individual differences contributed to problem solving efficiency and in what way. Participants' data were also analyzed to evaluate if learning occurred throughout the problem, as demonstrated by a reduction in legal and illegal moves in the second half of the problem. Individual states of the problem were examined in detail to determine if states differentially affected problem difficulty. Finally, in the second experiment, the cost of executing an illegal move was increased to determine if the number of illegal moves executed could be reduced without manipulating the interface or problem space.

In both experiments, participants solved the Hobbits and Orcs isomorph of a river crossing problem, which consisted of three hobbits, three orcs, a river and a boat. After learning the rules, participants were asked to work on the problem either silently or while

thinking aloud. In Experiment 2, half of the participants were informed that there would be an additional cost for each illegal move. After completion of the problem participants performed a task to assess their working memory span. In Experiment 1, participants also completed a questionnaire to assess their Need for Cognition (NFC) and Impulsivity.

In both experiments the performance of those participants that solved the problem while thinking aloud did not differ from those participants who worked on the problem silently. I was also unable to detect any influence of the individual measures I obtained for working memory, NFC and impulsivity. When the problem was split into halves the analyses revealed that participants executed fewer illegal and legal moves in the second half, demonstrating an improvement in performance. The comparison of the individual states revealed that states do differ in difficulty and that some states may contribute more to problem difficulty. In Experiment 2, increasing the cost of an illegal move resulted in a decrease in the number of illegal, but not legal, moves committed by participants.

These findings support the claim that instructing participants to think aloud as they work on a problem does not affect problem solving performance. I also found evidence that participants' performance may change as they gain experience with a problem space. The results also showed that the difficulty of a problem may lie within the individual states and not at the global level of the entire problem. In the second experiment, participants were able to improve their performance even when the difficulty level of the problem remained unchanged. I was also unable to determine any contribution of individual differences from the measures I obtained in both experiments.

CHAPTER 1 PROBLEM SOLVING

Problem solving has been a part of our nature since humankind's first thoughts and has remained an important part of human life. A problem can be defined as a question, matter, or situation to be considered, solved or answered. Problem solving is not only important because it allows us to do well in school or to advance our careers, but because we encounter problems outside of the classroom and outside of work on a daily basis. The solutions and answers we generate and the decisions we make affect not only the outcome of the situation, but may also affect those around us and potentially the direction of our own lives. Due to the potential impact of our decisions it would be beneficial to obtain a better understanding of problems and what factors contribute to masking the correct or optimal solutions, thus contributing to problem difficulty.

Kotovsky, Hayes, and Simon (1985) suggested a number of factors that contribute to problem difficulty.

However, the role of the consideration and/or selection of illegal moves in problem difficulty remains unclear, as does the contribution of illegal moves to problem difficulty. I explored whether illegal moves contribute to problem difficulty, in what ways and to what degree. Gaining insight into these questions would be valuable for helping others become more efficient problem solvers, because it could lead to techniques for decreasing problem difficulty.

A problem can be thought of as having two types of moves, legal and illegal, which help to define the problem space. A problem space consists of nodes that represent each

of the valid (legal) states of the problem; it is a map of all the legal problem states and their connections (Newell & Simon, 1972). These valid states are defined by the rules of the problem. If a move violates one of the rules, then the resulting state takes the problem solver outside of the legal problem space. Thus, illegal moves are those moves that are not included in a problem's problem space. However, we can also consider illegal moves as having their own problem space. This illegal problem space also consists of nodes or states that I will refer to as phantom nodes. Phantom nodes are those states that lie outside the legal problem space and do not exist according to the rules. Any state that was reached by violating one of the rules is a phantom node and is nonexistent in a map of the legal problem space. States that appear to be legal may still be considered phantom nodes if they were achieved through illegal moves. For example, if an illegal move leads to a state that is identical to a state contained in the legal problem space, that state is still considered a phantom node because it was obtained through an illegal move. Figure 1-1 displays a sample problem space consisting of both a legal and an illegal problem space. Moves to the left of the initial state do not violate any rules and are considered to make-up the legal problem space. In the problem shown, moves to the right of the initial state violate the rules; therefore any moves contained in these states are considered phantom nodes. A map of the actual problem space for the Hobbits and Orcs problem is displayed in Figure 1-2. Legal states are represented by white boxes and illegal states are represented by gray boxes.

This paper focuses on problem difficulty and how different types of moves contribute to and influence problem difficulty and wasted effort. Problem difficulty can primarily be thought of as the amount of wasted effort or extra moves that do not bring

the problem solver closer to the goal. For example, a problem may be considered very difficult if it takes the problem solver twice as many moves to solve the problem than are necessary to reach the solution. In addition, a problem may be considered very easy if no additional moves are required to solve the problem than what is needed to reach the solution. However, problem difficulty can also be assessed as the amount of time needed to find a solution. If a problem were solved in the minimum number of moves possible, but required extra time and thought to discover the solution then the problem may also be classified as very difficult, even though no extra steps were required. In any given problem there are two types of moves that may contribute to difficulty, legal and illegal.

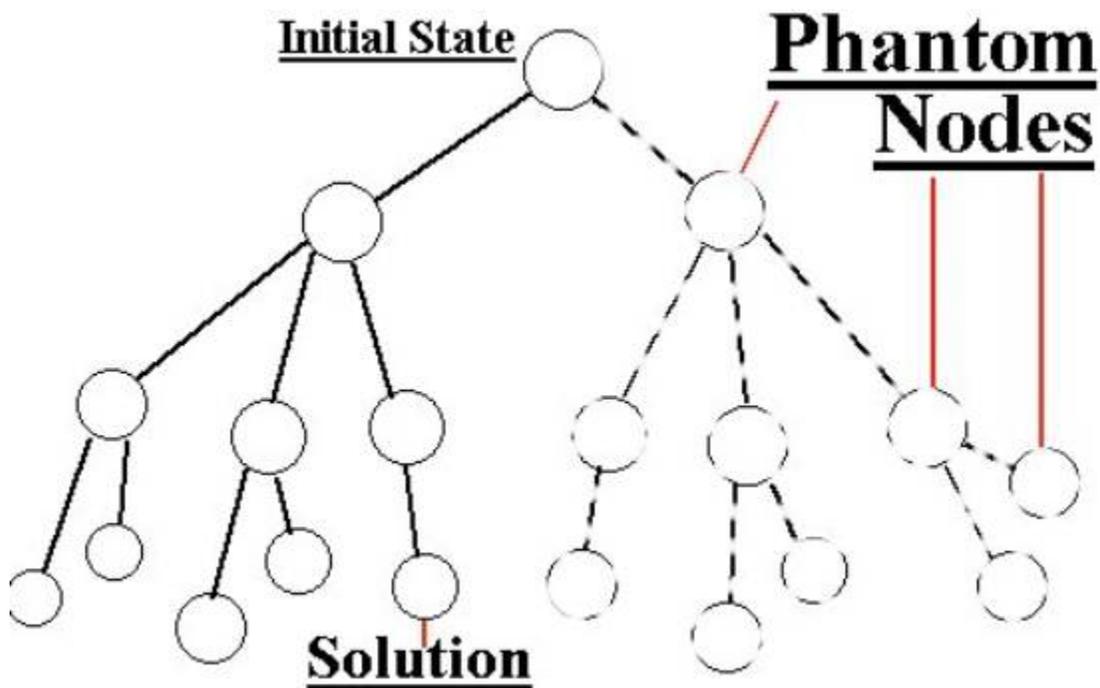


Figure 1-1. The circles represent nodes in the problem space. Nodes to the left of the initial state represent valid nodes that do not violate a rule. Nodes to the right, that are dashed, represent Phantom Nodes that are not valid because they violate a rule.

Legal moves are those moves that do not violate the rules, however they cause difficulty or wasted effort when they do not bring the problem solver closer to the

Sources of Difficulty: Legal Moves

Over the years, attempts have been made to understand what problem solvers do while working on a problem and how they solve problems. According to Anderson (1990), the most dominant features of human problem solving are that humans choose moves that avoid previously visited states and moves that are more similar to the goal state (a strategy called hill climbing) (Atwood & Polson, 1976). Also taken into account is the probability of achieving the goal, the cost of a specific move, and the amount of effort previously spent (Anderson 1990, Lovett & Anderson, 1996). In addition, various attempts have been made to understand what problems solvers are doing by creating computer models to mimic human performance. These models employ heuristic strategies to advance through a problem space. One such model is GPS (General Problem Solver), which uses means-ends heuristics to advance towards a goal (Newell & Simon, 1988). The means-ends heuristic involves assessing the distance between the current state and the goal state and then applying an operator to reduce that distance. Breaking down the problem into subgoals or smaller parts simplifies the problem and makes it more manageable and easier to solve in most cases. Various authors have also proposed other heuristics and strategies that are specific to particular problems.

However, problem-solving strategies such as means-ends and hill climbing are not always successful, and may contribute to problem difficulty. For example, a means-ends strategy would ultimately fail when applied to the Missionaries-Cannibals river-crossing problem (Jeffries, Polson, Razran, and Atwood, 1977). Jeffries et al. state that applying means-ends heuristics would result in reaching a state where no moves meet the legal move criterion for this strategy. It seems that choice of strategy and the effectiveness of heuristics are one component of problem difficulty.

Kotovsky, Hayes, and Simon (1985) proposed several other determinants that may contribute to problem difficulty. One determinant is the size of the problem space. When more states and choices in each state exist, then the difficulty of the problem likely increases because there are simply more options to consider. However, this is not always true, because other components of problem difficulty may be much more potent predictors of problem difficulty. Some problems have very small problem spaces yet are extremely challenging. The Missionaries and Cannibals problem, for example, only has a 16-node problem space, but seems very challenging to problem solvers who encounter the problem for the first time. Another determinant of difficulty is the internal representation of the problem, which depends upon the rules and how the problem solver perceives them. Some problems may be viewed as Change problems and others as Transfer or Move depending upon how a problem solver interprets the rules (Simon & Hayes, 1976). Problems that are viewed as Change problems are represented as having objects remain in their current location, while their properties are changed (e.g., Go). Transfer problems, on the other hand, are represented as objects being moved from one location to another (e.g., Chess). Kotovsky et al. consider Change problems to be more difficult because they showed that they take longer to solve and learning the rules and making judgments about problem legality also require more time compared to Move problems. Rule difficulty or how easy the rules are to learn seems to be another determinant of problem difficulty. Kotovsky et al. found that participants often begin working on a problem before they have adequately learned the rules. This causes them to refer back to the rules often throughout the problem until the rules are fully learned. The

discovering or defining of a legal move is also considered a determinant of problem difficulty.

Kotovsky and Simon (1990) found that the ambiguity of what constitutes a legal move was the major determinant of problem difficulty in the Chinese Ring Puzzle. In the Chinese Ring Puzzle, the goal is to remove all five rings from a bar. However, it is not obvious what physical manipulations will remove a ring from the bar. Kotovsky and Simon had participants unsuccessfully work on the Chinese Ring Puzzle for two hours. However, when a digital isomorph was created that clearly displayed what constituted a move, participants were able to solve the isomorph in under 30 minutes. Another determinant may include how demanding the memory load is while considering the next move. If a problem requires simultaneously holding many items in working memory, then the solver may choose a poorer move than if additional resources were available to aid in the decision process. How the problem solver's real world knowledge agrees with the rules of the problem may also be a factor of problem difficulty (Griggs & Cox, 1982; Kotovsky, Hayes & Simon, 1985). If the problem solver already knows the rules because they are part of everyday experience, then there is no need to allot resources to check the rules, meaning that more resources will be available for other problem solving activities. They also proposed the rule application hypothesis or the ease of applying the rules as a determinant of problem difficulty (Hayes & Simon, 1977). The rule application hypothesis suggests that the difficulty of conducting the tests that must be performed to determine the legality of moves is a major source of problem difficulty. In other words, the checking of a move against the rules to determine if the move is legal imposes

cognitive load, and to the degree that this is difficult the problem also becomes more difficult.

Sources of Difficulty: Illegal Moves

One further reason that problems are difficult is the presence of illegal moves. The most obvious contribution of illegal moves to problem difficulty can be seen when an illegal move is selected. Selection of an illegal move can either take the problem solver down a path that does not exist according to the rules or it may even reset or terminate the problem. In either case the problem solver is wasting time and resources by not advancing them through the problem. Selecting an illegal move may also result in terminating the problem completely.

Even when an illegal move is not chosen it may still contribute to problem difficulty. If a problem solver spends time and resources evaluating and considering an illegal move then this takes resources that could have been used to evaluate legal alternatives, thus contributing to wasted effort and ultimately problem difficulty. Illegal moves can be completely avoided, even without a map of the problem space, simply by checking the rules before making a move and correctly rejecting any illegal moves. Why, then, do people make these illegal moves if they can be avoided by checking the rules? Jeffries, Polson, Razran, and Atwood (1977) propose that problem solvers select illegal moves due to resource limitations. They believe that when a problem solver is performing at or near their resource limit they may miscalculate a future state or they may fail to check the future state for legality, thus resulting in the selection of an illegal move.

Atwood and Polson's model (1976) for water jugs involved the use of heuristics, which has broad generalizability to problem solving in other domains. Jeffries et al. (1977) extended Atwood and Polson's (1976) model for water jug problems to show its

generalizability to other problems such as the Missionaries-Cannibals and other river crossing problems. The model of Jeffries et al. (1977) is important because it attempts to demonstrate what participants may be doing as they solve a problem such as the Missionaries-Cannibals problem and because it makes assumptions about the selection of illegal moves. Understanding what a problem solver is doing is vital to understanding how we can improve performance. Their model consists of a three-stage process, which considers acceptable moves, finds a move leading to a new state, and finds the optimal move or makes a random move. A memory process is also included, which helps to determine if a state has previously been visited. Finally their model consists of an evaluation process, which includes the illegal move filter for testing the legality of a chosen move.

My main interest in the model of Jeffries et al. (1977) is that it makes assumptions about how and why problem solvers select and make illegal moves. To my knowledge this is the only model that includes an illegal move process model. According to Jeffries et al. it seems that participants consider illegal moves and then check them with the illegal move filter. This occurs after the move has been selected, but before the move is actually made. Most illegal moves will be discovered before they are made and a new move will be chosen. However, the process will not catch every illegal move and the problem solver may advance to an illegal state. For instance, if resource usage is at its limit then the problem solver may miscalculate the resulting state or he/she may never even initiate this filter at all. If resources are low, then decision-making ability may suffer. An example of this would be a teenager who just obtained a drivers' license and drives around talking on a cell phone with the radio blaring. The teenager's resources are

consumed by the cell phone and loud music and few resources are left for driving. The teenager may miscalculate a turn and cause an accident because there are insufficient resources available to calculate the correct angle for the turn. Jeffries et al. (1977) do not provide a process-based account of resource limitations; however, they assume that there is a fixed probability that the illegal move filter will be initiated. Jeffries et al. also state that if a move is not tested for legality then it will be chosen regardless of legality.

Jeffries et al. (1977) state that avoiding illegal moves also depends on how difficult it is to calculate the legality of a future state. In the Missionaries and Cannibals river crossing problem, there are two types of moves, easy-to-detect and hard-to-detect. Moves in which Cannibals outnumber Missionaries on the bank of the river with the boat are considered easy-to-detect. According to Jeffries et al. these moves will always be detected, because in their experiment participants were able to see the consequences on this bank before the move was made and they were able to correct their potential move. Hard-to-detect moves are those moves that place exactly one more Cannibal than Missionary on the bank of the river without the boat. These types of moves may not always be detected and may result in illegal moves. All other moves are easy-to-detect and should therefore always be detected, resulting in the selection of a new move.

In their 1977 paper, Jeffries and colleagues ran participants through four different isomorphs of the Missionaries and Cannibals river-crossing problem. Their intent was to observe how different representations of the task affected participants' performance. Jeffries et al. generated predictions of participants' performance based on their model for legal and illegal moves from each state for the four isomorphs and compared the predictions to participants' actual performance. Jeffries et al. obtained an $R^2 = .94$ for

legal and illegal moves combined, demonstrating that their model was able to predict many of participants' moves to legal and illegal states.

Another explanation of why people make illegal moves was proposed by Zhang and Norman (1994), who suggested that the difficulty of a task is a function of the number of rules represented internally versus externally. They suggested that problems are composed of internal and external representations, which together create the abstract structure of the problem. A rule is external if it does not need to be stated explicitly as a rule for the problem (e.g. if a ball cannot fit into a hole, then it does not need to be stated that the problem solver may not place the ball in the hole and the rule is external because the restriction does not need to be explicitly stated). A rule is internal if it must be stated explicitly in the instructions and retained in memory (e.g. if a ball can fit into a hole and it needs to be stated that this violates a rule then the rule is internal because it must be remembered by the problem solver). Zhang and Norman had participants work on various isomorphs of the Tower of Hanoi. They found that when the rules were presented externally versus internally, the problem became less difficult and participants tended to make fewer illegal moves. Zhang and Norman classified difficulty as the time to solve the problem, number of errors or illegal moves and number of steps to reach a solution (they did not dissociate between illegal moves and errors). This finding that changing the rules from internal to external decreases illegal moves also seems to support the claim of Jeffries et al. (1977) that memory load limitations increase illegal moves because making the rules external would also decrease memory load. In addition, Zhang et al. suggested that external objects not only act as aids for solving problems, but they also create a different representation of the problem.

Zhang and Norman (1994) presented five properties of external representations in their paper. (1) External representations can provide memory aids. (2) External representations can provide information that can be directly perceived and used without being interpreted and formulated explicitly. (3) External representations can anchor and structure cognitive behavior. (4) External representations change the nature of a task. (5) External representations are an indispensable part of the representational system of any cognitive task.

When Zhang and Norman (1994) manipulated the rules for the different isomorphs and changed them from internal to external, participants' performance greatly increased. The solution times, number of steps to solution and number of errors all improved as the rules became external. It seems that making the rules external removed the opportunity for participants to make illegal moves, thus increasing performance. According to Zhang and Norman, we could all become more efficient problem solvers if we altered the rules to make them all external. However, externalizing rules is not always an option because we often do not have control over the problem and if we did, altering the rules would likely change the problem and create a new one. Zhang and Norman's work demonstrates how problem representation can greatly influence the selection of illegal moves and the contribution of illegal moves to problem difficulty.

Kotovsky and Simon (1990) obtained additional supporting evidence that the type of information available to a problem solver and how the problem is represented influence problem difficulty. Kotovsky and Simon (1990) presented participants all the legal move options to an isomorph of the Chinese Ring Puzzle at each state. This manipulation reduced illegal moves and backtracking, decreased total number of moves

and decreased the total time required to solve the problem. Kotovsky and Simon had participants work on various digital isomorphs of the intensely difficult Chinese Ring Puzzle. They found that an isomorph without any information about the legal moves from each state, the No-Info isomorph, took approximately twice as long to solve and required many more moves than did the isomorph that offered the legal options at each state, the Lo-Info isomorph. Kotovsky and Simon also created an isomorph that not only displayed all legal moves, but also displayed the resulting state of each potential move from any given state; they called this the All-Info isomorph. They found that this confused participants, increasing the number of illegal moves made. However, giving an explanation of what the information actually meant helped participants lower the number of illegal moves made, increasing problem solving efficiency, but not to the level of participants in the Lo-Info condition.

Kotovsky and Simon found that the amount of information given to participants influenced the difficulty of the problem. Kotovsky and Simon (1990) showed that giving additional information about move legality decreased the number of illegal moves, but only if the information was understandable and not confusing. The Lo-Info condition simply told participants which moves were legal and which moves were not legal. The cost of using this information was very low and the information proved to be helpful. However, when all of the information about every move was presented, as in the All-Info condition, the cost of using the information increased because it was overwhelming and confusing and performance did not improve compared to the Lo-Info condition; in fact performance decreased even with an explanation on how to use the information. Additional information was helpful in decreasing illegal moves, but only to the extent

that the cost of using the information did not exceed the cost of making a move without it.

Many known determinants contribute to problem difficulty either by influencing problem solvers to select the wrong path or by increasing the amount of time and resources allotted to a problem. The representation of the problem and the information available to the problem solver influence problem difficulty (Kotovsky & Simon, 1990; Zhang & Norman, 1994; Simon & Hayes, 1976). Size of the problem space, ease of learning the rules, applying the rules, and retaining the rules, ambiguity of what constitutes a move, memory load limitations, and how the rules agree with real world knowledge are all known contributors to problem difficulty (Jeffries et al., 1977; Kotovsky, Hayes, & Simon, 1985; Simon & Hayes, 1976; Kotovsky & Simon, 1990; Hayes & Simon, 1977). However, the selection and the presence of illegal moves as a determinant of problem difficulty is the main focus of this research. The model of Jeffries et al. (1977) extended Atwood and Polson's (1976) model of problem solving by adding a process model, which assumes that illegal moves are selected and then evaluated by an illegal move filter. However, resource limitations may cause the illegal move filter to miscalculate a future state or it may cause the illegal move filter to never be initiated, resulting in the selection of an illegal move.

Purpose

The purpose of this paper was to explore the contribution of illegal and legal moves to problem difficulty. In Experiment 1, I replicated the findings of Jeffries et al. (1977) to show that both legal and illegal moves contribute to difficulty and that some states appear to contribute more to problem difficulty than others as displayed by an increase in illegal and legal moves in those states. I also attempted to discover what individual differences

contributed to participants' problem solving performance, especially illegal move making. In Experiment 2, I replicated the findings of Experiment 1 and in addition showed that increasing the cost of making an illegal move decreased the number of illegal moves in a river crossing problem, but not the number of legal moves (Experiment 2). This finding provides a possible alternative to the explanation given by Jeffries et al. on the selection of illegal moves, which they explained as arising from memory load limitations because I was able to show a decrease in illegal moves without manipulating the memory load.

In this paper, I also found support for Ericsson and Simon's (1993), review of evidence showing that when participants were instructed to think aloud while problem solving their performance remained relatively unaffected. Additionally, in Experiment 2 I found that instructing participants to provide additional information during verbal protocols did not significantly affect problem solving performance in any way. The additional instructions were given in an attempt to increase the amount of information elicited from the verbal protocols because as stated by Thomas (1974, pp. 258), "Some problems do not lend themselves as easily to protocol analysis. The Hobbits-Orcs problem is one of a class of problems in which an untrained subject's usual behavior is to make a fairly rapid series of moves with little verbal commentary."

CHAPTER 2 EXPERIMENT 1

The purpose of Experiment 1 was to determine when people make and consider illegal moves. That is, when do people fail to check the rules and proceed in making an illegal move, check the rules and reject an illegal move for a legal alternative, or check the rules, but still choose an illegal move? People may even fail to check a rule simply because they did not understand the rule (Kotovsky, Hayes & Simon, 1985). However, I was more interested in the cases where they did not check the rule even though the rules were well understood. Therefore, it was important that I was certain that participants fully understood the rules before beginning the target problem. Training on how to make moves and understanding what constituted an illegal move was included to ensure that this was the case.

Thinking Aloud

Determining if having participants solve a problem while thinking out loud affected problem solving ability was another important issue. Having participants think aloud helped provided insight into their thoughts and intentions as they worked on the problem. However, it was important to ensure that instructing them to think aloud did not affect their performance, so a control group that worked on the problem silently was included in the design. In Experiment 1, half of the participants solved the Hobbits and Orcs problem while thinking aloud the other half did so silently. Those participants instructed to think aloud had their voices recorded and their moves tracked for later analysis to help answer the foregoing questions.

Improvement

In addition, I was interested in how people change their performance throughout the problem. The particular performance indicators that I was interested in were how often the rules were checked, the number of illegal moves made, and how often an illegal move was considered and correctly rejected. I hypothesized that as participants gained experience with the problem they would improve their performance and make fewer illegal moves and also consider fewer illegal moves. However, it may have been that problem solvers did not change their behavior at all during the course of solving this problem. Participants may have consistently made illegal moves throughout the problem without showing any improvement at all. Jeffries et al.'s model assumes that there is a fixed probability that an illegal move will be checked for legality and a fixed probability that an illegal move will be correctly rejected after being checked for legality.

Individual Differences

Finally, I was interested in exploring what, if any, individual factors contributed to problem solving ability. This is important because it would be valuable to know and understand what factors contribute to problem solving ability because we may then be able to determine a priori what individual traits or skills make an efficient problem solver. After completion of the Hobbits and Orcs problem in Experiment 1, I had participants complete the operation span task (OSPAN) (Kane, Bleckley, Conway & Engle, 2001) to obtain a measure of working memory. Kane et al. have shown that working memory correlates with controlled attention tasks, such as eye movement and it seemed possible that controlled attention may be valuable in increasing problem efficiency on a task such as the Hobbits and Orcs problem. If controlled attention is a determinant in calculating future states and avoiding illegal moves then I hypothesize that

the OSPAN task would be negatively correlated to the number of illegal and legal moves. As working memory increases a participant's ability to calculate future states and correctly reject illegal moves and select the best alternative for legal moves should improve, decreasing the number of both illegal and legal moves.

I also took measures of participants' Need For Cognition (NFC) and Impulsivity. I predicted that as participants' NFC increased they would have a greater desire to do well on the task and they would exert more effort and show improved performance. Nair and Ramnarayan (2000) found that NFC positively correlated with participants' performance, in the form of improved sales and profit margin, in decision making and problem solving in a simulation of a business situation. I hypothesized that there would be a negative correlation between NFC and both illegal and legal moves. Participants with a higher NFC would have a greater need to do well on the task and would ultimately allot more resources to the task to decrease the number of illegal and legal moves.

I hypothesized that Impulsivity would be positively correlated to both illegal and legal moves. Participants with a high Impulsivity score would be more likely to use a trial and error strategy, trying moves without first calculating the resulting state. If participants were to randomly select moves without attempting to apply any reasoning or additional strategies, which I hypothesized those with high Impulsivity scores would do, they would show an increase in the number of both illegal and legal moves.

Methods

Participants

Participants were 62 undergraduates from the University of Florida who received course credit for their participation. All participants were above the age of 18. Ten participants proved unable to solve the problem in 20 minutes. They were assisted in

completing the problem and thus were not included in the analysis. This resulted in 52 participants in the analysis for Experiment 1.

Problem and Interface

The Hobbits and Orcs problem (which is also referred to as the Missionaries and Cannibals or River Crossing problem) consisted of one boat and six travelers, three of which were hobbits and three of which were orcs. All travelers began on the left bank of a river with the boat. The goal was to move all six travelers to the right bank of the river using the boat. However, the rules stated that the boat could only hold a maximum of two travelers at one time and at least one traveler was required in the boat for it to cross the river. The rules also stated that at no time could the orcs outnumber the hobbits on either bank of the river because the orcs would then kill the hobbits. A button was located on the display to reveal the rules. If at any time a participant forgot one of the rules they could click on the button to display three additional buttons, one for each of the three rules. Figure 2-1 displays the interface that participants saw after selecting one orc and after clicking the initial “Forget A Rule?” button, exposing the three additional buttons for each rule.

The problem was presented on a computer screen using a Visual Basic program. Participants used the mouse to click on and select the travelers and then they clicked on the boat to send the selected travelers to the other bank of the river. If the participants added too many travelers to the boat, allowed the orcs to outnumber the hobbits, attempted to move the boat with no travelers selected, or violated the rules in any other way they were notified via a message box and the illegal move did not occur.

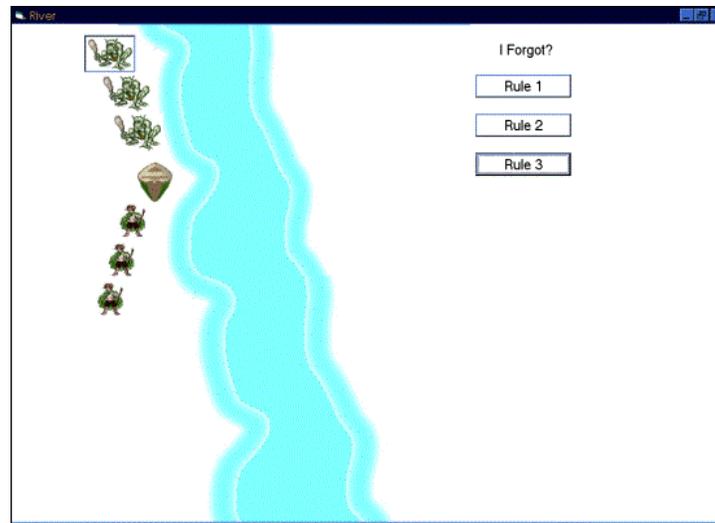


Figure 2-1. Above is the interface seen by participants in Experiment 1. In the initial state of the problem all six travelers began on the left bank of the river along with the boat. The top Orc is indented, indicating that it has been selected.

Procedure

Participants were randomly assigned to either the silent condition or the think aloud condition. Participants in the silent condition read a cover story for the Hobbits and Orcs problem and they also learned the rules. Before moving on, the participant was required to correctly recite all the rules from memory without error. Once the participant was able to recite the rules, the tutorial phase began. During the tutorial phase, the participant was shown an example problem on the computer and they were able to practice making moves with the mouse. During the tutorial phase the participant was instructed to click on the “Forget A Rule?” button and then to click on each individual rule button and asked to recite each rule aloud. Next, they were instructed to move one hobbit and one orc to the right bank and then back to the left bank. Then the participant received step-by-step instructions to violate each of the three rules. The participant completed the tutorial phase when they were able to correctly describe all three illegal moves. They also had to make moves on the example problem and violate all three of the rules after defining each one.

Next they had to make three legal moves and then they had to define the goal of the problem.

The participants were instructed that if they forgot a rule at any time they could click on the button in the top right corner to reveal the rules. Before beginning, the participants were asked to refrain from talking aloud while working on the problem. Then they were shown the target problem, and they were instructed to begin working on the problem.

In the think aloud condition participants first received think aloud training. They were given an operational definition of thinking aloud, similar to that of the description given by Ericsson and Simon (1993). More specifically, they were told to verbalize their thoughts and to say whatever came into their head, whether it made sense to a listener or not. Next, they practiced thinking aloud by imaging a familiar house and describing aloud everything they saw in the house. Then they read the cover story and completed the tutorial phase in the same fashion as those in the silent condition. Before beginning the problem, participants in the think aloud condition were informed that their voice would be recorded and they were reminded to think aloud. It was then explained that if they were silent for too long or if they were speaking too quietly, they would be reminded to keep talking. Next, they were asked to put on a set of headphones with a microphone. Then they were shown the target problem, and they were instructed to begin working on the problem.

In both conditions, testing began when the participant clicked on the mouse to initiate the program. In the think aloud condition, the experimenter tracked the considered then made illegal moves and the considered then not made illegal moves and their times

in minutes and seconds by attending to the verbal protocol and a timer, marking each occurrence on a score sheet.

If the participant was unable to solve the problem within the 20-minute time limit, they were then assisted in finishing the problem and their data were not included in the analyses. The maximum solution time of 20 minutes was chosen to restrict the session length to one hour.

If a participant was able to complete the Hobbits and Orcs problem without assistance before the 20 minute time restriction, then those participants were asked to complete the operation span (OSPAN) (Kane, Bleckley, Conway & Engle, 2001; Turner & Engle, 1989) assessment of working memory on the computer. They were also asked to complete a questionnaire on paper, which was used to obtain a need for cognition (NFC) (Cacioppo, Petty, Feinstein & Jarvis, 1996) and Impulsivity (Patton, Stanford & Barratt, 1995) score.

Illegal Moves Made and Considered

The focus of this experiment was on the illegal moves committed by participants. An illegal move was a violation of the third rule, which allowed the Hobbits to be eaten by the Orcs. Violations of Rule 1 or Rule 2, which involved trying to move the boat while it was empty or attempting to add more than two passengers to the boat, were considered errors (Jeffries et al., 1977). These moves were not analyzed. Figure 1-2 shows a map of the Hobbits and Orcs problem space, which includes all possible illegal moves from each state. The problem space is almost completely linear, allowing participants to either move forwards or backwards at each state. In most states there is only one move that will move the problem solver forward: however, there are two states, state 0 and state 9, where there are two moves that will take the problem solver closer to the goal.

In both Experiment 1 and 2, those participants who were in the think aloud condition had an experimenter attend to their verbal protocols. The experimenter coded all of the illegal moves, Rule 3 violations, which the participant considered, but did not make. An illegal move was classified as considered if and only if the experimenter was able to understand the exact move being considered, the participant realized that the move was illegal, and the participant rejected the move and did not make that move. This could be assessed by watching the participant's moves and attending to their verbal protocols. More specifically, considered illegal moves could be observed through removal of travelers from the boat and placement of the mouse over a traveler while the participant stated that they were not allowed to make that move, that this would violate the rules, or any other words that indicated that they knew that moving that specific traveler would result in a violation of Rule 3. Broad statements such as, "Any move I make will violate Rule 3," or "Anything I do will kill the Hobbits," were not counted as considered illegal moves; the statement had to refer to a specific traveler and a specific move.

Results and Discussion

Silent Versus Aloud Comparisons

A series of independent samples t-tests were conducted comparing the silent and think aloud groups to assess whether instructing participants to think aloud affected their performance. Ericsson and Simon (1993) reviewed evidence that instructing participants to think aloud had minimal effects on problem solving except that in some cases it increased the time needed to solve the problem. In my Experiment 1, I replicated the findings summarized by Ericsson and Simon; the analyses failed to detect any differences across conditions for total moves made, illegal moves made, legal moves made, illegal

moves made in the first half, and illegal moves made in the second half, all $t_s < 1$.

Proportion of total moves made which were illegal was also not significant $t(50) = 1.31$, $p = .196$. There were also no individual differences found between the two groups for OSPAN, NFC, and Impulsivity, all $t_s < 1$.

Ericsson and Simon previously reported that in some cases instructing participants to think aloud affected move and solution time. I was unable to detect any such effect in my first experiment. No differences were found between total time to solve the problem and average time per move, both $t_s < 1$. Due to technical problems, the total times and average times were not available for 18 of the 52 participants. As hypothesized, I was not able to detect any differences between the silent and think aloud groups on several different measures. In Experiment 1, the instruction to work on the problem while thinking aloud resulted in no detectable differences on problem solving ability or performance when compared to a control group that worked on the problem silently. Means and standard deviations are shown in Table 2-1.

Table 2-1. Silent and Think Aloud Comparisons, Experiment 1*

Factors	Silent	Think Aloud
Total Moves	34.81 (24.20)	34.88 (15.79)
Illegal Moves	5.19 (5.41)	5.69 (3.66)
Legal Moves	29.62 (20.64)	29.19 (13.99)
Illegal Moves in First	3.15 (3.06)	3.46 (2.08)
Illegal Moves in Second	2.04 (2.73)	2.35 (2.02)
Proportion Illegal	0.13 (0.10)	0.16 (0.08)
Total Time	6.06 (4.27)	6.31 (4.38)
Average Time	0.12 (0.04)	0.12 (0.06)
OSPAN	14.81 (7.80)	13.65 (6.13)
NFC	61.08 (11.62)	62.88 (10.68)
Impulsivity	62.69 (9.55)	64.92 (9.98)

*Means are located in the table along with standard deviations in parenthesis.

Improvement

Another important question was whether or not participants' performance changed or improved throughout the problem. The model of Jeffries et al. (1977) assumes that the probabilities for checking and selecting an illegal move are fixed so participants should not show any improvement. However, it seems plausible that as participants work on the problem they may be learning something about the problem and how to avoid illegal moves. If this were true, then participants might show improvement as they progress through the problem. The problem was segregated into the first and second half by calculating the total number of moves and by dividing by two; this resulted in the number of moves in both halves of the problem. I then calculated the number of illegal moves committed in the first half of the problem and in the second half of the problem.

For the following analyses I collapsed over the two groups and performed a paired samples *t*-test for all participants. I found that significantly more illegal moves were committed in the first half of the problem ($M=3.31$) when compared to the illegal moves made in the second half of the problem ($M=2.19$), $t(51) = 4.15$, $p < .001$. I also calculated the number of illegal moves considered and correctly rejected in each half of the problem for the think aloud group. Although the number of illegal moves considered was not significantly different, there was a trend in the right direction with a mean of 2.88 considerations in the first half and 2.08 considerations in the second half $t(24) = 1.60$, $p = .123$. These findings suggest that participants may have changed their performance as they progressed through the problem; they created and, although it was not statistically significant, considered fewer illegal moves in the second half of the problem. These findings do not support the assumptions made by Jeffries et al. that the probability of checking and rejecting an illegal move is fixed.

Individual Differences

I hypothesized that the individual difference measures obtained in this experiment would correlate with participants' problem solving performance. However, I was unable to detect any significant contributions of the three measures that I obtained. For this analysis and all following analyses of Experiment 1, I collapsed over the silent and think aloud conditions because no differences were detected between the two groups in previous analyses. The correlation coefficients for OSPAN and illegal moves, NFC and illegal moves, and Impulsivity and illegal moves were not significant. In addition, the correlations between OSPAN and legal moves, NFC and legal moves and Impulsivity and legal moves were also not significant. Correlations and significance scores are shown in Table 2-2. Unexpectedly, none of the three individual factors were significantly correlated with illegal or legal moves.

A significant correlation between working memory (WM), as assessed by the OSPAN task, and illegal or legal moves was not detected in the first experiment. The lack of a significant correlation does not support the claim made by Jeffries et al. (1977) that resource limitations are the cause of the selection of illegal moves (Jeffries et al. do not define resource limitations, it is assumed by the author that WM could be one such resource). If resource limitations are indeed the reason for the selection of illegal moves then those participants with a higher WM should have showed a decrease in the selection of illegal moves because they would have additional resources available to calculate future states correctly, resulting in a decrease in the number of illegal moves selected and a negative correlation between OSPAN and illegal moves.

Table 2-2. Correlations for Individual Differences, Experiment 1*

Factors	OSPAN	NFC	Impulsivity	Illegal Moves	Legal Moves
OSPAN	-----	.179 (.204)	-.249 (.075)	-.079 (.579)	.174 (.218)
NFC		-----	-.251 (.072)	-.132 (.352)	-.029 (.837)
Impulsivity			-----	-.004 (.976)	-.263 (.060)
Illegal Moves				-----	.531 (.000)
Legal Moves					-----

*Correlation coefficients are located in the table along with significance ratings in parenthesis.

The lack of a significant correlation between WM and legal moves is also unexpected, although to a lesser extent. Delaney, Ericsson, and Knowles (in press) have shown that when participants are instructed to plan they show improved performance in the form of fewer moves executed in reaching the goal. Theoretically, participants with higher WM could more easily hold states in memory as they plan and mentally explore future states, thus showing improved performance in the form of fewer legal moves. However, according to the models of Atwood and Polson (1976) and Jeffries et al. (1977) participants do not plan, so those participants with high WM would be less likely to show any benefits in the form of fewer legal moves executed. It is also important to note that the models of Jeffries et al. and Atwood and Polson argue that participants do not plan because they are unable, due to memory resource limitations. However, Delaney et al. have found evidence to dispute this claim showing that participants are able to plan their way to a solution in challenging multi-step water jug problems.

The lack of a significant correlation for NFC and illegal moves and NFC and legal moves could possibly indicate that the difficulty of the problem is resistant to motivational factors and participants' desires to do well. Those participants with high NFC scores are likely those participants that are willing to put forth the effort to do well on the problem, however the correlations were not significant.

I was also unable to detect a correlation between Impulsivity and illegal moves and Impulsivity and legal moves. I hypothesized that participants with high Impulsivity scores would have a positive correlation between both illegal moves and legal moves. I assumed that impulsive participants would be more likely to make both illegal and legal moves without considering the consequences of their moves. However, not only were the correlations not significant, but there was a trend for a negative correlation between Impulsivity and legal moves, $r = -.263, p = .060$. This means that participants with a high Impulsivity score were more likely to make fewer legal moves. This could be interpreted as the more impulsive participants having a greater desire to find novel states. Since the solution pattern of the Hobbits and Orcs problem is linear a participant could find the solution in the minimum number of moves by simply selecting a novel state every time. If those participants with higher Impulsivity scores were trying to find novel states to move to then they would have executed fewer legal moves resulting in a negative correlation.

State Versus State Comparisons

Another area of interest was determining if I would find the same patterns of results for a river crossing problem as those found by Jeffries et al. (1977). In their paper Jeffries et al. show that participants made more illegal and legal moves in specific states of the problem. When a paired samples t-test was conducted on the number of illegal moves created in each state, I found that there were significantly more illegal moves created in states 2 and 5 than in any other states. This replicated the findings of Jeffries et al. and Figure 2-2 demonstrates this pattern for both the current experiment and the results obtained by Jeffries et al.. However, it is also important to note that there were no illegal moves made in states 1B, 1, 3, 8, 9, 10, 10A, or 11. These eight states yielded the same

results, thus they are all reported as one. The results for Experiment 1 are listed in Table 2-3.

Table 2-3. State vs State Comparisons, Experiment 1*

Pair	Mean (Standard Dev.)	t (53)	Significance
State 2	2.70 (2.68)		
State 0	0.17 (0.38)	7.12	.000
State 1A	0.74 (1.05)	5.87	.000
State 4	0.35 (0.73)	6.39	.000
State 6	0.22 (0.60)	6.54	.000
State 7	0.35 (0.65)	6.74	.000
States 1B, 1, 3, 8, 9, 10A, 11	0.00 (0.00)	7.43	.000
State 5	1.26 (1.67)		
State 0	0.17 (0.38)	4.56	.000
State 1A	0.74 (1.05)	2.25	.028
State 4	0.35 (0.73)	4.54	.000
State 6	0.22 (0.60)	4.11	.000
State 7	0.35 (0.65)	3.83	.000
State 1B, 1, 3, 8, 9, 10A, 11	0.00 (0.00)	5.53	.000

*The top half of the table is a comparison between the number of illegal moves made in State 2 and all the other states. The bottom half of the table is a comparison between illegal moves made in State 5 and all the other states.

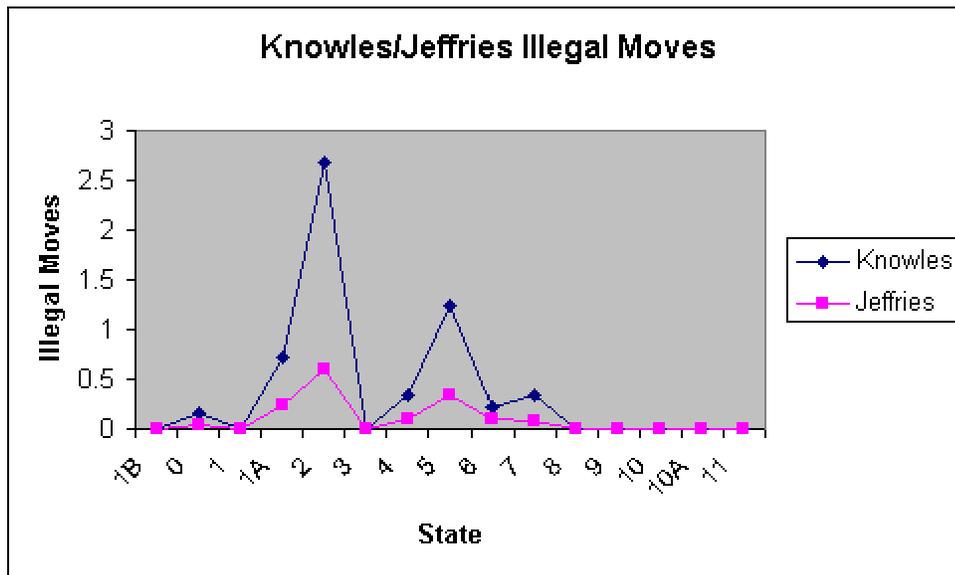


Figure 2-2. The Figure shows a comparison of illegal moves made in Experiment 1, by State and the illegal moves made by participants in Jeffries et al. (1977).

In addition, when a paired samples t-test was conducted on the number of illegal moves created in state 2 versus the number of illegal moves created in all the other states combined, excluding state 5, I found that there were significantly more illegal moves created in state 2 ($M = 2.70$) than all other states combined ($M = 1.83$), $t(53) = -2.48$, $p = .016$.

Although the results of Experiment 1 yielded the same pattern of results as obtained by Jeffries et al. (1977), their participants created fewer illegal moves in the Hobbits and Orcs isomorph. While participants in Experiment 1 of this study had a mean of 5.44 illegal moves those in the study of Jeffries et al. had a mean of 2.75 illegal moves. Some factors that may have contributed to this difference may have been the dissimilarities in the interfaces, sampling issues, and presence of experimenter in the room. In their interface participants were able to assess the result of the potential move on the bank with the boat and change the move before the move was made. This occurred because after selecting a traveler that traveler was removed from the bank and then it appeared at the bottom of the screen allowing participants to change their move after assessing the result on the bank with the boat. However, in my interface when participants selected a traveler they did not move, they became indented to show that they were selected, they did not leave the bank were they were positioned. Their interface allowed for more evaluation of each move, more effort to initiate a move because the program was not mouse driven, and more feedback as to the result of each move, which all may have contributed to a decrease in the number of illegal moves. In addition, all of my participants were recruited from the University of Florida and participated to complete a course requirement, whereas some of Jeffries and colleagues' participants were from the University of

Colorado and some of their participants were recruited through a newspaper ad and paid for their participation. Paid participants may have felt more of an obligation to do well on the problem since they were receiving a monetary reward for their participation. A third difference was that the experimenter in their study left the room while the participant worked on the problem and the participant was able to call the experimenter via a button if they needed anything. In my experiment, the experimenter remained in the room the entire time. Thus the participant may not have been able to relax and concentrate as well as the participants in the study conducted by Jeffries et al..

CHAPTER 3 EXPERIMENT 2

In Experiment 1, I obtained a pattern of participants' performance on the Hobbits and Orcs problem that was similar to that obtained by Jeffries et al.. I was also unable to detect any changes in participants' performance when they were instructed to think aloud as they worked on the problem. In addition, I found that participants made fewer illegal moves in the second half of the problem, possibly indicating that participants were learning and showing some improvement as they worked on the problem.

In Experiment 2, I increased the cost of making an illegal move to assess whether this manipulation would enable participants to improve their performance to a greater extent. If participants considered illegal moves and correctly rejected them more often when the cost of making an illegal move increased and they made fewer illegal moves, then this may provide an explanation for illegal move selection in addition to the memory load limitation hypothesis proposed by Jeffries et al. (1977). According to the memory load limitation hypothesis, people select illegal moves because they do not have resources available to correctly assess the resulting state. However, in both a low cost and high cost scenario the memory load should be very similar and should not affect performance between the two groups. Improved rule checking and illegal move rejection may lend support to the idea that participants simply forget to check the rules or they do not view checking the rules as important or necessary. It may be that the problem seems

too simplistic and checking the rules under a low cost situation seems wasteful because a trial and error solving technique may lead to a solution with minimal effort. However, in a high cost situation, where the problem solver receives a 30 second penalty after each illegal move, solving the problem using a trial and error strategy would take a great amount of time and it would be more costly to not check the rules.

In Experiment 2, I also set out to replicate and extend the findings of Experiment 1 that instructing participants to think aloud did not change their problem solving performance. However, in addition to having participants think aloud I implemented a new tool in obtaining verbal protocols in an attempt to increase the accuracy of the number of illegal moves considered by participants. Specifically, participants were instructed and received training on what to say whenever they considered certain moves. Participants were told that whenever they considered a move, even if they knew that they were not going to select that move, they were to say aloud all the travelers involved in that move. The purpose of these additional instructions and training were to increase the amount of information obtained through the verbal protocols. Showing no differences in performance for the silent and advanced think aloud groups with increased instruction and training for verbal protocols resulting in more information from the thoughts of the problem solver would indicate that more information can be obtained through verbal protocols without disrupting or changing problem solving behavior.

Methods

Participants

Participants were 90 undergraduates from the University of Florida who received course credit for their participation. All participants were above the age of 18. Six

participants proved unable to solve the problem in 20 minutes. They were assisted in completion of the problem and thus they were not included in the analysis.

Twenty participants were also not included in the analysis because they were run at the beginning of the semester and were classified as pilot data. These participants were not randomly distributed into the different conditions. The performance of these participants was significantly better than later participants, possibly due to motivational factors (because participants voluntarily sign up for experiments to complete a course requirement). This resulted in 64 participants in the analysis for Experiment 2.

Problem and Interface

The problem and the interface were similar to those used in Experiment 1 with a few minor exceptions. The boat was now located at the bottom of the screen and not in the middle, between the travelers. After initially clicking the travelers, they would appear at the bottom of the screen next to the boat. This was done to allow the participants the ability to view the result of their current move on the bank of the river where the boat was located because the selected travelers were now removed from the current bank and placed at the bottom next to the boat. This manipulation created an interface more similar to that used by Jeffries et al. (1977) because in both cases participants could assess the result of the potential move on the bank with the boat and change the move before the move was made. A display of the interface for Experiment 2 is shown in Figure 3-1. In the Figure, two Hobbits have been selected and they now appear next to the boat. The participant can easily see and determine that one Hobbit is left on the left bank with three Orcs, which will kill him. However, the participant can remove the travelers from the boat at the bottom of the screen at any time before clicking on the boat to complete the move.



Figure 3-1. This Figure shows the interface that participants saw in Experiment 2. Two hobbits have been selected from the initial state of the problem and the result of the move can be assessed on the left bank before the boat is clicked, completing the move.

Design

The design of this experiment was 2 Aloud (silent vs. aloud) x 2 Cost (no-cost vs. cost). Both variables were between-subjects.

Aloud variable

Participants were randomly assigned to either work on the problem silently or while thinking aloud, just as in Experiment 1.

Cost variable

Participants were randomly assigned to either a condition where there was no cost for violating Rule 3 or a condition where violating Rule 3 resulted in a penalty. Rule 3 states that if the Orcs outnumber the Hobbits on either bank of the river the Orcs will then kill the Hobbits. In the no-cost condition, if a participant violated Rule 3 they were notified via a message box and then they were allowed to continue working on the problem, as in Experiment 1. In the cost condition, if a participant violated Rule 3, then the screen turned black, except for some brief instructions, a text box, and a button

labeled “Go.” The participant was instructed to click on the “Go” button and words would then appear on the screen. Every time a new word appeared they were to say aloud a number between one and five, rating the pleasantness of each word where one was very unpleasant and five was very pleasant and three was neutral. After every Rule 3 violation in the cost condition, ten words appeared with a new word appearing every three seconds. After completing the cost task participants were returned to the problem where they last committed the violation.

Procedures

The procedures followed that of Experiment 1 except for some minor changes listed below. Those in the think aloud condition practiced thinking aloud by imagining that they were leaving school for the day and they were to describe the path that they take home and everything they see along the way, instead of describing a house.

After completing the tutorial, pictures of Hobbits and Orcs were displayed on the computer screen with different combinations of the travelers appearing every 1.2 seconds for a total of 14.4 seconds. Participants in the silent conditions were told that this was to help them better understand the characters so that they could follow the rules more easily. However, the main purpose of the task was to get those in the think aloud conditions to talk more often while working on the task and to refer to the characters by their proper names so that the experimenter could obtain a better sense of the some of the moves that the participants were considering. After completing this short task all participants were instructed to continue working on the practice problem. Those in the silent condition did so quietly, while those in the think aloud condition did so while thinking aloud. After solving the practice problem or after a minute and thirty seconds, whichever came first, the participants in the cost conditions were informed of the cost for violating Rule 3.

After the instructions those in the cost condition were instructed to begin working on the problem. Those in the no cost conditions did not receive these instructions and were instructed to begin working on the problem after the practice problem. As in Experiment 1, those participants in the think aloud condition were instructed to wear headphones with a microphone attached. As in Experiment 1, the experimenter attended to the participant's protocols, taking note of illegal made and illegal considered moves. In addition, the experimenter also marked whether the considered moves were easy-to-detect (violation of Rule 3 on the bank of the river where the boat is currently or the orcs outnumbering the hobbits by more than one) or hard-to-detect illegal moves (violation of Rule 3 on the opposite bank from where the boat currently is and the orcs outnumber the hobbits by only one) (Jeffries et al., 1977).

As in Experiment 1, if the participant was unable to solve the problem within the 20-minute time limit, they were then assisted in finishing the problem and their data were not included in the analyses. The maximum solution time was chosen to restrict the session length to one hour.

After completion of the Hobbits and Orcs problem, participants in all conditions were asked to complete the OSPAN (Kane, Bleckley, Conway & Engle, 2001; Turner & Engle, 1989) task on the computer. However, due to the null findings of Experiment 1, participants were not instructed to fill out a questionnaire to assess NFC or Impulsivity. The OSPAN task was included in Experiment 2 to replicate the findings of Experiment 1 due to the ubiquitous findings of correlations between working memory assessment tasks and other problem solving tasks (Kane et al., 2001; Turner & Engle, 1989).

Results and Discussion

A 2 x 2 factorial ANOVA was conducted for the following analysis.

Silent Versus Aloud Comparisons

In Experiment 2, I instructed participants to state all the moves that they considered in an attempt to obtain more elaborate verbal protocols. This was a new method so it was important to verify that it did not change participants' performance. In subsequent analyses I collapsed over the silent and silent cost conditions and the think aloud and think aloud cost conditions so that I could directly compare any differences between those participants who worked on the problem silently and those participants who thought aloud as they worked on the problem. As hypothesized, I replicated the findings of Experiment 1 and found no differences between the two groups for total moves made $F(1,60) = 1.53$, $MSE = 281.69$, $p = .221$, legal moves made $F(1,60) = 1.78$, $MSE = 227.28$, $p = .187$, or total time to complete the problem $F(1,60) = 1.02$, $MSE = 18.87$, $p = .316$. I also found no difference for average time per move, illegal moves made, Operation Span (OSPAN), or proportion of illegal moves made, all $F_s < 1$. Means and standard deviations for the silent and think aloud groups are reported in Table 3-1 and the graph in Figure 3-2 displays the illegal moves made by participants in the silent and think aloud conditions.

Table 3-1. Silent and Think Aloud Comparisons, Experiment 2*

Factors	Silent	Think Aloud
Total Moves	32.63 (17.80)	27.44 (15.86)
Illegal Moves	4.41 (3.71)	3.94 (4.09)
Legal Moves	28.22 (16.29)	23.19 (13.54)
Proportion Illegal	0.14 (0.09)	0.13 (0.08)
Total Time	6.90 (4.28)	5.80 (4.29)
Average Time	0.22 (0.10)	0.21(0.09)
OSPAN	14.16 (5.60)	13.38 (6.69)

*Means are located in the table along with standard deviations in parenthesis.

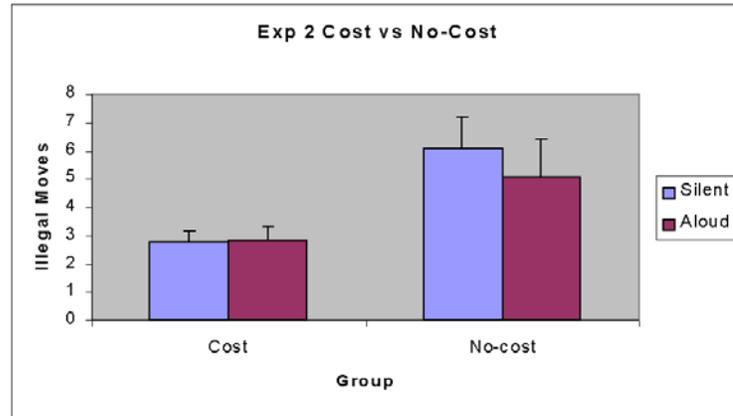


Figure 3-2. The graph shows a comparison between the cost and no-cost groups for illegal moves committed in Experiment 2.

Cost Versus No-Cost Comparisons

In the following section I collapsed over the silent cost and think aloud cost conditions and the silent no cost and think aloud no cost conditions so that I could directly compare any differences between those participants who received a penalty after violating Rule 3 and those who did not receive a penalty. I found no significant differences for total moves $F(1,60) = 1.92$, $MSE = 281.69$, $p = .171$, or OSPAN scores $F(1,60) = 1.99$, $MSE = 37.40$, $p = .164$. I also found no significant differences for number of illegal moves considered, legal moves made, total time to complete the problem, and number of easy-to-detect illegal moves made, all $F_s < 1$.

However, total illegal moves $F(1,60) = 9.11$, $MSE = 13.59$, $p < .005$, was significantly different indicating that those in the cost condition made fewer illegal moves than those in the no-cost condition. Figure 3-2 shows the differences between the cost and no-cost groups for number of illegal moves made. Proportion of illegal moves $F(1,60) = 4.67$, $MSE = 0.01$, $p < .05$, was also significant demonstrating that even though there was no difference between the total number of moves those in the cost condition

made a smaller proportion illegal moves. Average time per move $F(1,60) = 4.59$, $MSE = 0.01$, $p < .05$, was also significant suggesting that those in the cost condition were possibly more cautious and that they took significantly longer to make each move. Hard-to-detect illegal moves approached, but did not reach significance $F(1,60) = 3.58$, $MSE = 6.31$, $p = .069$. Means for the cost and no-cost groups can be found in Table 3-2.

Table 3-2. No-Cost and Cost Comparisons, Experiment 2*

Factors	No-Cost	Cost
Total Moves	32.94 (18.55)	27.13 (14.85)
OSPAN	12.69 (6.35)	14.84 (5.79)
Total Time	6.12 (4.11)	6.59 (4.50)
Illegal Moves Considered	5.50 (6.21)	6.81 (7.70)
Legal Moves	27.06 (15.92)	24.34 (14.30)
Easy-To-Detect Made	1.69 (2.47)	1.25 (1.13)
Illegal Moves	5.56 (4.86)	2.78 (1.72)
Proportion Illegal	0.16 (0.10)	0.11 (0.06)
Average Time	0.19 (0.08)	0.24 (0.10)
Hard-To-Detect Made	3.38 (3.16)	1.69 (1.49)

*Means are located in the table along with standard deviations in parenthesis.

Interactions

The results of the ANOVA revealed that there was not a significant interaction between Aloud and Cost for OSPAN $F(1, 60) = 1.00$, $MSE = 37.40$, $p = .321$. In addition, the interactions between Aloud and Cost for total moves, illegal moves, legal moves, proportion of illegal moves, total time to complete the problem, and average time per move were not significant, all $F_s < 1$.

Improvement

In this section I followed the same procedures as indicated in Experiment 1. Because there were no differences I collapsed over the silent and think aloud groups and did separate paired samples t -tests for the cost and no-cost groups comparing illegal moves made in the first half of the problem and illegal moves made in the second half of the problem. I also did the same analysis for the cost and no-cost groups for illegal moves

considered in the first and second half of the problem. The results indicated for the cost group that there were significantly more illegal moves created in the first half of the problem ($M=1.81$, $SD = 1.20$) than in the second half of the problem ($M=1.00$, $SD = 0.95$), $t(31) = 3.52$, $p = .001$. Although there appeared to be a trend in the same direction for moves considered in the cost condition the result was not significant $t(15) = 1.97$, $p = .073$ with a mean of 3.94 ($SD = 4.28$) in the first half and 2.88 ($SD = 3.70$) in the second half. When the same analysis was done for the no-cost condition the results revealed that there was not a significant difference for more illegal moves created in the first half ($M = 3.06$, $SD = 2.54$) than in the second half ($M = 2.53$, $SD = 2.72$), $t(31) = 1.44$, $p = .161$. However, there was a significant effect of more illegal moves considered in the first half ($M = 3.50$, $SD = 3.78$) than in the second half ($M = 2.00$, $SD = 2.73$), $t(15) = 2.70$, $p = .016$.

Individual Differences

As a result of the null findings between the silent and think aloud groups in both Experiment 1 and Experiment 2, I collapsed over the two groups for the following analysis so that I could further assess any differences between the correlations for the cost and no-cost groups. For both the cost and no-cost groups I was unable to find any correlation between OSPAN and illegal moves, OSPAN and considered moves or OSPAN and legal moves. For those in the cost condition I obtained a correlation coefficient of $r = -.020$, $p = .915$ for OSPAN and illegal moves, a correlation coefficient of $r = -.191$, $p = .479$ for OSPAN and illegal moves considered and a correlation coefficient of $r = -.215$, $p = .237$ for OSPAN and legal moves. For those in the no-cost condition I obtained a correlation coefficient of $r = -.163$, $p = .375$ for OSPAN and illegal moves, a correlation coefficient of $r = .169$, $p = .531$ for OSPAN and illegal moves

considered and a correlation coefficient of $r = -.088$, $p = .634$ for OSPAN and legal moves. In Experiment 2, I replicated the findings of Experiment 1 that working memory does not significantly correlate with participants' selection or consideration of illegal moves.

State Versus State Comparisons

The state vs. state comparison replicated the findings of Experiment 1, that is, there were significantly more illegal moves created in states 2 and 5 when they were individually compared to all other states of the problem. For this analysis I collapsed over the silent and think aloud groups and conducted paired samples t-test for both the cost and no-cost groups comparing states 2 and 5 to all other states in the problem. Just as in Experiment 1 there were no illegal moves committed in states 1B, 1, 3, 8, 9, 10, 10A, or 11. For the analysis these eight states are reported as one since they yield the same result. Means and t-values for the cost group for this analysis are located in Table 3-3 and means and t-values for the no-cost group are located in Table 3-4.

Table 3-3. Cost Group State vs State Comparisons, Experiment 2*

Pair	Mean (Standard Dev.)	<i>t</i> (30)	Significance
State 2	1.45 (1.12)		
State 0	0.10 (0.30)	6.97	.000
State 1A	0.26 (0.58)	5.11	.000
State 4	0.19 (0.48)	5.32	.000
State 6	0.03 (0.18)	6.88	.000
State 7	0.06 (0.25)	6.92	.000
States 1B, 1, 3, 8, 9, 10A, 11	0.00 (0.00)	7.21	.000
State 5	0.65 (0.80)		
State 0	0.10 (0.30)	3.44	.002
State 1A	0.26 (0.58)	2.11	.043
State 4	0.19 (0.48)	2.37	.024
State 6	0.03 (0.18)	4.05	.000
State 7	0.06 (0.25)	3.82	.001
State 1B, 1, 3, 8, 9, 10A, 11	0.00 (0.00)	4.50	.000

*The top half of the table is a comparison between State 2 and all the other states. The bottom half of the table is a comparison between State 5 and all the other states.

Table 3-4. No-Cost Group State vs State Comparisons, Experiment 2*

Pair	Mean (Standard Dev.)	<i>t</i> (30)	Significance
State 2	1.64 (1.90)		
State 0	0.27 (0.67)	3.89	.000
State 1A	0.79 (1.02)	2.65	.012
State 4	0.52 (1.03)	3.02	.005
State 6	0.36 (0.70)	4.05	.000
State 7	0.09 (0.38)	4.70	.000
States 1B, 1, 3, 8, 9, 10A, 11	0.00 (0.00)	4.95	.000
State 5	1.85 (2.21)		
State 0	0.27 (0.67)	4.13	.000
State 1A	0.79 (1.02)	2.82	.008
State 4	0.52 (1.03)	4.11	.000
State 6	0.36 (0.70)	3.70	.001
State 7	0.09 (0.38)	4.49	.000
State 1B, 1, 3, 8, 9, 10A, 11	0.00 (0.00)	4.81	.000

*The top half of the table is a comparison between State 2 and all the other states. The bottom half of the table is a comparison between State 5 and all the other states.

CHAPTER 4 GENERAL DISCUSSION

Problem difficulty has been evaluated in the past and many contributing factors have been noted. However, as yet little work has been done to assess the contribution of illegal moves to problem difficulty. In this paper, I sought to learn more about what factors influence problem difficulty. More specifically, I was interested in the role that the consideration and selection of illegal moves plays in problem difficulty. Two experiments were conducted to further assess to what extent the presence of illegal moves affects problem difficulty.

To assess when participants considered and rejected illegal moves, I had them think aloud as they worked on the problem. This technique enabled me to gain insight into the thoughts participants engaged in as they worked on the problem. Experiment 1 and Experiment 2 both included a control group that worked on the problem silently, this way I was able to assess if instructing participants to think aloud changed their problem solving performance in any way.

Thinking Aloud

In both experiments, I showed that instructing participants to think aloud as they worked on the problem did not significantly change or affect their problem solving performance. This lends support to Ericsson and Simon's (1993) argument that when participants are instructed to think aloud performance remains relatively unaffected. In the second experiment I implemented a new technique in an attempt to increase the output of participants' verbal protocols in a task that does not lend itself very well to

verbal protocols (Thomas, 1974). Participants were instructed to state every move they considered even if they knew that the move was incorrect and that they would not execute that move. I also had participants practice saying aloud different combinations of hobbits and orcs as they appeared at a fast rate on the computer screen so that they would know what to say when they were considering the different moves. The added instructions and training did not affect participants' performance indicating that it may be possible to probe deeper into problem solver's thoughts as they work on difficult problems without affecting their performance.

Individual Differences

Individual difference measures were taken in Experiment 1 to assess if there were significant contributions of WM, NFC, or Impulsivity to problem solving efficiency. Although I was hopeful that the results would indicate a contribution of at least some of the factors, I was unable to detect a significant relationship between any of these factors and performance on the task. The lack of significant correlations may possibly indicate that certain aspects of problem difficulty may be resistant to individual differences of specific skills, traits, and motivational factors. However, the lack of a significant correlation between WM and illegal moves may also indicate a flaw in the model of Jeffries et al. (1977). The model of Jeffries et al. makes specific claims about how illegal moves are selected and executed. If it were true that, as their model assumes, illegal moves are selected because of resource limitations, then we would have expected negative correlations between measures of cognitive resources (like working memory capacity) and illegal moves. However, no such relationship emerged.

The lack of a correlation between WM and illegal moves could also possibly be explained by the overwhelming difficulty of the task that affects both participants with

low and high WM, resulting in the lack of a significant correlation. Although this explanation seems plausible, it is not supported by participants' performance on the task. If all participants, those with both high and low WM, were overwhelmed by the difficulty of the task then we would have expected all participants to execute multiple illegal moves on the problem. However, this was not the case. Approximately six percent of participants in both Experiment 1 and 2 solved the problem without executing any illegal moves, 22 percent solved the problem with one illegal move or fewer and approximately 37 percent solved the problem with two illegal moves or fewer. Due to the low number of illegal moves executed by these participants it seems that the Hobbits and Orcs problem was not overwhelmingly difficult for all of the participants.

Illegal Move Selection

Jeffries et al. argued that illegal moves are selected and then subjected to an illegal move filter, which checks the selected move for legality. They also claimed that if a participant is performing at their resource limitation then the participant may fail to check the move for legality or he/she may miscalculate the resulting state and select an illegal move (Jeffries et al. do not define "resource limitations" so it was assumed that WM could be classified as a "resource"). However, from Jeffries et al.'s description of their model, it seems that high WM span would improve performance because additional resources would be available to check and correctly reject an illegal move. If this were true then I would have expected to obtain a significant negative correlation between WM and illegal moves. However, there was no such indication of a correlation in either Experiment 1 or Experiment 2.

The lack of a correlation between WM and illegal moves could possibly be due to inadequate power to detect the correlation. However, it is believed by the author that it is

more likely that the lack of a significant correlation was obtained because other determinants have a greater influence on participants' performance.

In Experiment 2, participants in the cost condition received a penalty for each illegal move, and this manipulation changed in participants performance. In the cost and no-cost conditions, the interface did not differ in any way, other than the penalty manipulation in the cost condition, and there were no known differences between the two groups in the amount of memory load imposed or the amount of WM required to solve the problem. The model of Jeffries et al. would predict that the penalty manipulation should not have affected the number of illegal moves committed by participants because it did not change the amount of resources needed to execute a move. However, the cost group chose and executed significantly fewer illegal moves ($M=2.78$) than participants in the no-cost group did ($M=5.56$). If resource limitations were the sole reason for the selection of illegal moves then the cost manipulation should not have affected the number of illegal moves executed. This finding of an improvement in participants performance when the cost of making an illegal move increases may indicate that there is an additional determinant to explain why participants select illegal moves.

An explanation for the selection of illegal moves could be the result of the problem solver's lack of an intuitive sense to plan. O'Hara and Payne (1998) found that increasing the cost of operator implementation increased participants' performance in the form of reduced solution lengths. Through verbal protocols, O'Hara and Payne were able to determine that when the cost of operator implementation increased, participants engaged in more planful search resulting in more efficient problem solutions. Despite the claim made by Jeffries et al. (1977) that participants do not plan because they do not have

the memory resources to do so, Delaney et al. (in press) showed that when instructed to, participants planned their way to a solution resulting in more efficient solutions.

However, there is no known evidence of participants planning in a river crossing or water jugs task without instruction to do so or without any other manipulation such as increasing the cost of making a move. I propose an additional determinant to the cause of the selection of illegal moves could be participants' lack of planning when the cost of making moves without planning is low. In the problem, if a participant violated Rule 3 they were notified via a message box and after clicking a button they were allowed to continue working on the problem from where they last chose the illegal move. The cost of making an illegal move was minimal so participants drive to check the move before initiating it was also minimal and probably not worth the effort. In this situation the cost of planning would be greater than the cost of selecting an illegal move and participants would not engage in planning without instructions to do so.

Although illegal moves have not been widely studied in the past, it is apparent here that they play a significant role in the contribution to problem difficulty. The selection of an illegal move increases the time and effort required to solve a problem. Illegal moves increase problem difficulty by leading the problem solver away from the goal increasing the total number of moves required to solve the problem.

Although I was not able to directly demonstrated it in this study, it is believed that even when illegal moves are not selected they still contribute to problem difficulty when they are considered. This is believed to be true because the consideration of illegal moves takes time and resources that increase the amount of effort and time needed to solve the problem, thus contributing to problem difficulty. Additional research is needed to further

assess the magnitude of the role played by the consideration of illegal moves in problem difficulty.

Legal Moves

Jeffries et al. (1977) argue that participants do not plan because they are unable due to memory resource limitations. If this were true then you would not expect to find a relationship between WM and legal moves, because participants would not show improved selection of legal moves if they were not planning and exploring multi-step moves. In both Experiments 1 and 2, I found no evidence of a relationship between WM and legal moves, which appears to support Jeffries et al.'s claim that participants do not plan in this task. Jeffries et al. argued that participants do not plan due to memory resource limitations. However, O'Hara and Payne (1998) and Delaney, Ericsson, and Knowles (in press) have found evidence to dispute this claim. O'Hara and Payne were able to determine through verbal protocols that when the cost of operator implementation increased, participants engaged in more planful search resulting in more efficient problem solutions. Delaney et al. showed that participants were able to successfully plan their way to a solution in a challenging multi-step water jugs problem when they were instructed to do so. Planning increased problem solving efficiency in a reduction in the number of legal moves needed to reach the solution. Research by O'Hara and Payne (1998) and Delaney et al. are evidence that planning can increase problem solving efficiency by increasing the quality of legal move selection. They also showed that participants are not restricted from planning by memory limitations, but possibly only because they do not intuitively engage in this strategy when the cost of making a move is very low, as it is in Experiment 1 and in the no-cost group in Experiment 2.

The above findings demonstrate the contribution of legal moves to problem difficulty. Legal moves can increase the amount of time and effort problem solvers spend on a problem, which increases problem difficulty. However, problem difficulty can be reduced when problem solvers engage in planning or other strategies that reduce wasted moves and wasted effort.

Improvement

Jeffries et al.'s model makes the assumption that there is a fixed probability that a move will be checked for legality and a fixed probability that an illegal move will be correctly rejected. According to this assumption it would be predicted that the number of illegal moves executed by participants would stay consistent throughout the problem. However, the findings of both Experiment 1 and 2 indicate that as participants work on the problem their behavior improves as experience with the problem increases. Simon and Reed (1976) found a similar finding when they had participants solve the Missionaries and Cannibals problem twice in succession and participants showed a decrease in the total number of legal moves needed to reach the solution.

In Experiment 1 and in the cost condition of Experiment 2, it was found that participants executed fewer illegal moves in the second half of the problem than in the first half. In Experiment 1 and in the cost condition of Experiment 2, the results also indicated that although there was a trend in the right direction participants did not significantly consider fewer illegal moves in the second half of the problem. In the no-cost condition in Experiment 2, although it was not significant, there was a trend in the right direction for participants to execute fewer illegal moves in the second half and a significant difference demonstrating that participants considered fewer illegal alternatives in the second half. I believe that the lack of significant findings for illegal moves

considered in Experiment 1 and in the cost condition of Experiment 2 and the lack of significance for illegal moves in the no-cost condition of Experiment 2 are the result of inadequate power to detect such differences. This is believed because both illegal moves and illegal moves considered in the second half are significantly different from those in the first half when the analysis is run after collapsing over the cost and no-cost group, thus increasing the power. In addition, the results are the same for those in Experiment 1 and those in the cost condition of Experiment 2 even though it would be expected that the no-cost condition would be the same as those in Experiment 1 because these groups are more similar. This could be explained through sampling errors and lack of power to detect significance in all cases.

The evidence of improved performance obtained in this study does not seem to support the claim made by Jeffries et al. that illegal move checking and illegal move rejection occur with fixed probabilities. Since participants are thinking about illegal moves less often, as displayed by a decrease in illegal moves considered, this may indicate that they have a better representation and understanding of the problem, which allows them to avoid illegal moves without even considering them. In addition, the number of illegal moves made also decreased in the second half of the problem; however it is difficult to determine if participants are improving at rejecting illegal moves or if this is just a result of fewer illegal moves being considered. The ratio of the number of illegal moves made and considered in the first half as compared to the ratio in the second half did not differ. This indicates that participants may actually be improving on both their ability to check and their ability to correctly reject illegal moves.

As participants gain experience from working on the task, they may learn or gain different techniques or strategies for avoiding illegal states. They may recognize specific states as being illegal and previously visited, which they now avoid because they know they are illegal. They may switch to a depth first search strategy after attempting several different moves and this would be beneficial in this task because the problem space is almost completely linear with only one solution path. Participants may realize that the problem is difficult and they may begin to plan, which I believe is possible based on the work of Delaney et al. (in press). The representation of the problem itself may change and improve for the problem solver resulting in a deeper understanding of the problem and improved performance. The above explanations for improved performance are speculation and it is possible that some, none, or all of them may influence the problem solvers performance.

State Differences

In this study and in Jeffries et al. (1977), specific states received more visits and produced more illegal moves compared to other states of the same problem. The two states that received the most visits and produced the most illegal moves were states 2 and 5, which can be seen in the map of the problem space in Figure 1-2.

It is possible that it could be argued that states 2 and 5 only produce more illegal moves because they are visited more often. However, states 2 and 5 have a higher proportion of the number of illegal moves made to number of times visited than any of the other states. The proportion of illegal moves made to the number of visits is 49% for state 2 and 45% for state 5. The next state with the highest proportion had 29% of the moves illegal and then 18% for the next highest state. It was also observed that some participants had a higher number of illegal moves than visits in states 2 and 5. This

occurred when a participant made multiple illegal moves before leaving the state. This evidence supports the claim that the selection of illegal moves makes a significant contribution to problem difficulty.

In state 2 of the problem there are two possible illegal alternatives, it is also one of the only states where there are two routes to backtrack away from the goal. In addition, if a participant does not look ahead or does not do at least some minimal planning then the correct move of transporting two orcs to the right bank seems like a dead end. The correct move may appear like a dead end without planning because participants may not realize that one orc would return to the left bank and if there are three orcs on the right bank and the boat can only hold two travelers then any move of the hobbits to the right bank would end with the hobbits being outnumbered on the right bank.

In state 5 of the problem there are three illegal alternatives and it is the only move in the problem, as noted by Jeffries et al. (1977), where two travelers must be returned to the left bank. This move of one hobbit and one orc to the left bank seems counterintuitive and more like backtracking than advancing. There may be additional determinants of state specific difficulty. However, additional research is required to further assess such determinants.

Conclusion

There are many determinants of problem difficulty. The selection and the mere presence of illegal moves is one such determinant. Illegal moves contribute to problem difficulty because they are mistaken for legal alternatives and because they are considered and not selected, which decreases a problem solver's ability to choose the best legal alternative. Increasing the cost of making an illegal move decreases the number of illegal moves, which indicates that there may be an additional explanation to the view

proposed by Jeffries et al. that illegal moves are chosen as a result of limited resources. This additional explanation could possibly be due to a factor within the control of the problem solver because participants were able to improve their performance when no direct manipulations were made to the problem interface or the problem space. I proposed that illegal moves might be selected due to the lack of an intuitive sense to plan when the cost of making an illegal move is low. This work gives light to the possibility that techniques may be adapted to decrease problem difficulty and increase problem efficiency. Discovery of such techniques would be a valuable finding with the potential for real world application.

In this project, participants were asked to give verbal protocols as they worked on the problem. In this process, a new technique was developed to facilitate future problem solving research. Additional instruction and training were initiated to obtain more elaborate and accurate verbal protocols. This process did not prove to affect participants' performance indicating that in the future we may be able to obtain additional information from problem solvers without affecting performance.

I believe that this work makes a significant contribution to furthering the understanding of illegal moves and their role in problem difficulty, but there are many questions that remain unanswered and additional research must be conducted to obtain a better understanding of the determinants of problem difficulty.

APPENDIX A
QUESTIONNAIRES

Need For Cognition

For each of the statements below, please indicate to what extent the statement is characteristic of you. If the statement is extremely uncharacteristic of you (not at all like you) please write a "1" to the left of the question; if the statement is extremely characteristic of you (very much like you) please write a "5" next to the question. Of course, a statement may be neither extremely uncharacteristic nor extremely characteristic of you; if so, please use the number in the middle of the scale that describes the best fit. Please keep the following scale in mind as you rate each of the statements below: 1 = extremely uncharacteristic; 2 = somewhat uncharacteristic; 3 = uncertain; 4 = somewhat characteristic; 5 = extremely characteristic.

1. I would prefer complex to simple problems.
2. I like to have the responsibility of handling a situation that requires a lot of thinking.
3. Thinking is not my idea of fun.
4. I would rather do something that requires little thought than something that is sure to challenge my thinking abilities?
5. I try to anticipate and avoid situations where there is a likely chance I will have to think in depth about something."
6. I find satisfaction in deliberating hard and for long hours.
7. I only think as hard as I have to.
8. I prefer to think about small, daily projects to long-term ones?

9. I like tasks that require little thought once I've learned them?
10. The idea of relying on thought to make my way to the top appeals to me.
11. I really enjoy a task that involves coming up with new solutions to problems.
12. Learning new ways to think doesn't excite me very much?
13. I prefer my life to be filled with puzzles that I must solve.
14. The notion of thinking abstractly is appealing to me.
15. I would prefer a task that is intellectual, difficult, and important to one that is somewhat important but does not require much thought.
16. I feel relief rather than satisfaction after completing a task that required a lot of mental effort?
17. It's enough for me that something gets the job done; I don't care how or why it works?
18. I usually end up deliberating about issues even when they do not affect me personally.

Impulsivity

For each of the statements below, please indicate how often you engage in the action. If you perform this action Rarely or Never please write a "1" to the left of the question; if you perform the action Occasionally please write a "2"; Often "3"; and Almost Always or Always "4". Note that in this set of questions use a 4-point scale.

1. I plan tasks carefully
2. I do things without thinking
3. I make-up my mind quickly
4. I am happy-go-lucky
5. I don't "pay attention"
6. I have "racing" thoughts

7. I plan trips well ahead of time
8. I am self controlled
9. I concentrate easily
10. I save regularly
11. I “squirm” at plays
12. I am a careful thinker
13. I plan for job security
14. I say things without thinking
15. I like to think about complex problems
16. I change jobs
17. I act “on impulse”
18. I get easily bored when solving thought problems
19. I act on the spur of the moment
20. I am a steady thinker
21. I change residences
22. I buy things on impulse
23. I can only think about one problem at a time
24. I change hobbies
25. I spend or charge more than I earn
26. I often have extraneous thoughts when thinking
27. I am more interested in the present than the future
28. I am restless at the theater or lectures
29. I like puzzles

30. I am future oriented

APPENDIX B
WORKING MEMORY

OSPAN

IS $(10 \div 2) - 3 = 2$? SEA

IS $(10 \div 10) - 1 = 2$? CLASS

IS $(7 \div 1) + 2 = 7$? PAINT

???

IS $(3 \div 1) - 2 = 3$? CLOUD

IS $(2 \times 1) - 1 = 1$? PIPE

IS $(10 \div 1) + 3 = 13$? EAR

IS $(9 \times 2) + 1 = 18$? FLAME

IS $(9 \div 1) - 7 = 4$? BIKE

???

IS $(8 \times 4) - 2 = 32$? BEAN

IS $(9 \times 3) - 3 = 24$? ARM

IS $(4 \div 1) + 1 = 4$? GROUND

???

IS $(10 \div 1) - 1 = 9$? HOLE

IS $(8 \times 4) + 2 = 34$? DAD

???

IS $(6 \times 3) + 2 = 17$? KID

IS $(6 \div 3) + 2 = 5$? FORK

IS $(6 \times 2) - 3 = 10$? JAIL

IS $(8 \div 2) + 4 = 2$? HAT

IS $(8 \div 2) - 1 = 3$? LAMP

???

IS $(9 \div 1) - 5 = 4$? CAVE

IS $(6 \div 2) - 2 = 2$? BACK

IS $(7 \times 2) - 1 = 14$? HALL

IS $(6 \times 2) - 2 = 10$? FERN

???

IS $(2 \times 2) + 1 = 4$? MAN

IS $(7 \times 1) + 6 = 13$? WORLD

???

IS $(3 \div 1) + 3 = 6$? DRILL

IS $(10 \div 1) + 1 = 10$? CALF

IS $(4 \times 4) + 1 = 17$? FISH

IS $(3 \times 3) - 1 = 8$? CHEEK

???

IS $(3 \times 1) + 2 = 2$? BREAD

IS $(4 \div 2) + 1 = 6$? GERM

IS $(5 \div 5) + 1 = 2$? DOCK

???

IS $(2 \times 3) + 1 = 4$? GAME

IS $(9 \div 3) - 2 = 1$? NERVE

IS $(10 \div 2) - 4 = 3$? WAX

IS $(5 \div 1) + 4 = 9$? TIN

IS $(10 \times 2) + 3 = 23$? CHURCH

???

IS $(7 \div 1) + 6 = 12$? BEACH

IS $(3 \times 2) + 1 = 6$? CARD

???

IS $(6 \times 4) + 1 = 25$? JOB

IS $(9 \div 3) - 1 = 2$? CONE

IS $(8 \div 1) - 6 = 4$? BRASS

IS $(9 \times 1) + 9 = 1$? STREET

???

LIST OF REFERENCES

- Anderson, J. R. (1990). *The Adaptive Character of Thought*. Hillsdale, NJ: Lawrence Erlbaum Associates.
- Atwood, M. E., & Polson, P. G. (1976). A process model for water jug problems. *Cognitive Psychology*, 8, 191-216.
- Cacioppo, J. T., Petty, R. E., Feinstein, J. A., Jarvis, W. B. G. (1996). Dispositional differences in cognitive motivation: The life and times of individuals varying in need for cognition. *Psychological Bulletin*, 119, 197-253.
- Delaney, P. F., Ericsson, K. A., & Knowles, M. E. (in press). Immediate and sustained effects of planning in a problem-solving task. *Journal of Experimental Psychology: Learning, Memory, and Cognition*.
- Ericsson, K. A., & Simon, H. A. (1993). *Protocol Analysis: Verbal Reports as Data*. Cambridge, MA: MIT Press.
- Griggs, R., & Cox, (1982). The elusive thematic-materials effect in Wason's selection task. *British Journal of Psychology*, 73, 407-420.
- Hayes, J. R., & Simon, H. A. (1977). Psychological differences among problem isomorphs. In N. J. Castellan, D. B. Pisoni, and G. R. Potts (Eds.), *Cognitive Theory* (pp. 21-41). Hillsdale, NJ: Erlbaum.
- Jeffries, R., Polson, P. G., Razran, L., & Atwood, M. E. (1977). A process model for missionaries-cannibals and other river-crossing problems. *Cognitive Psychology*, 9, 412-440.
- Kane, M. J., Bleckley, M. K., Conway, A. R. A., & Engle, R. W. (2001). A controlled-attention view of working memory capacity. *Journal of Experimental Psychology: General*, 130, 169-183.
- Kotovsky, K., Hayes, J. R., & Simon, H. A. (1985). Why are some problems hard?: Evidence from tower of Hanoi. *Cognitive Psychology*, 17, 248-294.
- Kotovsky, K., & Simon, H. A. (1990). What makes some problems really hard: Explorations in the problems space of difficulty. *Cognitive Psychology*, 22, 143-183.

- Lovett, M. C., & Anderson, J. R. (1996). History of success and current context in problem solving: Combined influences on operator selection. *Cognitive Psychology, 31*, 168-217.
- Nair, K. U., & Ramnarayan, S. (2000). Individual differences in need for cognition and complex problem solving. *Journal of Research in Personality, 34*, 305-328.
- Newell, A., & Simon, H. A. (1972). *Human Problem Solving*. Englewood Cliffs, NJ: Prentice-Hall.
- Newell, A., & Simon, H. A. (1988). GPS, a program that simulates human thought. In A. M. Collins, E. E. Smith, (Eds.), *Readings in Cognitive Science: A Perspective from Psychology and Artificial Intelligence* (pp. 453- 460). San Mateo, CA: Morgan Kaufmann, Inc.
- 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.
- Patton, J. H., Stanford, M. S., & Barratt, E. S. (1995). Factor structure of the Barratt Impulsiveness Scale. *Journal of Clinical Psychology, 51*, 768-774.
- Simon, H. A., & Hayes, J. R. (1976). The understanding process: Problem isomorphs. *Cognitive Psychology, 8*, 165-190.
- Thomas, J. C. (1974). An analysis of behavior in the Hobbits-Orcs problem. *Cognitive Psychology, 6*, 257-269.
- Turner, M. L., & Engle, R. W. (1989). Is working memory capacity task dependent? *Journal of Memory and Language, 28*, 127-154.
- Zhang, J., & Norman, D. A. (1994). Representations in distributed cognitive tasks. *Cognitive Science, 18*, 87-122.

BIOGRAPHICAL SKETCH

I was born in Huntington, New York, in 1978 and moved to Tampa, Florida, when I was young. I graduated from Tampa Catholic High School in 1996 and I began attending Florida State University that summer. I majored in psychology and received my Bachelor of Science degree from Florida State University in 2000. I started attending the cognitive psychology program at the University of Florida in 2001.