Title
What am I supposed to do? Problem Finding and its impact on Problem Solving.

Author
Holman, Daniel Matthew

Publication Date
2018

Peer reviewed|Thesis/dissertation

UNIVERSITY OF CALIFORNIA, MERCED

What am I supposed to do?
Problem Finding and its impact on Problem Solving.

A dissertation submitted in partial satisfaction of the requirements
for the degree Doctor of Philosophy

in

Cognitive and Information Sciences

by

Daniel Matthew Holman

Committee in charge:
Professor Jeffrey Yoshimi, Chair
Professor David C. Noelle
Professor Michael Spivey

2018

The Dissertation of Daniel Matthew Holman is approved, and it is acceptable in quality and form for publication on microfilm and electronically:

Co-Chair (*if applicable*)

Chair

University of California, Merced

2018

Dedicated to my mom, dad and sister,

without whom I would not have been able to reach this point,

and whose love, support and belief in me have been invaluable.

I also dedicate this work to the many friends, family members,

colleagues and communities that have supported and encouraged me to succeed.

I am immensely grateful to you all, and fully aware of how I fortunate I am to

have more names to list than I possibly could.

Chapter 1 : Problem Finding and Problem Solving 1
1.1 Definition of Problems 3
1.2 The Problem Cycle 6
1.3 Prior Experimental Work 13

Chapter 2 : Literature Review 17
2.1 Problem Finding 19
2.1.1 Problem Discovery 20
2.1.2 Defining Problems 22
2.2 Problem Solving 24
2.2.1 Insight Problems 24
2.2.2 Insight vs. Noninsight 32
2.2.3 Analogous Problem Solving 37
2.2.4 Experts 40
2.2.5 Procedural Problems 42
2.2.6 Understanding vs. Repeating 44

Chapter 3 : A Taxonomy of Problem Types 46
3.1 Elaboration of Problem Features 48
3.2 The Problem Types 54

Chapter 4 : Experimentation 62
4.1 Two Models of Problem Solving in a Simple Game 64
4.2 Experimental Setup 68
4.3 Experimental Conditions 72
4.3.1 Conditions 73
4.3.2 Experimental Implementation 77
4.3.3 Logging Player Data 78
4.4 Dependent Variables 81
4.5 Results 83

Chapter 5 : Conclusion 100

TABLE 1: ALL PROBLEMS BUT PROBLEM 8 CAN BE SOLVED WITH B-2C-A, BUT 7 AND 9 CAN BE SOLVED WITH A+C; 6, 8 AND 10 WITH A-C (ADAPTED FROM LUCHINS, 1942). 30
TABLE 2: CLASSIFICATION OF PROBLEM TYPES 53
TABLE 3: THE FOUR PROBLEM VARIATIONS SELECTED FOR USE IN OUR EXPERIMENT. 73
TABLE 4: EXPERIMENT CONDITIONS 73
TABLE 5: DEPENDENT VARIABLE DESCRIPTIONS 81
TABLE 6: SUMMARY OF THE SIX MAIN LINEAR REGRESSION MODELS REPORTED IN THIS SECTION. NOTE THAT IN EVERY CASE BUT TIME SPENT THAT PROBLEM TYPE WAS A SIGNIFICANT PREDICTOR OF BEHAVIOR. 84

FIGURE 1: THE PROBLEM CYCLE 6
FIGURE 2: THE EXPERIMENTAL ENVIRONMENT. 15
FIGURE 3: THE TWO-STRING PROBLEM. 26
FIGURE 4: THE BOX AND CANDLE PROBLEM.. 31
FIGURE 5: GRAPH OF INSIGHT VS. NON-INSIGHT. 34
FIGURE 6: THE TUMOR AND LASER PROBLEM 38
FIGURE 7: PARALLELOGRAMS. 45
FIGURE 8: GEARS PROBLEM 55
FIGURE 9: WATER-JUG PROBLEM. 56
FIGURE 10: THE TOWER OF HANOI. 57
FIGURE 11: THE TRAVELING SALESMAN PROBLEM. 58
FIGURE 12: AN EXAMPLE OF THE ROGUE-LIKE GAME 69
FIGURE 13: THE ENTIRE GAME WORLD. 71
FIGURE 14: OPENING A WALL BY PRESSING A BUTTON.. 72
FIGURE 15: *CLEAR PROGRESS*, AND *UNCLEAR PROGRESS* 75
FIGURE 16: A DETAILED PLOT OF ONE'S ACTIVITY IN THE GAME.. 79
FIGURE 17: PLOTS OF THE GAME GRID. 80
FIGURE 18: THE LENGTH OF TIME PLAYERS SPENT IN THE GAME. 85
FIGURE 19: TIME SPENT FOR ALL SUBJECTS. 86
FIGURE 20: TIME SPENT FOR ALL FOUR CONDITIONS USING ONLY THE SUBJECTS WHO FOUND ALL EIGHT RED BUTTONS. 87
FIGURE 21: A HISTOGRAM OF THE TOTAL PERCENT OF THE GAME BOARD VISITED BY PLAYERS. 88
FIGURE 22: PERCENT OF THE GAME BOARD VISITED BY PARTICIPANTS 90
FIGURE 23: PERCENT OF THE GAME BOARD VISITED BY SUBJECTS WHO FOUND ALL RED BUTTONS. 91
FIGURE 24: HISTOGRAM OF THE NUMBER OF PATHS TAKEN BY PLAYERS. 91
FIGURE 25: PATHS 92
FIGURE 26: THE NUMBER OF PATHS TAKEN BY PLAYERS IN THE COURSE OF PLAYING THE GAME. 94
FIGURE 27: HISTOGRAM OF PLAYERS PRESSING BUTTONS MULTIPLE TIMES.. 94
FIGURE 28: AVERAGE BUTTON REVISITS 96
FIGURE 29: HISTOGRAM OF HOW MANY BUTTONS PLAYERS WERE ABLE TO FIND AND PRESS. 96
FIGURE 30: THE TOTAL NUMBER OF RED BUTTONS FOUND BY PARTICIPANTS. 98
FIGURE 31: HISTOGRAM OF THE NUMBER OF TIMES PLAYERS TRIED TO END THE GAME.. 98
FIGURE 32: THE NUMBER OF TIMES PLAYERS TRIED TO END THE GAME BY PRESSING THE BLUE BUTTON 100

Acknowledgements

First and foremost, I need to thank my family for getting me to where I am today. Without my family's love, support, and encouragement, I have no doubt that I would not have made it this far.

I'd also like to thank some wonderful teachers and educators I've had over the years that bolstered my love of learning and of finding new ways to think about things. My preschool principal, Margaret Riley, with whom I continued to visit well into college, was the founder of the Montessori school I attended, instilling in me early on to learn by doing and discovering. Mr. Mason, my middle school principal, spent time after school with me to learn math, and after discussing my interests with my dad, I started to learn to program in the fifth grade, despite not knowing what to call the thing I wanted to do (a recurring theme for me). Mr. Lyons, my high school history teacher was an avid supporter of everyone having different learning styles, and encouraged students to do what they needed to do and to find their own ways to understand the material. Clark Elliott, who taught the first cognitive science class I took, continues to be a source of fascination long after introducing me to emotional reasoning and sharing his struggles with decision making. I'd also like to thank the faculty and staff of Shimer College, for being exactly what I needed in a school, and giving me opportunities to learn and explore in many different ways and over many different areas. There are many more names that I could list that deserve to be mentioned, including friends, relatives, and friends that are closer than family, and I hope when they read this they know that I'm thinking of them.

I, of course, owe a debt of gratitude to those I've worked with at UC Merced, beginning with Michael Spivey, who met with me and my dad for an interview after I initially applied, and took me into his lab, offered me a chance to co-author my first published work, and helped me to conceive of and run my first research experiment. A further debt is owed to both David C. Noelle and Jeff Yoshimi, who not only comprise the rest of my committee, but were there to help me when I felt I needed to redirect my efforts but was unsure how.

I want to extend an extra thank you to Jeff for taking me on as a mentee. We had talked about problem solving and wanting to work on something together, but I know he hadn't expected to become my primary advisor, and be so deeply invested in my ideas when we first started. I am very grateful that he took a chance on me, and I have benefited greatly from his patience, availability, thoughtfulness, encouragement, and prodding.

Curriculum Vitae

Daniel Holman
1205 Devonwood Drive, Apt. 103
Merced, CA 95348
(847) 769-8802
dholman@ucmerced.edu

Education

University of California, Merced, Cognitive & Information Sciences 2012-2018
Ph.D.

DePaul University, Chicago, Interdisciplinary Studies 2007-2010
M.A.

Shimer College, Chicago, Liberal Arts 2005-2007
B.A., with Distinction

Teaching

University of California, Merced, Cognitive & Information Sciences

- Introduction to Cognitive Science, Teaching Assistant 2014
- Introduction to Cognitive Science, Teaching Assistant 2015
- Philosophy of Cognitive Science, Teaching Assistant 2015
- Introduction to Cognitive Science, Teaching Assistant 2016
- Introduction to Ethics, Teaching Assistant 2016
- Contemporary Moral Problems, Teaching Assistant 2017
- Introduction to Philosophy, Teaching Assistant 2017
- Introduction to Ethics, Teaching Assistant 2018
- Philosophy of Cognitive Science, Teaching Assistant 2018

Awards

Eugene Cota-Robles Fellowship 2012-2013

Publications

Holman, D. M. & Spivey, M. J. (2016). Connectionist models of bilingual word reading. In R. Heredia, J. Altarriba, A. Ciesliska (Eds.), *Methods in Bilingual Reading Comprehension Research.* Springer.

Rigoli LM, Holman D, Spivey MJ and Kello CT (2014) Spectral convergence in tapping and physiological fluctuations: coupling and independence of 1/f noise in the central and autonomic nervous systems. *Front. Hum. Neurosci.* 8:713. doi: 10.3389/fnhum.2014.00713

Abstract

In Chapter 1, a definition for *problem* is introduced, along with a model of the *problem cycle*, which is used to conceptualize the relationship between normal cognition, problem finding and problem solving. Initial experimental efforts to explore problem finding are discussed. In Chapter 2, previous work on problem finding and problem solving is reviewed. The few existing accounts of the stages of problem solving are surveyed, and the older literature on different kinds of problems and how they are solved is reviewed in detail. Chapter 3 builds upon the first two chapters by introducing a new taxonomy of problems, highlighting four dimensions that differentiate problem types and explaining each type at length. Chapter 4 describes an experiment developed to test the idea that problems can be organized along the lines of the taxonomy introduced in Chapter 3, and that these differences determine how they try to solve a specific instance of a problem. A simple game is developed and behaviors in the game are tracked, confirming the hypothesis that different problem types are solved in different ways, that are specific to the type of information available to the problem solver. The last chapter summarizes the dissertation and describes avenues for future research.

Chapter 1: Problem Finding and Problem Solving

Problem solving is a component of a larger process governing general cognition. We routinely encounter information and decision tasks in daily life, many of which are dispatched unconsciously and without causing disruption to the primary focus of attention. Sidestepping debris on the ground while walking, hitting the button to unlock your car a second time because it didn't work the first time, etc., might be viewed as problems, but are not the sort that research into "problem solving" generally investigates. The majority of problem solving research is concerned with the ways to go about addressing specific, defined problems. While there is some variety in the particular problems explored, including visual search, value judgments and predictions or trial and error, the research begins with a problem that has already been established, and demands conscious attention, and tries to explain how such a problem may be solved.

It will be argued that an aspect of problem solving in this classical sense is typically overlooked, namely the cognition and discovery of problems themselves. Before a problem can be engaged cognitively, it has to be recognized and formulated. The period leading up to the stage where one has a clear problem to deal with is important, and has consequences that extend into the problem solving activity itself. The problem determination phase will determine how you approach the problem, how well you are able to solve it, and so on. In addition, during the process of attempting to address a problem there is an element of re-evaluation that takes place, both in terms of clarifying or altering the problem with new information, and in considering the context and implications that may lead to a determination that a different problem altogether should be attended to that is more important, and may even nullify the presently perceived problem.[1] Attempts to solve problems can, themselves, involve the recognition and resolution of a number of component, or sub-problems to enable the solving of a larger one.[2] This is to say that even when a problem is being interpreted in a particular way, there is a degree of exploration of alternative possibilities that occurs. This is evident from studies of insight showing that how problems are solved relies on how they are perceived, as well as in neuroscientific studies that indicate tendencies to maintain some openness to different possibilities even when one choice has been selected (Dieciuc, Roque, & Folstein, 2016).

---

[1] For example, when trying to get a computer program to run, a problem might arise in the form of a command that fails to work. While trying for some time to address the error, re-structure the command, or alter code around it to make it work properly, it may be discovered that the program can be run without that component, or that the task could be achieved in an entirely different way, and so even though the immediate problem doesn't get resolved, switching to a different method (changing to a different problem) can essentially solve the larger issue. Suppose a program needs to loop many times over data, and the programmer writes a recursive algorithm that keeps stumbling. At some point, it might be decided that it's not worth trying to be recursive, and that a series of loops is good enough.

[2] Newell and Simon (Newell & Simon, 1972) famously explain how a problem may be solved by iteratively breaking it down into smaller and smaller problems until a problem with a clear solution is found, solving that one, and working backwards up the stack until the entire problem gets resolved. Their resulting General Problem Solver (GPS) model has been demonstrated in computer science using a number of problems of the same general kind—things that require a number of steps to complete—but it lacks a specific implementation, isn't necessarily the most efficient method and is somewhat limited in the types of problems is can be applied to.

1.1 Definition of Problems

Roughly speaking, a problem is something that prevents or impedes the attainment of a desired condition in such a way as to require conscious intervention in order to overcome it and proceed to other concerns. This definition is intended to allow for a degree of flexibility regarding the determination of whether a particular circumstance qualifies as a problem or not, while maintaining the characteristics of a problem that are commonly found in the problem solving literature. This definition also encompasses cases that might not be considered to be problems in the classic sense, but make sense to include based on the specification of problem types that are described in Chapter 3.

One virtue of this definition is that it allows us to say things like “When I got up to leave, I knocked over a glass of water, but nobody cares, so that’s not a problem,” while at the same time including things like “When I got up to leave, I knocked over a glass of water, my stuff isn’t waterproof, so I had to deal with the problem of wet belongings before I left,” as well as “When I got up to leave, I knocked over a glass of water without even realizing it—I didn’t know until much later when I was trying to figure out why my TV wouldn’t turn on, and traced it back to the water having been spilled on the power cord!” In other words, a circumstance or event doesn’t necessarily constitute a problem itself, but it is a problem if it immediately impacts something you want (or don’t want), or is later discovered to be responsible for an undesirable circumstance.

Similar definitions have been proposed in the literature. For example, Duncker describes a problem as follows: "A problem arises when a living creature has a goal but does not know how this goal is to be reached. Whenever one cannot go from the given situation to the desired situation simply by action [i.e., by the performance of obvious operations], then there has to be recourse to thinking" (Duncker, 1945). Duncker's definition essentially characterizes problems as any case in which the means of reaching a goal are unclear, although it doesn't provide a reason for the lack of clarity. This is an important aspect to point out. The present definition, and the subsequent detailed analyses, emphasize that a problem can involve multiple dimensions of unclarity or uncertainty. One of these is uncertainty about whether a goal has been reached, as opposed to how it may be reached. For example, if your goal is to be healthy, one might still have the problem (in the present sense) of not knowing whether at the end of a day or month or time period whether one has achieved that goal. Duncker's definition does not extend to cases of this sort.

The current conception of problems involves at least the following four dimensions of clarity or uncertainty. They will be discussed in greater detail in Chapter 3, but since they will be referenced several times before then here is a brief overview. These four dimensions are *Possibilities*, *Course, Progress* and *Goal-Reaching*, each of which can be either clear or unclear. These dimensions specify what information about a problem is available to the solver. They can be summarized in the form of four questions:

Clear / unclear possibilities: Are the rules of the system, or the options available, known?
Clear / unclear course: Is it obvious what should be done next, or which choice to make?

Clear / unclear progress: Can the solver tell if they are closer or farther from completing the problem, and whether their choices are making a difference?
Clear / unclear goal-reaching: Does the solver know whether or not they've succeeded in solving the problem?

Depending on what aspects of a problem are known, it is considered to be a distinct type of problem. The combination of clear or unclear in each dimension results in a list of sixteen problem types.[3]

It should be noted that while these categories of problem type are new, this is not the first effort to observe variations in problems or solution methods. De Groot, for example, in describing Selz's characterization of problems, gives a description that applies to at least two of these problem categories. "In the simplest case, one is able to solve a problem since one 'possesses' the necessary know-how. That is, the solution itself is not at one's disposal but rather the means to reach the goal is available in some form. In principle, two cases can be distinguished: (1) the subject may consciously know how to proceed, or (2) he may have an automatic solution complex available" (De Groot, 1946). The distinction here is between simply being able to carry out the actions that lead to a solution, without necessarily knowing how it works, and having a specific, conscious, plan. In either case, this is a type of problem that one needs to solve in the sense of carrying out, rather than finding a means of carrying out, the solution for. This will be discussed further in Chapter 2.

---

[3] An attempt was made to find a structure or grouping within the sixteen types that might generalize to broader categories. While there are a few possible arrangements, nothing has so far seemed to strongly suggest that they should be divided in this way. However, as will be shown in 4.3, Table 3, it can be useful to combine dimensions at times.

1.2 The Problem Cycle

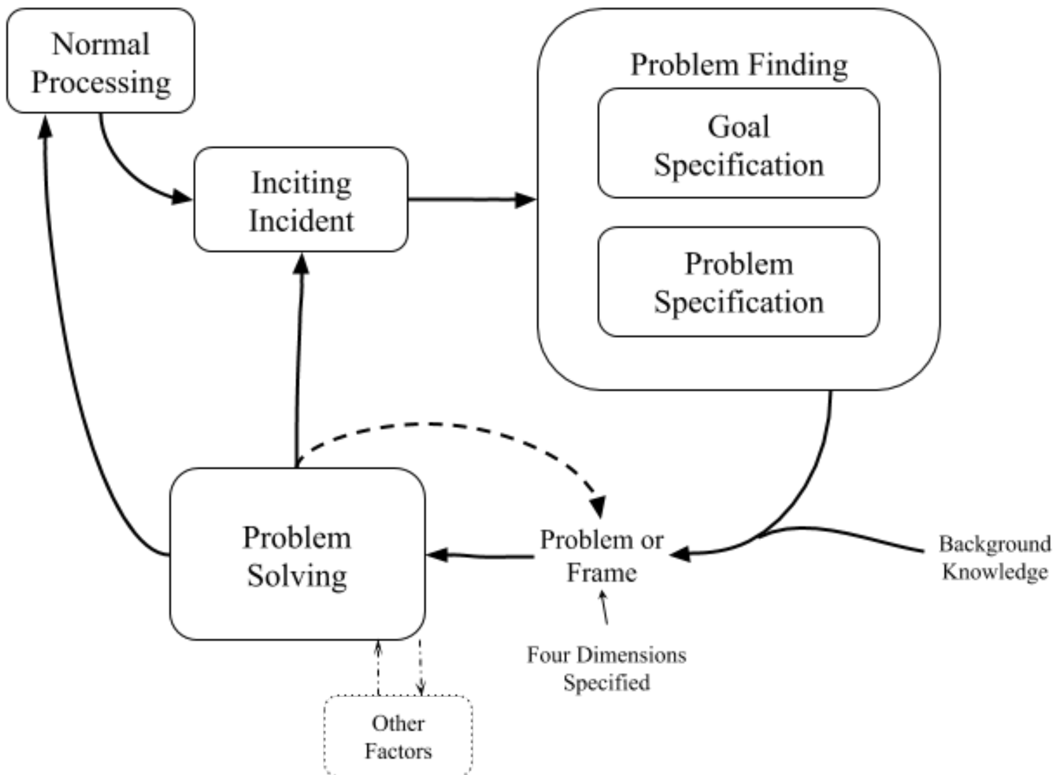

Figure 1: The Problem Cycle represents a problem as a departure from one's normal routine, sparked by some interruption and followed by a process in which the problem is identified and formulated. Once found, the problem solving stage works to resolve the issue and return to normal. The problem or frame may be changed throughout the course of solving, however, and other influences including time pressure, mood and fatigue play a role in managing the problem. In the course of finding a solution, other incidents may occur, leading to another iteration through the cycle as well.

Figure 1 depicts the general sense of how the problem cycle works. This could be drawn in different ways, and is not meant to indicate definitive boundaries between separate processes. Rather, it shows a cyclical, iterative system responsible for dealing with problems, and specifically seeks to emphasize and describe the precursor to problem solving, namely problem finding (see section 2.1).

Normal Processing is used as an umbrella label for the myriad processes typically being carried out as a part of everyday cognition, considered for present purposes as everything that is *not* problem-oriented. Much of this can be characterized as task execution, or performance of routines and functions that occur without issue.

The initial entry point to the problem cycle is the *inciting incident*, or source of trouble that requires conscious attention to be resolved. It is important to stress that this incident is not synonymous with the problem, as will be explained shortly, but is ostensibly the impetus for finding a problem. An inciting incident can be immediate, as in a loss of power while at work, or delayed, such as noticing an error in a program that may have changed what it produced from what it was intended to produce. Often, the incident is perceived as something having gone wrong,[4] though this is certainly not always the case.[5] Fundamentally, a problem exists as something preventing the direct attainment of a goal.[6] Impasses, as described in problem solving literature, could be considered to be inciting incidents as well, however, in the present account it is not necessary to know

---

[4] Or going wrong, or going to go wrong.

[5] Getting interrupted by a surprise party, or winning the lottery aren't typically seen as something going wrong, but they still constitute a change of circumstances requiring attention and qualify as problems insofar as one needs to figure out what to do in response.

[6] A thing that can "go wrong" and that can be recognized as having gone wrong, even if it wasn't ever consciously identified in advance as a project or goal one had. While pursuing the goal the agent might have no conscious awareness of it as a goal. The requirement is only that the agent can later come to realize that the goal was not being reached. Purposeful behavior can initially be largely unconscious, though it is a requirement in the current definition that we can later become conscious of failing to achieve your goal or purpose. You might not even realize that you had a goal at first until you later realized you failed to succeed at it. This definition allows that animals can engage in goal-directed behavior, since they can notice something has gone "wrong" in an action sequence. A dog that can't get through a fence tries to find a way around it (Köhler). We can't say whether they are conscious or not, but they are doing the same kind of process of recognizing a problem and trying to solve it.

one's goal in order to experience an inciting incident, and so does not fit the classic examples of an impasse.[7]

Once attention has been drawn by the inciting incident to the existence of a problem, the next step before it can be solved is determining what the problem actually is, and how to consider it. This is the stage most lacking in prior work on problem solving,[8] and is described here as *problem finding*. Briefly, problem finding begins when an inciting incident occurs, and works to figure out what to do as a result.

As (Gershman, Horvitz, & Tenenbaum, 2015) rightly observe, the incident that launches problem finding is not isolated from other concerns. Part of problem finding, as described next, is the need to decide what to attend to in terms of importance. Suppose you were just finishing breakfast, and getting up from the table to go to work when you accidentally knock over a bowl, leaving shards of glass and milk all over the table and floor. This certainly constitutes an inciting incident, as it interrupts and prevents what the next actions would have been (for instance, putting the bowl in the sink), but before necessarily figuring out what to do about the spill, it may be decided that it isn't as important to deal with at all as it is to get to work on time. In a way, this could be seen as dismissal altogether, indicating that it isn't actually a problem (in which case, not all inciting incidents or interruptions of normal processing necessarily lead to problems that have to be solved, which doesn't negate the fact that problems rely on inciting incidents as starting points), but it can just as easily be said to be deferred to a later time in favor of

[7] An impasse is described as the point during problem solving that the agent seems to be stuck, not knowing what to do next. This assumes that the goal is already known, and the problem has been set, but because it can trigger another round of problem finding, it qualifies as an inciting incident.

[8] See Chapter 2 for other researchers with similar claims.

more pressing interests. Presumably, not dealing with the broken glass before leaving for work doesn't inherently mean it isn't a problem at all, since it will still need to be dealt with later. Relatedly, it's perfectly valid to include the actions of other people when considering problems. Choosing not to take direct or immediate action, and then having someone else resolve the issue for you does not oppose this model any more than someone giving away the answer to a riddle would mean it wasn't a tricky question.

Problem Finding consists of two subcomponents, *goal specification* and *problem specification*. The first of these is used to determine what to attempt to achieve as a consequence of the inciting incident,[9] for example if someone was working on a portable computer and received an alert that the power was very low, they might specify a goal of saving all of their work before the battery is drained, or they may decide that their goal is to find an available power outlet before that point. The second step, problem specification, selects from a variety of interpretations and ways in which the goal may be satisfied, producing a kind of template corresponding to the form of the problem to be solved. For example, if the inciting incident was a sudden downpour, the goal might be specified as "keep from getting wet between here and the car," and the problem may be specified as "what can I use to keep me dry while I'm outside?"[10]

---

[9] This is a critical part of problem discovery, and should be further explored in detail. For the present, it must suffice to acknowledge that the job does occur, without a discussion of how.

[10] Alternate problem specifications include "how can I avoid going outside," and "how fast should I walk to stay as dry as possible?" Both of these share the goal of minimizing contact with the rain, but are different conceptions of what to pursue to achieve that goal.

The *model* or *frame* of the problem represents the problem to be solved as it is considered by the problem solver, and includes any related background information that may be useful in working out a solution. This is also where the four dimensions of problem type may be applied, as one's consideration of the problem will impact the ease or efficiency of solving it. Using the rain example, the frame of the problem might include information like "common ways to keep out the rain are umbrellas, ponchos, trash bags or newspapers," as well as whether there are any readily available. Thus, from a general sense of the problem as "what can I use to keep me dry while I'm outside," the more specific model of the problem may become "I need to get to the car without getting wet and I don't have an umbrella with me. There are probably trash bags around that would work. Where can I find one?" At this stage the problem is recognized as involving clear or unclear possibilities, course, progress and goal-reaching, which, in turn, determines the type of problem to be dealt with.

This representation of the problem is what *problem solving* begins with, and operates on. This is generally recognized, even if implicitly, in the problem solving literature. When problem solving is referenced in the academic literature, it begins by presenting problems that have already been formulated. A matchstick problem, for instance, consists of some arrangement of matches that form a shape, and the solver is asked to modify the shape to produce a specific result by making a limited number of allowable changes to the layout. The problem solver begins with a defined goal and an understanding of the task before them.

The dashed line leading from problem solving back to the model or frame indicates that in the midst of problem solving, what is known about the problem, including how it is considered, may change over time, resulting in a change in the way that problem solving continues. In the midst of the matchstick problem, for example, the solver might realize that a different interpretation of what is before them can facilitate a solution they otherwise wouldn't have considered. Another example is realizing in the 9-dot problem that lines can extend beyond the boundary of the dot diagram: at this point the problem specification changes and the solver can approach the problem in a new way (the possibilities have changed).

Attached to problem solving is an acknowledgement that a wide range of *other factors* may also affect the process. One's ability to effectively solve a problem will differ based on their mood, focus, interest, level of energy, and so on.[11] In fact in the experiment conducted as a part of this work (Chapter 4), one's level of engagement with an activity appears to play a central role in participants' behaviors.

Ultimately, finding the solution to a problem (and/or implementing it) can lead back to normal processing. However, either by creating a new issue as a result of the solution, or by encountering troubles while trying to solve a problem, the cycle can loop right back to the inciting incident phase, leading to potentially nested or chained problem cycles.

[11] Also, given that numerous processes in cognition occur simultaneously, it is certainly possible that a single inciting incident gives rise to multiple problems being found and worked out in parallel, progress towards one can impact another. The inclusion of this feature in the diagram allows and accounts for this as well.

The following example illustrates these stages and a possible path through the cycle. Suppose that you are in your office, working on filling out a long form on your computer (*normal processing*), when the power suddenly fails, and you've lost your connection to the network (*inciting incident*). As a result, your work/routine is interrupted, and you begin *problem finding*. Considering the variety of concerns that are raised by this interruption (loss of work, etc.), you decide that you need to check in with your family to make sure they're ok (*goal specification*). Consequently, your *problem-specification* takes the form of the question "How can I reach my family to check on them without power or a connection to the internet?"

As you proceed toward solving the problem, your *model* is built, you *frame* the problem in a specific way that contains a number of points, including "I want to contact my family," "I would normally use the internet to contact my family," and miscellaneous background knowledge, such as knowing where they are, what they're likely to do, other ways of communicating, etc. This leads you into *problem solving*. The specific problem you go to work on is a *creative problem*. This is based on *Clear goal-reaching*; you'll know if you've made contact or not. *Clear progress*: if you have multiple people to talk to, you could consider this clear if you count it as how many you've reached versus how many you need to. *Unclear possibilities*: you know a number of things you could do, including waiting for the power to come back, trying a landline, walking over to where they were, etc., but probably don't have a strong feeling of knowing *everything* you could try in the moment. *Unclear course*: you don't have a procedure to follow in cases like this, and are not sure which of the possibilities available to you would be best to try first.

This is only one of several paths that could have been taken. If, for example, you had a plan in place for what to do when the power goes out, you would have a clear course, producing what will be referred to as an *Insight 1* type of problem instead. Similarly, if you wanted to know that your family as alright, but didn't need to talk to them, you might go about asking anyone around you for more information. Is the power out in just your office? The building? The state? Is the cause known? These could be different tracks towards the same goal of determining the safety of your family that don't involve directly talking to them.

At this point you can go through the process of solving the problem in order to (hopefully) return to normal processing.

1.3 Prior Experimental Work

As part of this dissertation, two experiments were conducted. The first attempted to elicit and observe the discovery of a problem. The second experiment, described in Chapter 4, directly manipulated the problem specification component in order to study how different types of problem solving behavior are related to how one specifies a problem. The first experiment is described here, and is important because it shows how difficult it is to directly study the problem finding phase of the problem cycle.

In this first experiment, an environment was created with something unusual enough to be likely to gain attention, but it did not give any instructions whatsoever to participants. They were expected to discover the issue and form the problem on their own. The environment consisted of a small room in which there was a shelf containing several curious items. The purpose of the items was to create a likelihood that subjects would reach out with their hands. A motion sensor was used to track each subject's position, and respond to the motion of their hands, playing musical notes when specific movements occurred.

The hope was that participants would enter the room with no prior information about it, and leave with the understanding that there had been a problem to solve, and that this problem was to figure out the right way to move their hands.[12] At first, notes would rise as a subject's hands got closer together, and a triumphant sound would occur at the closest point, with the supposition that the participant would deduce that they were meant to bring their hands together. Later, notes were associated with the physical space above the shelf, and if a sequence was played, subjects were rewarded with chocolate (this version is shown in Figure 2). The sequence would grow over time so that they would have to figure out that they needed to follow a set pattern rather than be rewarded at random.

[12] Since the intent of the experiment was to observe the activity of discovering a problem, it seemed appropriate to have provided a specific problem to find. However, in this case, there were multiple potential problems, including determining whether the sounds were random or causal, what the cause of the sounds might be, and even trying to discover whether there was something the subject was meant to look for at all. Future versions of this study would likely benefit from greater constraints on the design in order to more clearly discern when a problem has been detected and to remove spurious results.

Figure 2: The experimental environment, showing the shelf of interesting objects, and the chocolate dispenser on the right. Subjects were told not to cross the line on the floor but were otherwise given no instructions or restrictions other than to stay in the room for ten minutes.

While there was a specific problem/task in place, the experimental interest was on the discovery of the problem rather than the solution. At the time, physical recordings of body position over time were intended to be used to reveal moments of insight or understanding, at which point the realization of circumstances would be given away by a change in motion. Unfortunately, the level of data analysis needed to make such discoveries was beyond the experimenter's expertise, and engagement or exploration by participants was significantly lower than anticipated.[13] Some degree of significant

[13] Many subjects stood still, or leaned against the wall and waited out the time rather than touching any of the objects in front of them. When asked why, reasons given ranged from "It's not mine," to "I didn't know if I was supposed to." In one case that is still bewildering, a subject stopped exploring after chocolate was dispensed, and later stated that "when the chocolate came out, I thought I had done something wrong."

difference between people who got closer to the problem than others was shown, but not quite as had been intended, or to the same extent.

The basis for using motion capture to identify flashes of insight or understanding was a 2009 study by Damian Stephen et. al., (Stephen, Dixon, & Isenhower, 2009) in which participants' eyes were tracked while attempting to determine which way a particular gear in a set would rotate. Given repeated exposure to this task, it was found that subjects would first trace the outline of the gears, later changing strategies to alternating left or right for each, and ultimately simply counting the number of gears and deciding which way one would turn depending on whether the number was odd or even. By performing analyses on the recorded logs, Stephen was able to show that subtle shifts in movement could accurately predict a change of strategy several trials in advance of the change, demonstrating the effect of an unconscious process searching for more optimal methods. In the experiment described here, it was hoped that similar alterations in movement could be uncovered that could be used to indicate the points at which a participant had begun to notice the sounds, and start testing their environment.

There were several reasons that this experiment did not turn out as anticipated. One reason was that the motion tracker that was used to provide audible feedback to participants in order for them to identify the problem produced anomalous noises throughout the trial which interfered with their ability to determine the cause of the sounds. Another reason was that the Stephen paper made it clear that the significant factors they had found weren't derived from simple, directly recorded data, but rather from an analysis of cross-recurrence of the entropy of their logs, but attempts at replicating their methods to apply to this study were unsuccessful.

A subsequent experiment, discussed in Chapter 4, was concerned with the second part of problem determination: problem specification. For this study, a scenario was created in which the inciting incident and goal-specification had already taken place, but by providing different details to participants, the problem-specification stage effectively directed subjects into four different models of the problem, and consequently four different types of problem to solve. If problem solving behavior is dependent on how a problem is framed, then there should be measureable differences in the way individuals carry out the task depending on the experimental condition. By altering how well subjects were able to gauge their progress and providing variations in the instructions of a simple computer game (whereby some participants knew more about what they were supposed to do than others), this experiment successfully showed that how much a player knows does produce changes their behavior in predictable ways. For example, players with more information performed better in the game and made fewer errors and made more direct movements.

Chapter 2: Literature Review

As a phenomenon, the practice of solving problems has been studied extensively over the last century. Beginning in earnest in the 1920s, researchers like Köhler, Wallas and Maier began to examine how both humans and other animals would behave when confronted with a puzzle, or some sort of obstacle to overcome in an effort to understand the reasoning processes involved. As the work continued, certain factors were discovered which would need to be accounted for in explanations of problem solving, such as the differences in the amount of time required for subjects to solve completely novel problems compared to those who had performed a number of similar ones already, or the tendency to get “stuck” on some problems before appearing to suddenly find a solution. Maier recognized that a particular problem could be viewed in different ways, and suggested that for some problems it would be necessary to reconsider the problem before it could be solved. In fact, Maier proposed that familiarity with the components of a problem could actually inhibit the discovery of a solution that required using objects in non-standard ways. Luchins went on show that having found a solution to similar problems could make it more difficult to solve problems of the same type if they couldn’t be resolved in exactly the same way. Questions were raised not only about how problem solving works, but whether there were actually different classes of problem altogether, and if so, how to differentiate them from one another and provide a model of problem solving that would accommodate these differences. More recently, some researchers have begun to draw attention to another area related to handling problems; their discovery.

The literature on problem solving historically focused on how problems are directly solved. They would start with a problem and study how participants attempted to solve it. This older literature is reviewed in section 2.2. Over time the emphasis began to shift to the framing of a problem. This led to discussions of different types of problems. This in turn has led to some (largely implicit) discussion of what is referred to in this work as problem finding. This more recent literature is reviewed next in section 2.1.

2.1 Problem Finding

As explained in section 1.2, the present conception of handling problems involves a stage in which a problem is found and formed before any attempt at a resolution can be made. Several accounts of problem finding exist in the literature that are similar to the present view. These accounts generally see problem finding as distinct from problem solving, and necessarily a precursor to it. They also tend to agree that this aspect of problem solving has been insufficiently explored, and does not get the attention it deserves, considering how important it is to the overall process of solving problems. (None of these authors go particularly deeply into the process and so the point about insufficient study still stands). Many of these authors describe a sequential progression through several stages of problem solving but also qualify this by allowing that these processes will unfold in different orders and can be nested in various ways. These are all points that being argued for here as well.

Roughly, this model for problem finding is an account of the parts of the process that begin with something occurring (or having been found to have occurred) that will need a conscious response, and end with the establishment of problem to solve. This problem finding stage is composed of two important components: *goal-specification* and *problem-specification*. Goal-specification refers to the decision of what one's goal is in response to the incident, which can widely vary due to a number of factors, while problem-specification is the general sense of what the problem is considered to be.

These two components have been described in the literature, but not distinguished and related in precisely the way that is done here. In this section, a review of what previous discussions have been found is expressed, organized around two persistent themes: first, problem discovery (in present terms, roughly, the inciting incident plus goal specification), and second, problem defining (in present terms, roughly, problem specification).

2.1.1 Problem Discovery

A number of earlier researchers have taken pains to point out that problems must first be uncovered before any attempt at problem solving can be made. Thomas, for example states directly that “The first stage of problem solving is problem finding,” and “Until a problem is found...it cannot be solved” (Gilhooly, 2012). Davidson and Sternberg concur, asserting that “Problem solving does not usually begin with a clear statement of the problem; rather most problems must be identified in the environment; then they must be defined and represented mentally” (Davidson & Sternberg, 2003).

Others make the distinction between things we know how to do, and things that must be figured out. Mackworth, for example, defined problem solving as “...the selection and use of an existing program from an existing set of programs…” in contrast to problem finding, which is “...the detection of the need for a new program by comparing existing and expected future programs” (Mackworth, 1965). Similarly, Getzels “...[made] the distinction between presented problem situations and discovered problem situations" (Arlin, 1975). Arlin puts this together nicely, saying of Mackworth and Getzels that "The suggestion is that the stimuli, or in Mackworth's terminology, ill-defined problems situations, initiate the problem-finding process" (Arlin, 1975).

All of this relates also to an earlier body of work that divided problems into routine and nonroutine, or "productive" and "reproductive" depending on whether the solution strategy is already known. When first encountering a problem that one does not know how to handle, it becomes necessary to work out how the problem may be solved, producing a method or technique by which the same (or a very similar) problem could be dealt with again. This is what Duncker, Wertheimer, etc. refer to as "productive thinking." Alternately, a problem may arise for which the solver already has a strategy available that they simply must apply, producing the solution again in "reproductive thinking." Wertheimer illustrates several examples of both productive and reproductive thought in his work, including finding the area of a parallelogram (see Figure 7). The first time a student is shown a parallelogram, they struggle with figuring out the area inside it, but once they are taught a method for doing so, they can re-apply it to future shapes given to them.

These descriptions generally support the present model, in which an inciting incident triggers problem finding. While Getzels, Mackworth and others recognize that finding a problem can result in something that is more or less well-defined, the present work improves on the distinction by providing sixteen categories of definition to account for not only how well defined a problem may be, but what parts of the problem are or are not well defined. These are explained at length in section 3.1.

2.1.2 Defining Problems

Once a problem is determined to exist, it must also be formulated in some way. This issue actually presents itself in a few different ways in the literature. While many researchers emphasize the importance of how a problem is framed, there is more than one interpretation of what it means to form a problem. For Maier, it indicates something of a different focus or understanding of the problem. In his two rope experiment (see section 2.2.1), he relates that subjects would try different strategies depending on what they considered the difficulty to be; some subjects would frame the issue as not being able to reach far enough, while others would see it as not being able to keep the ropes from moving.

Another version of framing a problem involves how specific it is in its setup. In some cases, problems are considered to be difficult due to a lack of definition, implying that giving a stronger definition to the problem is not only possible, but would make it easier to solve. For example, an ill-defined problem like *make a good dinner* is difficult to accomplish, in part, because it's unclear what qualifies a dinner as *good*. Increasing the specificity, for instance *make a dinner that uses ingredients already in the house, takes less than an hour to make, and tastes good enough that no one complains*, makes the problem easier to solve if for no other reason than offering clear means of telling whether or not it has been achieved.

A third sense of a problem's form is the literal structure it takes, in terms of wording or depiction. A problem that is presented in words might be much more difficult than the same problem given as an equation, for example, or a picture. As analogical transfer shows (section 2.2.3), recasting a problem in different terms can also lead to solutions that were otherwise difficult to arrive at.

Finally, problem framing can refer to the explicitness and completeness of a problem as it is expressed. Amarel, citing "...the relationship between problem formulation and problem solving efficiency…" (Amarel, 1981) explains that how a problem is presented can dramatically impact the amount of effort required in solving it. Generally, the more information that can be used to constrain the problem at the outset, the simpler it will be to work out the solution.

Amarel uses the missionaries and cannibals problem to support this, showing that paring down the potential search space by eliminating paths that fail or don't advance the solution cuts down the effort involved in finding a successful path.[14] Similarly, Langley points out that even in machines meant to make new discoveries, "...the first step in using computational discovery methods is to formulate the discovery problem in terms that can be solved using existing techniques" (Langley, 2000). This means that work has to be done to express a problem in a specific format before a computer can begin to work on it.[15]

According to the account of problem finding developed in this dissertation, an inciting incident triggers the process that then determines a goal to address, and constructs a problem to be worked on. These stages are similar to earlier work by

---

[14]One wonders how well this would apply to things like tic-tac-toe, in which, theoretically, the possible choices can be cut down as one progresses, but it would be difficult to apply in advance. Amarel isn't quite saying that it all has to be done in advance, though, but rather that the appropriate application of information can be used to continuously restrict the possible choices.

[15] While not quite what Langely was talking about, this brings to mind the different ways of representing math problems in computer science, such as infix notation versus reverse Polish notation.

Mackworth and Getzels who "...imply three elements that are necessary for an operational definition of problem finding. They are: (1) a problematic situation; (2) an opportunity for subjects to raise questions; and, (3) a way of categorizing the questions once raised. The latter is necessary to single out the *general question*" (Arlin, 1975). The general question is approximately what is referred to here as the product of problem-specification.

2.2 Problem Solving

The literature on problem finding is in some ways a response to the much larger body of research on problem solving. One of the assertions about the importance of the problem finding stage is the impact that occurs as a result of this process, particularly regarding how easily or even how likely it will be that the selected problem will be solved. In the problem solving literature, efforts have been made to determine what makes some problems harder than others, and whether or not there are fundamentally different kinds of problems.

2.2.1 Insight Problems

Much of the early work on problem solving (beginning around the 1920s) was done by Gestalt psychologists such as Max Wertheimer and Wolfgang Köhler. In particular, they chose to focus their attention on insight problems. Novick and Bassock define an insight problem as follows: "the solution to an insight problem appears suddenly, accompanied by an "aha!" sensation, immediately following the sudden restructuring of one's understanding of the problem" (Novick & Bassok, 2005). They go on refer to Duncker's classic definition: "The decisive points in thought-processes, the moments of sudden comprehension, of the Aha!,' of the new, are always at the same time moments in which such a sudden restructuring of the thought-material takes place" (Duncker, 1945; Novick & Bassok, 2005).

The idea was to focus on tasks whose solution required not only multiple steps, but also multiple stages of problem solving. Different strategies would have to be employed at different stages in the process. Typically in an insight problem the obvious strategy only works initially, and at some point an impasse is reached. Van Lehn (1989) summarizes.

> The earliest experimental work on human problem solving was done by Gestalt psychologists, notably Köhler, Selz, Duncker, Luchins, Maier, and Katona. They concentrated on multi-step tasks where only a few of the steps to be taken were crucial and difficult. Such problems are called insight problems because the solution follows rapidly once the crucial steps have been made.

An excellent example of an insight problem is Maier's 1931 two-string problem: two lengths of rope suspended from the ceiling of a room, along with a chair, a pole, pliers and some paper. The task was to tie the two ropes together, which were too far apart to grasp simultaneously (depicted in Figure 3). The solution to the problem is to use the pliers as a weight, swinging one rope like a pendulum while grasping the other in order to bring them close enough together to reach. Maier found that subjects frequently had a difficult time finding this solution, and attributed this to the fact that it required using an object in a different way than is usually considered for that object. This idea, called "functional fixedness," argues that when an object has a known, typical use, that it will be considered with that use in mind, at the expense of other possible affordances it might provide. In other words, knowing that a pliers is used for gripping and rotating apparently makes it harder to consider that it can also be used as a simple weighted object. Additionally, Maier demonstrated the usefulness of hints to overcome fixedness, finding that "accidentally" brushing against a rope and causing it to swing would lead 23 of 37 participants who hadn't solved the problem after 10 minutes, to do so in the next 60 seconds (Maier, 1931).

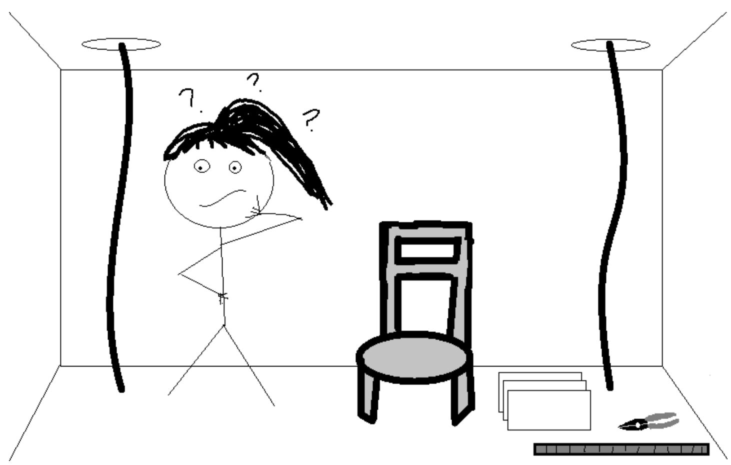

Figure 3: The Two-String Problem features two lengths of cord hanging from overhead. They are far enough apart that one cannot reach one while holding onto the other. Given a number of objects to use, such as a chair, pliers and paper, the subject is asked to use the items to help them tie the two strings together. (Guidone, 2018)

An often overlooked aspect of this now famous problem is that the room actually contained several kinds of objects, offering multiple potential solutions by which the strings might be joined. It was only the pendulum-style solution the experimenters were really interested in, as it was the less obvious of available possibilities, as well as the least-often found, and specifically required non-traditional uses for the available tools. However, the other solutions illustrate an aspect of how the problem is perceived, which relates directly to the problem specification component of problem finding. Different solutions, such as holding down one rope, tying a cord to one to make it longer, or using a pole to reach the other rope further demonstrate different perceptions of where the difficulty in solving the problem lies, or rather, various specifications of a problem from a common starting point. If one's reach seems too short, one may look for a different solution than they would if they were considering the ropes to be too short. Considering what makes a problem difficult was explored later by Newell and Simon.

Another commonly referenced insight problem is known as the "9-dot problem," first described by Maier in 1930. For this, a subject is presented with a set of nine dots arranged as a grid, and asked to connect them by drawing no more than four straight lines without lifting the pen.[16] Most people struggle quite a bit to solve this task, and "[t]he expected solution rate for this problem under laboratory conditions (e.g., a time limit of a few minutes) is 0% (MacGregor et al., 2001)." (Kershaw & Ohlsson, 2004). Those who do manage to complete the task typically experience the "aha!" moment of sudden insight, and it has been explained by many (such as Duncker, Maier, Ohlson, Newell and Simon, etc.) as a restructuring issue; initially subjects are constrained to stay within the bounds of the nine dots whereas the solution requires going outside them. It's almost another form of functional fixedness, except that instead of being prevented from finding

[16] In this case, the goal is essentially clear, like the Tower of Hanoi, but not entirely, since the necessary arrangement of lines is unknown.

a solution due to the assumption of how an object is meant to be used, it is an assumption of limited space available.[17] As it did with the two ropes problem, giving subjects a hint has proven to be the catalyst to finding the solution for the 9-dot problem (Weisberg & Alba, 1981, and others), but even with a hint the majority of individuals are unable to come up with the solution. In Weisberg and Alba's study, control subjects were given 20 attempts to solve the problem, while another group was given 10 attempts followed by a restructuring hint, and 10 more attempts. "No subject in either condition solved the problem in the first 10 tries, and no subject in the control condition ever solved the problem (excluding those who had seen the problem before). However, 20% of the restructuring hint group solved the problem in the second 10 tries" (Novick & Bassok, 2005).

For both the 9-dot and two-string problems, subjects' inability to easily find a solution is predicated on some obstruction in their grasp of the problem. Either there are constraints being placed on what can be done stemming from the assumption of rules not stated, or there are possible uses that are not considered due to expectation of use from previous encounters, meaning, for instance, that one will be unlikely to consider using match as a writing instrument because that's not how it is typically used.

Another example of prior experience influencing how problems are solved comes from A. S. Luchins' water-jug problem (Luchins, 1942). This problem is frequently used to illustrate a phenomenon known as the *Einstellung*, or set, effect, in which people display a tendency to stick with a known solution strategy despite the availability of

---

[17] Which are both cases where the possibilities aren't entirely clear from the outset.

simpler choices. In the water-jug problem, subjects are presented with the task of measuring out a desired amount of water by using a combination of three different sized containers.[18] The steps to the solution to the first in a series of these problems also works for the second, third and so on, such that nearly all of them can be solved the same way (see Table 1). Some of them, however, can also be solved more easily, and in fewer steps, yet participants typically don't apply the simpler method. As long as the one they've been using still works, they tend to favor it over something different. Additionally, the instances that can't be solved with the same technique prove to be more difficult for subjects, and are solved less often, despite having simpler solutions than the established strategy.[19] While it's not quite the same as functional fixedness, the inability to use the tools available in a different way than you've been doing certainly seems related.

[18] If, for example, you are given a 9 gallon container and a 4 gallon container, then asked to measure out 5 gallons, you can do so by first filling the 9 gallon container and then emptying it into the 4 gallon container until the second is full. What remains in the 9 gallon container is 5 gallons.

[19] What this means is that a particular instance of the problem could be given on its own, and a subject would find the solution, but if that same instance comes after a series of other instances that all share a solution method, they will have a much harder time--even if the solution to the different instance is actually easier than whatever they've been doing already.

Table 1: All problems but problem 8 can be solved with B-2C-A, but 7 and 9 can be solved with A+C; 6, 8 and 10 with A-C (Adapted from Luchins, 1942).

Problem	Capacity of Jug A	Capacity of Jug B	Capacity of Jug C	Desired Quantity
1	21	127	3	100
2	14	163	25	99
3	18	43	10	5
4	9	42	6	21
5	20	59	4	31
6	23	49	3	20
7	15	39	3	18
8	28	76	3	25
9	18	48	4	22
10	14	36	8	6

Karl Duncker first introduced a now-classic insight problem in 1945: the “candle problem,” also called the “box and tacks problem” (Duncker, 1945). For this challenge, subjects were presented with a candle, a box, and some tacks, and asked to mount the candle to a wall such that burning it wouldn’t drip any wax on the floor (see Figure 4). He presented the materials to one set of participants by putting the tacks, candle, and matches into the box, while providing the items separately to another set of participants. In a strong example of functional fixedness, the participants given all materials in the box found the problem much more difficult than the other subjects, as they tended not to consider the box as a component they could use, seeing it only as the means of providing them with other tools (in other words, seeing the box as only a box). This study was replicated in 1952 by Robert Adamson, showing participants given empty boxes to be twice as likely to be able to solve the problem (Adamson, 1952).

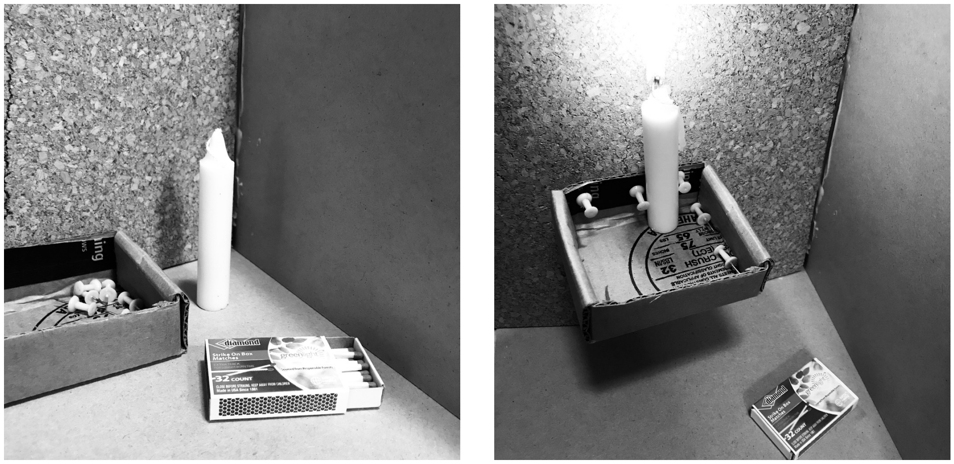

Figure 4: The Box and Candle Problem. On the left, a box of tacks, a box of matches, and a candle are given to a subject who is then asked to affix the candle to the wall in a way that won't allow the candle to drip. On the right, the solution to the problem, using the box as a platform for the lit candle.

In 1951 Birch and Rabinowitz investigated the effects that familiarity can have in novel circumstances, particularly highlighting the difficulties it can cause (Birch & Rabinowitz, 1951). While being an expert for a type of recognizable problem can indeed lead to solving similar problems more quickly, they showed that it can also make it much more difficult to solve problems that require a different use of familiar components. This conclusion has come up again in subsequent work (as well as re-examinations of earlier research) as one of the classic issues of insight problems: perceived or self-imposed limitations on the circumstances can prevent the finding of a useful solution.

2.2.2 Insight vs. Noninsight

With insight problems it seems as though people arrive at solutions suddenly, rather than by gradual progression. However, do they qualify as a genuinely distinct category of problem? This question extends back to when the concept of insight was first formulated. Wertheimer described the different ways a single problem may be interpreted, as did Maier and Dunker, who suggested along with Ohlsson in 1984 that the insight component of insight problems is a change of perspective, or the framing of a problem (Duncker, 1945; Maier, 1931; Ohlsson, 1984; Max Wertheimer, 1959). This would suggest that insight and noninsight problems aren't all that different, but that instead an insight problem is simply one that is improperly understood, or misinterpreted.[20]

One of the first efforts to rigorously distinguish insight from non-insight problems was carried out by Metcalfe and Wiebe (Metcalfe & Wiebe, 1987), who tried to distinguish between insight and noninsight problems by demonstrating the suddenness that appears to occur upon finding the solution to an insight problem as opposed to the more gradual progression associated with noninsight problems. Choosing algebra as an example of noninsight, they presented subjects with one or the other type of problem, and asked them every 15 seconds to report, on a scale of 1 to 7, how close they felt they were

[20] In this view, the only difference between insight and noninsight problems is correctly grasping the question--in other words, a problem that is hard because it was misread could be easily solved if the subject realizes the mistake, and this would account for the "aha" moment.

to finding the solution. Their results (Figure 5) do indeed support the general description of insight problems as being solved in a much more sudden burst.

Figure 5 shows the way these results have typically been presented in subsequent reviews.[21] The vertical axis represents how close to the solution each subject was (or, more accurately, how close they reported feeling up until they reached the solution), with each level higher indicating a step closer to the solution. The horizontal axis shows the time, in seconds, prior to reaching the answer. The dashed line shows the gradual progression through the algebra problem over time, from 3 to 4 to 5 to 7 on the confidence scale, while the solid line depicts the much more rapid ascension from 2 to 3 to 7 for the insight problem.

[21] The initial study used a series of bar graphs to portray their findings, but it isn't as visually obvious as the plot.

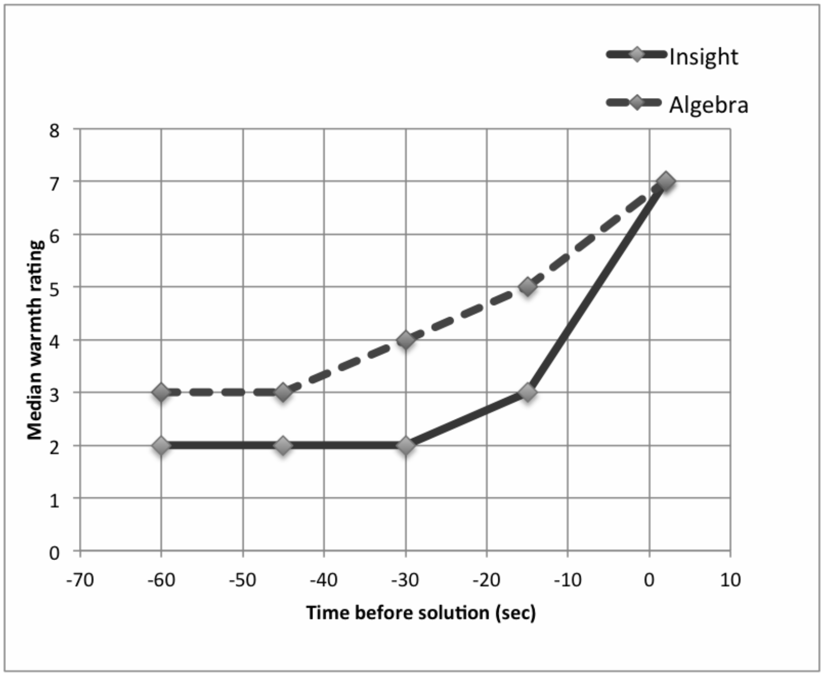

Figure 5: Graph of Insight vs. Non-Insight showing the results of Metcalfe and Wiebe's (1987) experiment on feelings of closeness. The dashed line corresponds to subjects' reported feeling of closeness in ten second intervals prior to solving an algebra problem. The solid line indicates the same report for an insight problem, showing a much more sudden rise from uncertainty to solution.

Later research built on the "closeness" idea that subjects report, often demonstrating that individuals have a sense of how far along they are even without knowing the end result.[22] However, it seems that this feeling of closeness may not be entirely accurate as a measure of how close one really is. "Research into intuitive problem solving has shown that objective closeness of participants' hypotheses were closer to the accurate solution than their subjective ratings of closeness" (Reber, Ruch-Monachon, & Perrig, 2007). In other words, people are actually closer to the solution

[22]When working through a problem, people can often express a sense of being close to, or far from, a solution before they've found it. In some cases, they can even guess at the answer before they reach it, with surprising accuracy.

than they think they are. Metcalfe and Wiebe had compared reports of how close people thought they were to how close they actually were for algebra problems, and found them to be comparable, so they assumed that the report of closeness for insight was similarly indicative of actual closeness, which Reber et al., disprove in other cases.

This idea of competing activation of unconscious processes has been described in the insight literature as an explanation for sudden breakthroughs—they weren't sudden at all, but were gradually being built up separately until reaching a threshold where the secondary interpretation overtakes the original (unsuccessful) view of the problem. There is evidence that alternate ideas persist in categorization, effectively demonstrating that we keep other possibilities in mind even after having made a categorical decision, which fits the "multiple competing interpretations" theory of insight quite well (Dieciuc et al., 2016).

The question of whether insight really is sudden or not has recurred in the literature with arguments on both sides. Durso et al. (Durso, Rea, & Dayton, 1994) investigated this issue using the barroom puzzle: "A man walks into a bar and asks for a glass of water. The bartender points a shotgun at the man. The man says 'Thank you,' and walks out."[23] The solution to this problem typically pops into mind suddenly and fully intact, accompanied by an irresistible feeling of "aha" Moreover, the solver has no awareness of incremental progress toward the goal such as that which accompanies

---

[23] This brings up a question of puzzles versus problems. As described, there doesn't appear to be anything to solve in this case. The question seems to be missing, but it would be something like "At first, this scenario is confusing and doesn't make sense, but with a little more context, it could. What additional information would resolve the apparent confusion, and make the scenario make sense?"

search solutions"[24] (Novick & Bassok, 2005). He gave several pairs of words to subjects, initially starting from before reading the problem, then right after, and finally at ten minute intervals until they found the solution. At each point, he asked whether the pairs of words were similar or dissimilar, (examples of word pairs include pretzel/shotgun and surprise/remedy). They had already established that those who solved the puzzle saw the relationships in the word pairs while those who hadn't did not.

To be clear, participants in this study did not feel like they were getting closer to a solution at any point—they struggled until having a moment of insight, at which point they had an answer. The results from responses to questions about how related pairs of words were, however, suggested a gradual progression to the solution, as the word pairs were first reported as unrelated, then partially related, then completely related. This somewhat contradicts the conclusions of Metcalfe and Wiebe (Metcalfe & Wiebe, 1987), but was supported by subsequent work by Novick and Sherman (Novick & Sherman, 2003). Concerned with the question of whether the repeated polling was, itself, interfering with the process, they did another study in which they asked subjects to quickly decide whether anagrams could be unscrambled into English words. The concluded from that study, that "pop-out solutions arise gradually through the accumulation of relevant partial information (also see Bowden & Jung-Beeman, 2003)." In other words, insight problems are also solved by gradual processes, but such processes are not consciously available.

[24] This is taken to mean that for problems that require a number of steps to complete (like the tower of Hanoi, or a maze), there is some awareness either that a move has brought you closer to the goal, or that at least it narrows down the path that will. In either case, there is a sense of progress that is lacking for insight problems.

2.2.3 Analogous Problem Solving

It is possible that some of the difficulty in solving insight problems stems from a lack of relevant experience with similar situations. Gick and Holyoak (Gick & Holyoak, 1980) took a closer look at this by providing analogies that could be useful for dealing with something unfamiliar, effectively creating a different type of hint (cf. the discussion of Maier and the two-string problem above). They took Duncker's (Duncker, 1945) radiation problem (destroy a tumor without destroying surrounding tissue) and replaced everything in the setup to create a functionally identical situation, in which the same solution could be applied, but that it was more obvious what the solution was when put in different terms relating to an army scenario (troops must invade a city without revealing themselves). This is referred to as analogical transfer: taking a solution from one scenario and applying it to another.

The radiation problem consists of presenting the subject with a simple image of a filled-in oval with a larger oval around it (see Figure 6). The inner oval represents a tumor that must be destroyed, and the outer oval is the body containing it. Subjects are told that a laser can be used to destroy the tumor, but that it will also damage healthy tissue, and that their task is to indicate where lasers should be positioned in order to only affect the tumor. Most individuals struggle with, and fail to find a solution to this task without some extra guidance. The solution is to use multiple lasers at once, such that the lower intensity won't hurt the healthy tissue, but crossing beams within the tumor will increase the intensity to an effective level, only within that area.

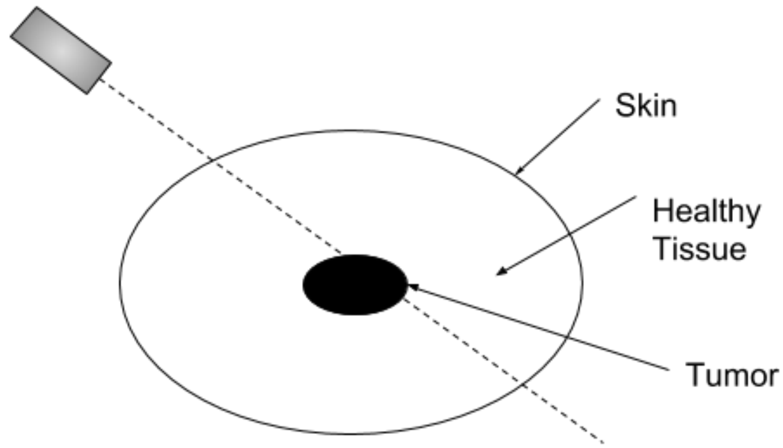

Figure 6: The Tumor and Laser Problem. Subjects are shown a version of this image, and told that a laser strong enough to damage a tumor would also destroy healthy tissue, but a lower powered one that would leave the healthy tissue alone would be insufficient to destroy the tumor. They are asked to figure out how to use the laser to damage only the tumor.

In the analogous case, subjects were told a story in which an army attempted to invade a city, but couldn't take the entirety of the forces along the main road because they would be discovered. Instead, they split up and came from multiple directions, regrouping once they reached the city. The subjects provided with this story then went on to apply similar logic to the tumor, introducing multiple lasers that would concentrate on one spot. Later studies (Grant & Spivey, 2003; Thomas & Lleras, 2007) using this problem have shown that it isn't necessary to have an analogy in order to find the solution, and that where a subject's gaze is most often directed is linked to how likely they are to solve the problem quickly. Another related study substituted visual diagrams for the story, showing a single line of large arrows pointed at a spot followed by several small arrows pointing to a single point of convergence (Pedone, Hummel, & Holyoak, 2001). In this case, having been given the diagram before the problem increased the rate of solution, and offering a hint to the subjects that the diagram might help them also increased their success rate.[25]

> An important factor that mediates spontaneous retrieval and use of analogous solutions is people's understanding of the learned example. Chi, Bassok, Lewis, Reimann, and Glaser (1989) investigated this issue in the domain of physics, using problems from elementary mechanics. They found that learners who understood the logic of textbook examples spontaneously applied the example problems' solutions to analogous test problems that differed from the learned examples in many respects. However, poor learners failed to recognize the structural similarity between the examples and the novel problems (Novick & Bassok, 2005).

In general, if the reasoning behind an example is well understood, then it is more likely that the same kind of reasoning will be used in other circumstances. A textbook might give an example of a penny being dropped from a building, and explain how to determine the building's height by how long it takes the penny to hit the ground. If the next problem involves a skydiver that takes a set amount of time to fall and asks how high up they were, those who understood the components of the first problem should have no problem with the second, while those who don't see the similarities will find the second to be just as difficult as the first.

---

[25] In the analogical transfer cases, participants are effectively given the answer for the analogous case, either directly in the story (Gick & Holyoak, 1980) or visually in a diagram (Pedone, Hummel, & Holyoak, 2001), although they might not understand the meaning at first in the later version. It would be interesting to test this capacity by first presenting a problem, and then, when the subject reaches an impasse where they cannot find the solution, to offer them an analogous problem. If they can solve the analogous problem despite not solving the main one, this would go a long way towards supporting the effect of how a problem is presented making a difference in solvability. If they could then return to the original problem and make the connection between the two to solve it, the use of analogy for problem solving and for education in general would be further supported. Such an experiment may be difficult to perform since it would require a way to determine the point of impasse, but with the proper controls in place, it seems like it would be both possible and beneficial to do.

2.2.4 Experts

Building off the ability to transfer from one problem to the next, it is a known phenomenon that "experts" can solve novel problems with relative ease if the new problem has characteristics of those that they are well acquainted with. Above and beyond simply applying solutions that are already known to new scenarios, they seem to rapidly pick up on the features that are important and necessary whereas novices have a much more difficult time discerning what aspects of a problem are significant. As described by Novick & Bassok (2005):

> A number of studies have found that experts' attention is quickly captured by meaningful configurations within a presented stimulus, a result that calls to mind the Gestalt view that problem solving is related to perception. In contrast, novices' attention is focused on isolated components of the stimulus. Perhaps the earliest research investigating this issue comes from the domain of chess (Chase & Simon, 1973; de Groot, 1966). In the typical study, subjects view 20 or more chess pieces arranged on a chess board for 5 seconds and then have to immediately reconstruct what they saw on a new chess board.

In the study described above, experts were typically able to recreate what they had seen, while newcomers could not. An explanation for the discrepancy is that experts would recognize groups or arrangements of pieces, owing to experience with the game. Novices, in contrast, would try to recall each piece on the board individually, and would be unable to do so with much accuracy.

Examples of studies on experts in other domains include “Circuit diagrams (Egan & Schwartz, 1979), computer programming (McKeithen, Reitman, Rueter, & Hirtle, 1981), medicine (Coughlin & Patel, 1987; Myles-Worsley, Johnston, & Simons, 1988), basketball and field hockey (Allard & Starkes, 1991), and figure skating (Deakin & Allard, 1991)” (Novick & Bassok, 2005). Although the specifics may vary according to their field, these studies have all shown that experts are able to perform faster and more reliably than beginners.

As in the chess example above, experts appear to direct their attention differently over time. Relations to attention “are a matter of emphasis and degree. With increasing expertise/knowledge, there is a gradual change in the focus of attention and in the problems that are seen as related, and the extremes are not quite as extreme as summaries of the differences often suggest (e.g., Deakin & Allard, 1991; Hardiman, Dufresne, & Mestre, 1989)” (Novick & Bassok, 2005).

Accordingly, expert problem solvers tend to focus on specific, relevant components of a problem more than novices, whose attention is spread more widely. In other words, not only are they able to apply solutions to similar problems, but they readily attend to the features that are necessary to the solution. One explanation for the difference between how well (or frequently) experts may solve a problem as contrasted with novices is that the experts are more familiar with recognizing analogous situations to ones they know how to solve. Someone who has seen numerous problems of the same sort will be more likely to recognize a new one that follows the pattern, and attempt the same solution. In this view, experts aren’t so much better at solving problems as they are good at noticing that a problem is of a known form and applying a previous strategy to it.[26]

2.2.5 Procedural Problems

While the emphasis on insight problems came from a desire to reveal different stages of problem solving (such as reconsidering what is known, or finding easier paths), other work has shifted away from insight in favor of simply studying multi-step problems. Insight problems have consistently been harder to define than other types of problems[27], and are difficult to model. Solving something like algebra problems, or the Tower of Hanoi, on the other hand, represent tasks that can be accomplished by using known rules applied to a situation with a clear goal. Wertheimer would take these sorts of problems and make systematic changes to them in order to determine how well the solution is ultimately understood, and he preferred to stick to these more concrete scenarios over "insightful" ones. The solution to a noninsight problem isn't necessarily immediately apparent, but it doesn't require outside knowledge apart from the situation and instructions given. There are, of course, various levels of expert for a given class of problem, but much of the expertise can be characterized as simply being familiar with the

[26] It is worth noting, however, that this can backfire as well. If a problem is *not* of the sort that an expert is accustomed to, but has enough similarities to appear as though it is, the expert may in fact have a much harder time with it. This is related to the *Einstellung* effect, but to the point that experts fail to realize that the problem itself is different, not just that their solution won't work.

[27] There does not appear to be a single agreed definition of an insight problem. There are, however, a variety of similar descriptions.

type of problem and how to solve it, rather than an intrinsic ability to solve it “better” than those encountering it for the first time. In fact, the familiarity aspect gives rise to two subsequent investigations: expert systems, and the hampering effects of experience.

Unlike typical insight problems, there is no need to reframe the situation or do anything differently than expected for these multi-step tasks. Duncker, in 1945, called these “thinking problems,” and Newell and Simon referred to them as “move problems” almost 30 years later (Duncker, 1945; Newell & Simon, 1972). With a more concrete problem, Newell and Simon decided to break down the steps involved in working through the search-space of possible actions. Search-space, or problem-space, is a way of describing all of the possible states that could exist for a given problem, and how they relate to one another. At any given moment the problem is in a particular state—the arrangement of pieces on a game board, for instance—that is different from other possible states. There are a limited number of options (choices, or moves that can be made) at any one state, and each leads to another state. Moving through the states of a problem with the intention to reach a specific one (the goal) can be thought of as “searching” through the problem-space. In order to effectively navigate the search-space, Newell and Simon proposed dividing a large goal that one doesn’t yet know how to achieve into smaller steps, or subgoals, that could incrementally move towards the solution. This logic, as applied to the Tower of Hanoi, could be expressed as follows: the goal is to move the whole stack to the rightmost peg, but only one disc can be moved at a time, so the first subgoal is to move one disc to the rightmost peg, because that will at least be closer to the goal than where we started. There has since been a great deal of work, particularly in the area of computer science, that has demonstrated different strategies for most efficiently navigating a search-space, and as long as a problem is well-defined enough to be described in these terms, it is possible, although not always reasonable, to try to solve it with one of these methods.

2.2.6 Understanding vs. Repeating

In his book *Productive Thinking*, originally published in 1945, Max Wertheimer wrote about the importance of understanding over simply memorizing and replicating (Max Wertheimer, 1959). According to Wertheimer, students are regularly "taught" by repeated examples to such a point that they can solve similarly structured problems by following the prescribed steps, but fail to understand what they are actually doing, or the relationship between the components of the problem, and therefore are incapable of solving *very similar* but functionally different problems.

Wertheimer demonstrates the concept of understanding in a number of ways, including asking students to find the area of a parallelogram (which they had just learned to do), but rotated by 90°. If they understood the relationship of the height and width, and that one end corresponds to the other end, then, he argues, it shouldn't matter how the shape is oriented. He found, however, that students tried following the identical instructions for drawing a line down from the top of the shape, and another one over to it, and since that wouldn't work when the parallelogram was turned, they were unable to find the area (see Figure 7). The same holds true for other situations, ranging from algebra to sums of series—in each case he showed that a grasp of how the components relate to one another should allow any variation to be solved with equal ease, but that most students would struggle when problems look more difficult.

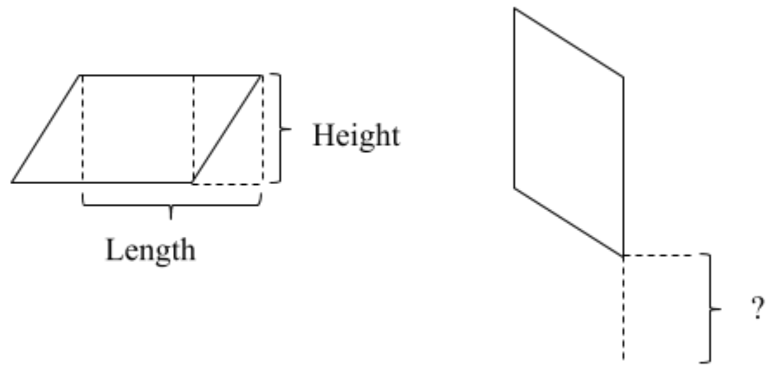

Figure 7: Parallelograms. On the left, an illustration of how to find the area of the parallelogram by calculating the length by the height. A typical lesson shows a line drawn to the right from the bottom right, and down from the top right to form the rectangle. On the right, the same method of adding lines won't work for the same figure when rotated 90 degrees.

One of the most intriguing things that Wertheimer points out is that the same problem, with the same solution, can be thought of in different ways. Depending on how the problem is considered, the meaning of the parts of a formula can differ greatly—n could mean, for example, "the middle number from a group" or "half the number of pairs in that group." The formula works out the same way in either case, but the interpretation and understanding of it is very different, and this difference in conception explains how some problems appear easier or more difficult to different people, even though they consist of the same parts. Combined with the previous assertion regarding the understanding of how parts of a problem relate to each other, this provides something of a framework for how two people could have the same information to work with, and come to different solutions to the same problem; they view the problem differently. This all echoes back to Maier's two-string problem with its multiple possible solutions, along with the significance of problem-specification within the problem cycle.

Chapter 3: A Taxonomy of Problem Types

In this chapter a new classification system is presented, which describes problems as distinct from one another on the basis of what information is available to the problem solver. A variety of problems that are difficult in different ways were selected, and attempts were made to try to specify the aspects of each problem that made them challenging.[28] By analyzing what information is known about these, what methods may be used to work on them, and what makes them hard, a few different dimensions that distinguish problems from one another were found. While other researchers have sought to either divide problems into categories by descriptive means (for example, problems that can be solved by following a series of steps versus problems that can't be solved without changing the consideration of the problem) or by determining whether a problem is well defined or ill defined, the present system accounts for specific ways in which problems may be more or less well defined, and suggests that the difficulty of a problem is associated with both how well defined it is, and what information is actually known.

[28] These included the nine-dot problem, Tower of Hanoi, two-ropes, math, Rubik's cube, path finding, combinatorial problems, matchstick problems and more.

In a similar way, David Jonassen (Jonassen, 2000) emphasized that different problems involve different cognitive processes and skills, and should therefore be taught in different ways in order to be effective. Jonassen considered a variety of problems, looking for the features that stood out to make them more or less well defined and difficult, and arrived at a set of 11 categories of problem, which he arranged on a continuum from well-structured to ill-structured. In addition, he included a number of descriptors alongside each group to indicate a potential range within these categories. However, his categories do not directly specify what aspects or features differ from one type to the next. For example, his first type of problem "Logical Problems," corresponds to five separate types of problem in the present taxonomy. The taxonomy developed here is also more explicit about how the variations among problems work.[29]

Other taxonomies exist in the literature. Several researchers have divided problems into various groups and classes based on particular qualities they may share or lack. Among those who have introduced problem taxonomies are Newell & Simon (Newell & Simon, 1972), Greeno (Greeno, 1974), Greeno and Simon (Greeno & Simon, 1988), Perkins (Gardner, 1990), Kaufmann (Kaufmann, 1990) and Goel (Goel & Pirolli, 1992). Each of these systems of classification are built with the assertion that problems can be fundamentally different from one another, requiring different means of approach and with inherently different levels of difficulty.

---

[29] It should be noted, however, that both Jonassen and the current taxonomies are meant to allow for variation, rather than being considered absolute divisions.

3.1 Elaboration of Problem Features

The taxonomy presented here is organized around four dimensions: "Possibilities," "Course," "Progress," and "Goal-Reaching," each of which can be either clear or unclear. Considering all possible combinations of clear and unclear for these four dimensions yields a list of sixteen distinct problem types.

Clear possibilities refers to whether or not the agent is fully aware of all the possible choices available to them in solving the problem. This amounts to knowing the rules in a game or the tools (and their uses) at the problem solver's disposal. This is a significant feature of insight problems, which are often difficult specifically because the individual believes that they know all of their options when they do not (see Insight 1 and 2 below). Amarel describes "The *rules of action* in a system *P* specify a possible next situation...as a function of certain features in previous situations" (Amarel, 1981). In his paper, Amarel makes five progressive manipulations to the representation of a problem to make it easier to solve, and the first of these is to define and constrain "...a unique choice of a set of feasible actions" (Amarel, 1981).

When possibilities are clear (a "Yes" in this column) it is clear at each stage of the activity what you can or cannot do. Board and card games are typically good examples of problems in which the possibilities are clear—at any point in the game there are known, well-defined rules governing a player's options. Cases of unclear possibilities include playing a game without first learning all the rules,[30] operating equipment without training or a manual, and social interactions.

Insight problems (which have been split into two types) typically rely on someone *incorrectly believing* that they know what their options are, and must overcome the self-limitation in order to consider solutions that will actually work. Many more general problems (those in day-to-day life for instance) are not well constrained at all. If you want to get somewhere on time, for example, you may be able to use various forms of travel and leave at different times, including ones you may not at first consider. It's not a clearly delimited set of choices for those problems, which can be a stumbling block—how can one solve a problem without knowing what they are allowed to do?

Amarel lends support to the importance of knowing ones options by examining how a problem can be made much easier to solve. Using the missionaries and cannibals problem as an example, he notes that "the verbal statement of the M&C problem does not induce a unique choice of a set of feasible actions" (Amarel, 1981). In other words, from the description of the problem alone, one's *possibilities* are *unclear*. The very first thing he does in order to simplify the problem is to examine the scenario to determine "…a 'reasonable' set of elementary actions that are assumed to be feasible and that satisfy the given constraints" (Amarel, 1981).

Clear course refers to whether or not the agent typically knows what to do in order to progress through a problem. A "yes," here means that there usually isn't a struggle to determine how to proceed because the next necessary action is relatively

---

[30] Which is common for complicated board or card games, and particularly for role playing games, which can have multiple books of rules and can be overruled by whoever is running the game.

obvious. In a game of poker, for example, the possibilities are clear (call, fold, raise, etc.), but the course is not. The game "war," on the other hand, always has a clear course, as there are no real choices to make at all. How clear the course is can change at different junctures along a problem, so describing a problem as having a clear course is meant to convey that it is generally clear what to do more often than not.

A number of researchers have depicted problem solving as navigating through a web of possible states in search of one in particular. Lovett and Anderson explain that numerous potential routes to a solution, and that "Problem solving can be viewed, then, as finding one of the few paths that lead from the problem's initial state to its goal state…however, at each step in the solution there are typically several operators that can be applied, thus forcing a selection" (Lovett & Anderson, 1996). By this understanding, a problem is harder if there isn't a known method for choosing which path to take in pursuit of a solution. Duncker also identified a lack of a clear course as part of his definition of a problem itself, saying "The subject has a goal; the goal becomes a problem if it cannot be reached immediately by obvious actions" (Newell, 1980). One of the most direct references found in support of using clear course in the present taxonomy comes from Sternberg, who linked how well a problem is defined to this very factor, stating "Ill-defined problems are characterized by their lack of a clear path to solution" (Davidson & Sternberg, 2003).

Clear Progress refers to whether or not one can distinguish if their efforts are getting them closer to the solution.[31] For a jigsaw puzzle, progress is obvious—if more of the puzzle is getting solved you can see it. Navigating a maze from the inside, however, makes it difficult to track progress—you don't know if you're getting closer to the end, or to a wall.[32]

Curiously, explicit references in the literature of problem solving to the sense of progress being important have not been found, but it has nonetheless been a component of prior work. The feeling of closeness reported in the study of insight problems, for example (see 2.2.2), implies that a sense of progress plays a part in assessing a problem, and a fair bit of the literature on computer based problem solving involves making "cost" decisions for how to search through a problem space by comparing where they are to where they'd like to be, and seeking out the moves that bring them closer at each move.

The last distinguishing feature is the determination of a solution, or *Clear Goal-Reaching*. When arriving at the conclusion of some efforts, is it clear whether the problem has been solved? Again, for some problems the answer is yes—you can readily tell when a puzzle, a Rubik's cube or a maze have been solved. Alternatively, there are problems for which it is difficult to know if the solution was found, either because the

[31] Recall that in insight problems like those studied by Durso, subjects would frequently claim that they didn't think they were making any progress toward a solution, even though they were doing so.

[32] Another potential example to use is the game Mastermind. In this game, you must find a particular combination, typically four or five colors or numbers. There are a limited number of guesses, and after each is made, you are given feedback that tells you how many colors are right, but in the wrong place, how many are right and in the right place, and how many are wrong. You don't know which is which, though. This allows for a sense of progress, because you know you're partly right at various points, and can make your next guess in such a way as to confirm which part is right or wrong (in fact, it would be reasonable to say that your success *depends* on having clear progress).

definition of the goal was ambiguous enough to allow some room for interpretation[33], or because there simply isn't an easy way to know if the goal was satisfied or not. Did you make a good impression? Did you find the shortest route? Is the painting really "done?"

Goal reaching has been discussed in the literature by Middleton, who also chose to categorize problems by how well-defined they are, describes design problems[34] as "…a type of problem that is not yet adequately addressed within the cognitive literature. [They] are complex because they are ill-defined…and that the complexity is conceptually different from ideas of complexity in existing problem-solving literature" (Middleton, 1998). Middleton's argument is that design problems are difficult because the goal isn't concrete. "It is argued that design problems are always ill-defined in that the customer for the design is unable to specify completely, the nature of the solution required" (Middleton, 1998). In addition to not knowing in advance what the solution should look like, it is possible that a variety of potential solutions could serve equally well, meaning that there is no single state to strive for, but rather a set of conditions that aren't necessarily stated. This perfectly describes the dimension of *goal-reaching* in the present problem categories, which is used to identify whether or not it is easy to determine if the problem has actually been solved.

Another way to think of these four categories is in terms of four questions the agent can ask about a problem: "Do I know what I can do?" "Do I know what I should do?" "Do I know if I'm getting any closer to the goal?" "Do I know if I've reached the goal?" Asking these questions about a problem allows us to classify it, and specify what

[33] Ill-defined goals/problems are frequently cited as the source of difficulty in problem solving.

[34] Literally referring to problems involving design, such as clothing or artwork.

makes it different from another problem. Typically, the less that is clearly known regarding a problem, the more challenging it is to solve. Combining each possible set of answers to these questions yields a list of 16 different types of problems. For 11 of these, good examples were readily available, and so meaningful names were given to those. The ones without a name or example may be added at a later time.[35]

Table 2: Classification of Problem Types

Clear Possibilities	Clear Course	Clear Progress	Clear Goal-Reaching	Problem Type	Examples
Y	Y	Y	Y	Direct	Jigsaw Puzzle
Y	Y	Y	N	Lingering	Writing
Y	Y	N	Y	Exploratory	Water-Jug
Y	Y	N	N		
Y	N	Y	Y	Direct-Sequence	Tower of London
Y	N	Y	N	Undirected	Traveling Salesman
Y	N	N	Y	Unguided	Rubik's Cube
Y	N	N	N		
N	Y	Y	Y	Insight 1	Radiation Problem
N	Y	Y	N		
N	Y	N	Y		
N	Y	N	N		
N	N	Y	Y	Creative	Candle Problem
N	N	Y	N	Unbound	Lose Weight
N	N	N	Y	Insight 2	Two-String Problem
N	N	N	N	Loose	Medical Diagnosis

[35] In this way, it is not unlike the periodic table of elements, leaving placeholders for what may be discovered later. It is also possible that some of these qualifiers are dependent on others, which would preclude some versions from arising.

3.2 The Problem Types

While the preceding section describes the features used for categorizing problem types, this section gives a brief account of each of the sixteen types of problem, in order to delineate the unique features of each one.

A ***direct problem*** is the simplest to solve. It is one where it is clear at each stage what actions are allowable, it is clear what should be done, one's progress is clear, and it is clear when one has attained the goal (clear possibilities, clear course, clear progress, and clear goal-reaching). This is the most straightforward kind of problem—there is little to no question or ambiguity involved in addressing it. You know what needs to be done, and how to go about it, as well as whether you're making progress or have finished. A jigsaw puzzle is a prime example of a direct problem, as is the "gears problem," shown in Figure 8 in which subjects are asked to determine which way a single gear is turning based on the direction of another gear along a series of interlocking gears. In either case, the task simply needs to be done, and there may be different specific strategies, but there is no case in which the individual is stuck not knowing how to proceed.

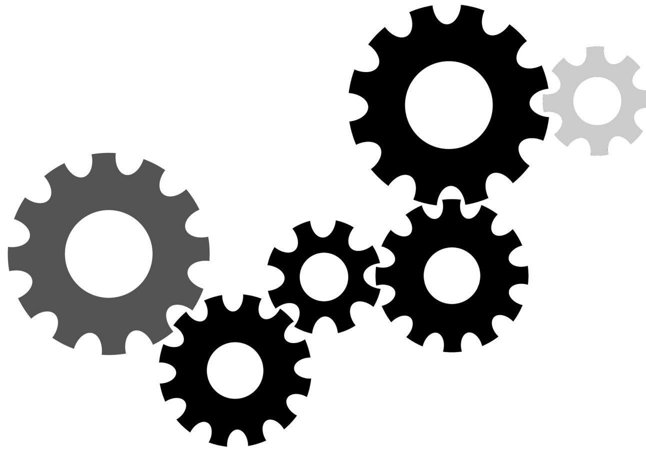

Figure 8: Gears Problem: If the leftmost gear rotates clockwise, which way will the rightmost gear turn?

A ***lingering problem*** is almost as simple as a direct one, except that it's uncertain whether the objective has been reached (clear possibilities, clear course, clear progress, unclear goal-reaching). Writing a paper, painting a picture or attempting to explain something to someone else may be lingering problems, as there are few means by which to definitively determine when you're "done."[36] Alternatively, this could also refer to long distance and/or long term projects for which it simply isn't knowable if the endeavor has been successful. A time capsule to be dug up in 1000 years, a probe sent into to space to orbit a distant celestial body, or even attempts to direct global climate may fall into the realm of lingering problems.

[36] Note that not all instances of these examples are necessarily lingering problems, but some of them certainly are. For example, a trained artist who has an established style and materials available has clear possibilities (based on what canvases, paints and brushes are on hand), a clear course (do a rough outline, fill in large areas, add details, flourishes, shadows, etc.), and clear progress (more colors, shapes, components have been included), but no definitive point at which it is clear that the painting is truly complete.

Any instance in which you know everything about the problem except your own progress is an ***exploratory problem*** (clear possibilities, clear course, unclear progress, clear goal-reaching). This means you have to play around a bit with your options, "exploring" the problem space, but within certain limits. The water-jug problem (Figure 9) is a good example of this, since it's usually not difficult to get close to the solution, but takes a little trial and error to finalize it (see 2.2.1 above for a description of the problem, and Table 1 for an example).

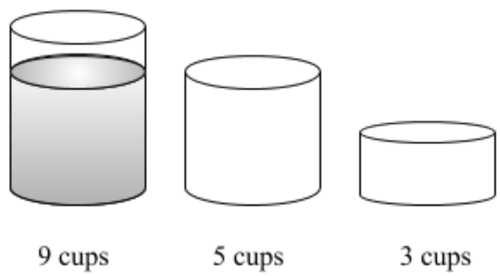

Figure 9: Water-Jug Problem: Using only these three jugs, measure out exactly seven cups of water.

The Tower of Hanoi provides a good model for ***direct-sequence problems*** (clear possibilities, unclear course, clear progress, clear goal-reaching). These are mostly straightforward in the sense of what can be done and whether the goal has been attained, but the solution relies on doing a number of steps in a specific order, and it is this order that is not always immediately obvious—particularly because it often involves taking steps that undo previous progress.[37] The Tower of Hanoi (Figure 10) asks for a stack of

[37] Other, similar problems like the Hobbits and Orcs or Rubik's cube have the same issue of needing to backtrack. In these cases, the progress may appear to be illusory, but it is still generally clear whether one is

discs to be moved from one peg to another, with the caveats that only one disc can be moved at a time, and no disc may be placed on top of a smaller disc. To complete this task requires moving discs multiple times, including putting them back on the pegs they were on before, which might be seen as backtracking. Solving these types of tasks typically involves carrying out a specific sequence of moves—things have to be done in the right progression in order to reach the goal. Ironically, the fact that there is a clear sense of progress for a task like the Tower of Hanoi may be a factor that contributes to its difficulty, as it requires making movements that clearly lead *away* from the goal.

Figure 10: The Tower of Hanoi consists of a stack of disks on pegs. The task is to move the entire stack to one of the other two pegs, moving only disk at a time, and without placing a larger disk on top of a smaller disk.

An ***undirected problem*** is one with a great deal of possible choices and for which it is very difficult to know what the "next" thing to try should be (clear possibilities, unclear course, clear progress, unclear goal-reaching). Similarly, it's hard to know that

---

closer to the end than to the beginning, because at a minimum, the current state can be compared to the goal state at any time. A possibly more intuitive example of a case where everything is clear but the course would be Pac-Man, in which the rules, progress, and goal-reaching are clear, but it isn't always obvious which way to go or what order to collect pellets in.

the goal has been reached, usually because it involves not only finding "a" solution, but also finding "the best/shortest/fastest" solution, or a solution strategy that would work in any variation of the problem.[38] The classic example of this is the traveling salesman problem (Figure 11), which asks the question of what the shortest route would be in order to travel to each area on a mapped region. The undirected part comes from lack of guidance regarding how to select paths to try, as well as the issue of not knowing the best one without testing every combination possible.

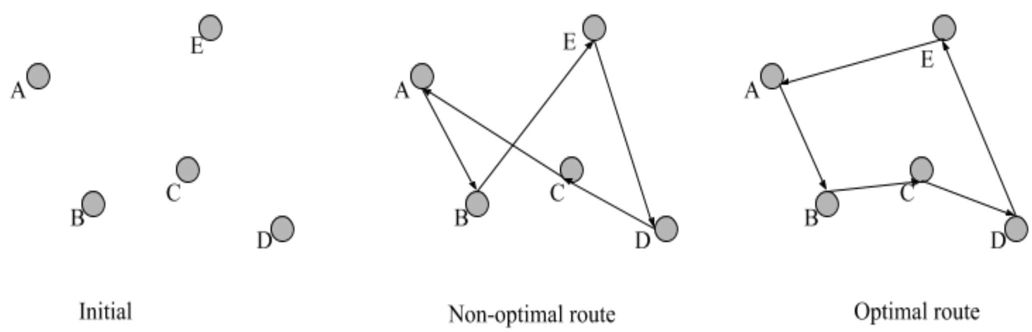

Figure 11: The Traveling Salesman Problem asks for the shortest path possible that passes through all of a number of scattered points. It is based on a salesman trying to visit all of the cities in a given area in the most efficient manner.

---

[38] Consider something like the water-jug problem, in which a solution strategy may work for many instances, but would not work for *all* instances. Another example is the Turing machine assignment given in the Philosophy of Cognitive Science class, in which students are asked to write a program that will scan forward through a series of letters until it finds a particular word, then write out their names. Many students "solve" this by searching for the first one or two letters of the target word, or matching every character to the example script. While their program produces the asked-for result, it would not work if minor changes were made to the input. This isn't necessarily an undirected problem, but it might be one in the sense that people may be mistaken about reaching the end.

An ***unguided problem***, like an undirected problem, involves not knowing what to do from a given point, but with the further complication of not knowing when progress is being made (clear possibilities, unclear course, unclear progress, clear goal-reaching). With an undirected problem, it is at least clear what paths have been taken, and even within a particular attempt, which areas have been covered or not. An unguided problem, on the other hand, lacks the signs of obvious progress, but does have a definitive end. Sliding puzzles and the Rubik's cube are instances of unguided problems—within an attempt it's not clear which moves have already been made, what to do next, or whether you're making headway at all (for the most part), but when it's solved, there's no ambiguity. There may be some sense of progress up to a point, but often there is a stage in which the agent gets stuck, for example trying to get the last few squares in place and only seeming to be shifting them around. Contrast this with a jigsaw puzzle, which lacks the instance of being uncertain whether a move helped progression toward the solution—either the puzzle is more complete, or it isn't.

Insight problems have been divided into two subtypes. The first, ***Insight 1***, describes one of the classic types of insight problems, specifically the case where the subject doesn't actually know what their options are (unclear possibilities, clear course, clear progress, clear goal-reaching). This may be a form of self-limitation, or simply not fully grasping the problem as presented. Duncker's radiation problem exemplifies this type of problem. This task (described above in section 2.2.3) asks subjects to use laser-power to destroy a tumor without harming the surrounding tissue. The solution to the problem is to use multiple lasers, which will have sufficient strength only where the beams cross. The reason people tend to fail at this problem is initial belief that they are only allowed to use one laser. Once they realize (or are told) that more lasers can be used, the solution rate rises substantially.

A ***creative problem*** involves doing something that is, at first, unintuitive (unclear possibilities, unclear course, clear progress, clear goal-reaching). It may involve self-limitation, but even when the limitation is lifted it isn't immediately clear what to do. An example of this is the box and candle problem (see 2.2.1, Figure 4), where subjects are given a box of tacks, matches, and a candle, then asked to affix the candle to the wall. Many people struggle with this problem because they assume that the box is meant to hold the tacks, rather than being a tool at their disposal. Even if they do know that the box is something they can use, however, it still takes some effort to work out how to employ it.

Unbound problems are identified by the fact that *only* the progress is certain (unclear possibilities, unclear course, clear progress, unclear goal-reaching). An example of an unbound problem may be to "lose weight" (assuming you have not specified a specific target weight) in which it isn't clear what can be done (there are many potential options), it isn't clear what the next thing to do is, and there is not a definitive goal state (how does one decide that they have successfully completed "losing weight?") On the other hand, it is clear that at any given point one is either closer or further to that goal than they were previously.[39]

---

[39] Another example is "learn a foreign language." There are many possibilities, not all of which are readily known, it isn't clear what steps to take (i.e., focus on grammar or vocabulary, etc.), and no specific point at which you could say that you've successfully learned the language, as that can mean different things (such as knowing enough to have a conversation, or to read a book, or to move to a country where that is the primary language, etc.). You do, however, have a sense of knowing more of less of the language over time, so the progress being made is comparatively clear.

Differing from the first type of insight problem, ***Insight 2*** is arguably the more common type, in which very little at all is known regarding the problem (unclear possibilities, unclear course, unclear progress, clear goal-reaching). Not only are the options nebulous, but it's a challenge to know what to try, and whether progress is being made. The 9-dot problem, some matchstick problems, or the two-string problem would fall under the heading of insight 2.[40]

Loose problems are the least well-defined (unclear possibilities, unclear course, unclear progress, unclear goal-reaching). Every part of a loose problem is open to conjecture and uncertainty. An exemplar for this is a medical diagnosis—it's typically unclear what one is looking for, there are a variety of tools and markers available, little to go on in terms of finding if one is getting closer to the right diagnosis or not, and even when one is made it may be incorrect. This doesn't mean it's impossible to make proper diagnoses, or that they are uncommon; it simply means that the task itself is unconstrained enough as to be highly case-based. Goel makes similar observations about this example, stating that "Medical diagnosis is a form of problem-solving where it is also not possible to specify the goal state with accuracy."[41] This is based on the fact that "...the features of the problem are not always easily observable and...the cause or causes of the problem may be the result of a number of complex interactions."

---

[40] Perhaps the change of one of the other features from unclear to clear could be considered overcoming the insight problem, and turning it into a problem of another type (depending on what was made clear).

[41] According to a reading of Goel, medical diagnosis might actually fall into one of the categories of the present taxonomy that don't yet have a name because a good example wasn't readily available. This would be a case of only clear possibilities. Goel's reasoning is that "Medical problems do, however, contain a reasonably well defined search space defined by procedures (solution paths) that are deemed to be legal, ethical and appropriate at any given time." In other words, the available options are clear.

Chapter 4: Experimentation

Although a number of researchers (Gardner, 1990; Greeno & Simon, 1988; Jonassen, 2000; Kaufmann, 1990 and others) have sought to divide problems into more refined categories, either by specific differences or, more often, along a spectrum of complexity, it is surprising that no previous work has been done to experimentally demonstrate that the type of problem a person solves influences the way they solve it. Researchers have focused on specific problems in order to examine how solving them may work, and these have been used in part to emphasize that not all problems are alike, as the proposed methods of solution for one don't readily work for another. However, no apparent attempt has been made to take the same general problem and recast it into different types in order to test the hypothesis that there would be changes in performance as a result. To an extent, this makes sense, since the point of arguing that "Problems are not equivalent, in content, form, or process" (Jonassen, 2000) is more readily demonstrated by identifying examples of problems that are substantively dissimilar in nature. However, the same researchers also point out that the categories they propose are not differentiated by rigid and strict boundaries. Jonassen (2000), for example, directly states this position:

> It is also important to note that these problem classes are neither absolute nor discrete. Additional analysis of hundreds or even thousands of problems is needed. Additional research may possibly identify new categories or reorganize the existing categories. Likewise, these classes are not discrete or independent of each other, that is, they are not mutually exclusive categories. So, there are necessarily similarities and overlap among the classes.

Given that the boundaries between problem type are not clearly delineated, it seems all the more important to be able to experimentally demonstrate that there are, in fact, important differences from one to the next. Using the visible light spectrum as an analogy, it would be difficult to look at a rainbow and identify precisely where the boundary between red and orange is, or between orange and yellow, but that does not mean that red, orange and yellow are indistinguishable from one another.

What follows is an exploratory study of how the information available for establishing a problem type impacts performance in a simple game. A pilot study was first performed, and an analysis of the data gathered in this initial study was used as a basis for framing specific hypotheses before conducting the second study. The results of the second study are described here. Several models of problem solving are used to frame a series of hypotheses about how participants will behave in a game depending on what information they are provided with. A simplified version of the problem taxonomy of Chapter 3 is presented and predictions are made about expected behavior based on what type of problem the solver is presented with. These hypotheses are largely confirmed, though there are important exceptions and several complex cases emerge. However, it should be emphasized again that the study was exploratory, and that multiple variables and subsets of the data were considered, increasing the likelihood of type 1 errors. In the conclusion, ways of improving the study and further lines of follow-up research are discussed.

4.1 Two Models of Problem Solving in a Simple Game

Figure 1 shows that just before problem solving can begin, the problem is built, or framed, in a certain way, and this relies in part on what information the solver has, resulting in a problem of a particular type (Table 2). This motivates an *information-based model*, according to which the ease and efficiency with which a problem is solved depends on how much and what kind of information about it is available to the solver. The more information an individual has, the faster, more efficient and more successful they should be when solving a problem. For example, given the problem of finding out whether a library has a particular book, someone familiar with the Dewey-decimal system should find the book faster than someone with no idea such a cataloging style exists. Someone who knows the layout (or has a map) of the library should be even faster and more direct. Someone who knows that libraries have comprehensive lists of their holdings should be even faster, and more efficient, as they don't have to travel to the shelf it should be on to find out if it's there. In each case, the more that is known, the faster and easier it is to answer the question.

In the domain of videogames, this is expected to manifest as players taking less time, taking shorter or more direct routes and performing better (i.e., completing more of the task, getting a higher score, etc.) when the have more complete information.

A second model emerged based on consideration of the pilot data, where it seemed that some of the game elements had an impact on behavior in the game that had less to do with the amount of information players had and more to do with how engaged the were by the game, for example the presence of elements providing direct feedback about their performance. Pilot subjects reported feeling bored or unfocused, for example, if the map was too large and information was insufficient to offer guidance. This was the basis for including *Other Factors* in the *Problem Solving* stage of the Problem Cycle in section 1.2. There is a literature regarding engagement with videogames, largely focused on how to keep players motivated enough to play for extended periods of time. The degree to which a player feels challenged and rewarded for their efforts is directly tied to the amount of time they choose to spend playing. Sharek & Wiebe (2015) explain the importance of mediating the difficulty of tasks in order to maintain interest:

> By mixing easier tasks with increasingly more difficult tasks, a game may be able to maintain player curiosity and interest while also allowing them to decompress after a particularly challenging task. For example, if a player becomes cognitively overloaded, an easier task will allow them to maintain gameplay while releasing cognitive resources so they can face the next difficult challenge.

This *engagement-based model* sometimes agrees with the information-based model, but in some cases it makes different predictions, suggesting that, in theory, these two processes largely support but sometimes compete with one another.

The *information-based* model makes several broad directional hypotheses about performance in simple games and tasks. Regarding the amount of **time** involved in solving a problem, this model predicts that it will take longer to solve a problem given less complete information. Someone assembling a device or piece of furniture with a set of instructions to follow should do so in less time than someone with no guide, or with no idea what the final construct should look like or how to use the components and tools.

Experience with puzzles and games also tends to reduce the time taken. For example, the world record for the Rubik's cube is now 4.22 seconds, while those who are not experts, even when given step-by-step instructions, take closer to 3.6 minutes (Zhang et al., 2009). In this context, experience can be understood to be another form of available information, as someone who has done the same type of problem many times over has a better sense of what to do to be successful, what to avoid, and how to evaluate their own progress than someone new and unfamiliar to the task.

Availability of information also has an impact on **performance**, in the sense of making fewer mistakes or errors and doing a better job overall. Knowing about non-standard movements like castling in chess gives a player more choices that can help them play better. By the same token, knowledge of probabilities and patterns greatly increases the chances of winning in Blackjack, where something as simple as remembering that the odds of a good outcome from asking for another card when you already have more than eleven points makes a difference.

When an agent is more confident about their situation and has more information to work with, their **movement** tends to be more direct and precise, as opposed to jerky or hesitant. Adults walk much more smoothly and deliberately than infants or patients in rehab that are learning how to move and balance safely. A dancer or gymnast can move from one position to the next in a smooth, fluid manner that those less practiced cannot. Even something as simple as walking across a dark room would be different if one was familiar with the space what to expect. Knowing that the room is large and empty of obstacles would lend itself to large strides, whereas being unaware of furniture, walls, or things on the floor results in cautious, uncertain movement.

With less definite information, there is a related increase in **confusion** and floundering. Players unfamiliar with a videogame and given no instructions just press buttons randomly to try to figure out what they can do. People who don't understand how traffic signals or elevators operate tend to press those buttons repeatedly when they don't see immediate results. With little else to go on, the same thing gets tried multiple times, like attempting to start a car over and over when it fails at first, in hopes of a different outcome.

The more that is known about a problem or task, the fewer **unnecessary actions** that are taken. There are typically a range of possible choices to make throughout the course of problem solving, not all of which are useful in finding a solution, and the more clearly the outcomes of different choices are known, the less often they are taken. This ties directly into the *Einstellung* effect (see section 2.2.4), demonstrating that when a useful process is found, it tends to be favored over looking for, or trying out, alternative solutions.

The engagement-based model makes slightly different predictions in some cases. Greater information in this case includes a familiarity with common practices in game design that rewards certain player behaviors. Even when players are aware of the course objectives (i.e., find this item, complete this task, reach the next point before a timer runs out, etc.), they typically take an exhaustive approach, trying to reach every area possible just in case there is something there worth finding, regardless of how difficult it may be to do, or how likely it is that they are expected to, and many games encourage such action by placing extra rewards in out-of-the-way places. If a game is engaging enough, players spend more time playing than they have to just to complete the game, and go far out of their way in directions that are clearly not where they game is leading them in hopes of being rewarded for doing so. Similarly, they repeat actions that don't have apparently positive results, either in the same or similar circumstances. A player might try hitting every wall in a room just to make sure there isn't a hidden reward, even if it costs them something to do so, or they might try asking a non-player character a question 100 times to see if the 100^{th} answer is different.

4.2 Experimental Setup

To test whether differences in problem solving behavior depend on what type of problem is being solved, a scenario was created in which the same general problem could be given enough variation that it would be classified into different types of problems without changing the overall task. A computer game environment was chosen for this purpose, as it allows for some simple manipulations of information given to subjects, and therefore the type of problem perceived, and to then gather detailed data on player performance.

The experiment was designed as a small, 2d game of exploration, in the style of roguelike dungeon crawlers. Games of this type are both common enough that they would not be unfamiliar to subjects, yet variable enough that expectations when playing are minimal. Such games are also adaptable in numerous ways. Figure 12 shows an example of a typical game in this genre.

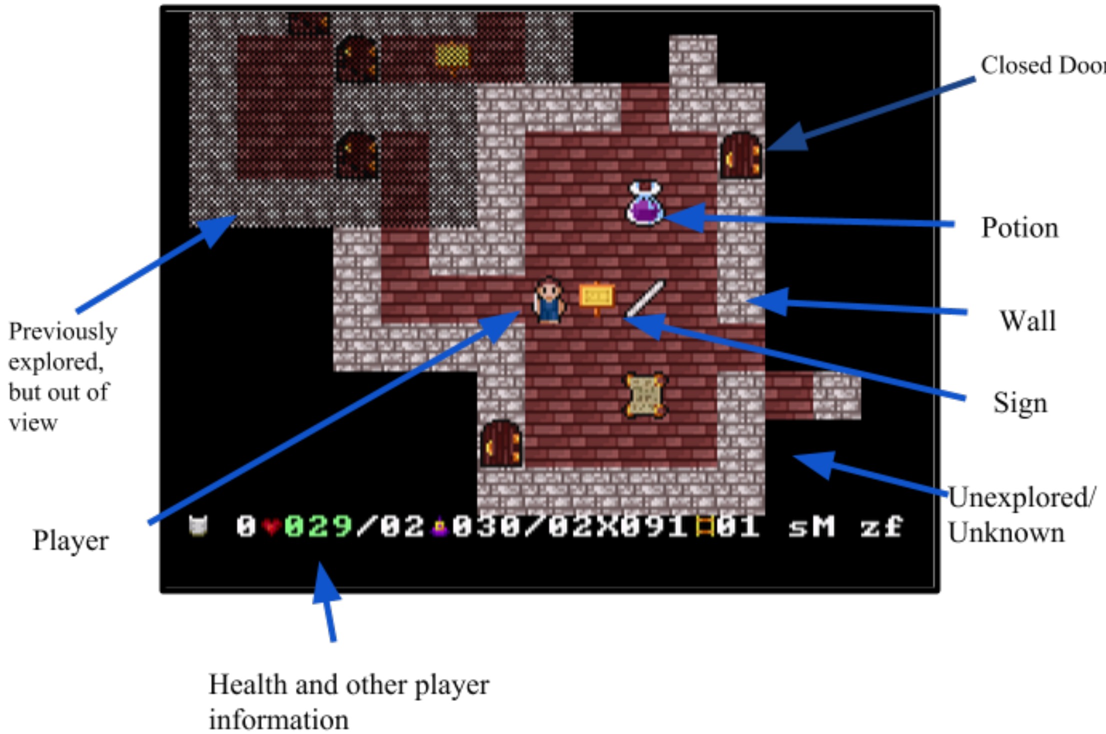

Figure 12: An example of the rogue-like game, Powder (Jeff Lait, n.d.) with a number of common elements indicated.

Environments such as this provide a world for a player to interact with, as well as rules and boundaries that govern their choices and goals. Typical tasks include collecting objects, exploring beyond the known area, solving puzzles, finding a way through a maze and completing objectives in order to proceed to other regions.

According to the *information-based* model of problem solving, depending on how much, and what specific information a player has, they should take longer or shorter amounts of time to complete a game of this type, explore more or less of a world map, move either cautiously or confidently, find greater or fewer objectives, try the same actions repeatedly or only once, and either try to move on before finishing an area, or only when they're actually done. These should all be observable from data collected during a game, and different problem types should have different information which would change these outcomes.

For this experiment, an open source JavaScript game engine called Phaser was used, (Davey, 2013/2018) along with the Tiled map editor (Lindeijer, 2011/2018) with which a unique environment was designed. All of the graphics for this project were created using pixel art tools, primarily Piskel.[42] Phaser was chosen because it runs entirely within a web browser,[43] making it possible to perform the experiment using online subjects.

The game consists of a small robot used to explore a simple 2d environment. Along with walls, there are trees, buttons, spikes and a teleporter on the game map that instantly moves the player across the board. The general problem given to each player (with varying details as will be explained) is to find and press all red buttons on the map, followed by a blue one when they are done, however it is expressed in four different forms to demonstrate how degrees of clarity lead to different solution behaviors. The map extends beyond the player's field of view, such that they cannot see the entire space at once, but must move from area to area in order to explore. The map itself is comprised of a 20x20 grid of tiles,[44] while the player's view is just over 6x6 tiles at a time (Figure 13).

---

[42] www.piskelapp.com

[43] Input to the game was only accepted via keyboard, meaning that phones and tablets couldn't be used to play (unless a keyboard had been attached).

[44] The size was chosen after a few potential maps were tried out and thought by pilot subjects to be too large.

Figure 13: This shows a zoomed out view of the entire game world. The grey rectangle in the center depicts the player's field of view at the start of the game. As the player moves, their view shifts to follow, but remains at this size (note that in the game the robot appears at about half the size shown above).

Each red button, when pressed, triggers one section of the wall to open, providing additional pathways for the player as they find more buttons. Some of these simply offer shortcuts or different potential routes. However two of the red buttons cannot be accessed without first opening part of a wall with another red button. One of the buttons is also hidden from view behind a tree. Figure 14 shows how one wall opens from pressing a button.

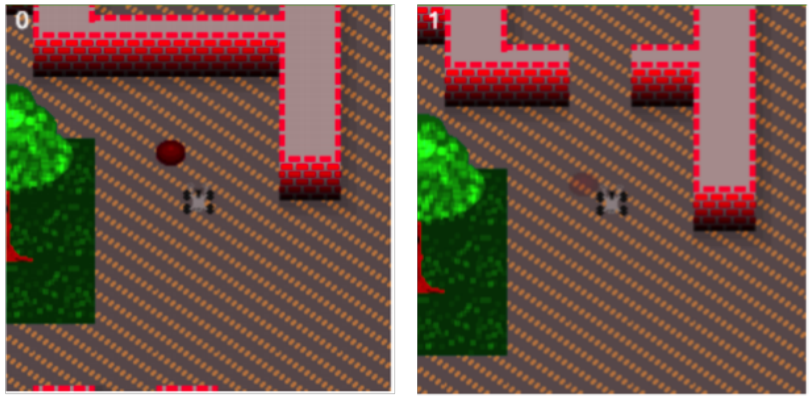

Figure 14: Opening a wall by pressing a button. Left: the button has not been pressed; Right, after pressing the button a section of wall opens.

4.3 Experimental Conditions

To test the information-based hypothesis, the amount of information available to players controlling robots in the world was varied, along with their sense of progress. In terms of the sixteen problem types presented in section 3.2, it was decided to constrain the study to four of them. To do this, *progress* was selected as one dimension that could be easily manipulated. *Course* was then made unclear for all cases by not including obvious means of choosing where to go or in what order. Finally, *possibilities* and *goal-reaching* were combined into a single dimension called *rules*. *Unclear rules* corresponds to both *possibilities* and *goal-reaching* being unclear, while *clear rules* indicates that they are both clear. The result is a smaller table comprised of four problem types, and thus four experimental conditions (see Table 3).

Table 3: The four problem variations selected for use in our experiment.

Clear Progress	Clear Rules	Name	Example
Y	Y	Direct-Sequence (Fully Informed)	Tower of Hanoi
Y	N	Unbound (Progress Only)	Lose Weight
N	Y	Unguided (Rules Only)	Rubik's Cube
N	N	Loose (Uninformed)	Medical Diagnosis

The "Name" column in Table 3 denotes problem type, i.e. which row of the larger table each of the four problems correspond to. The parenthetical names are alternative names used to formulate and discuss the hypotheses discussed in section 4.1; these names are quick to parse in terms of which experimental condition they correspond to, and useful for making predictions and interpreting results more easily.

4.3.1 Conditions

A 2x2 factorial design was used to study problem solving behavior relative to Table 3. The two factors were *progress* and *rules* (see section 4.3, Table 3), each of which could be *clear* or *unclear*.

Table 4: Experiment Conditions

		Rules: Clear	Rules: Unclear
Progress	Clear	Fully Informed	Progress Only
	Unclear	Rules Only	Uninformed

The four conditions were differentiated primarily via instructions given at the beginning of the game. In all four conditions, the following instructions were provided.

```
Welcome! In this experiment you will be given control
of a small robot. To move it, use the arrow keys, or
the W,A,S and D keys. Holding the 'shift' key while
pressing the arrow or WASD keys will allow you to move
faster.
```

Clear vs. Unclear Progress

To distinguish between *clear* and *unclear progress*, several interface items were either activated or not. In the *clear progress* case, each button that is pressed fades out, making it visibly clear whether it has been pressed, and a number at the top of the screen keeps track of how many buttons the player has reached throughout the game. This is explained to the players in the instructions preceding the start of the game. See Figure 15.

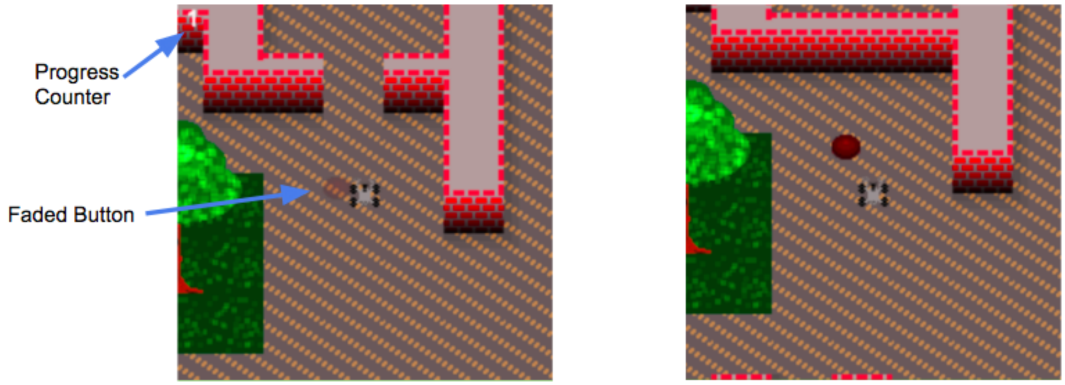

Figure 15: On the left, a case of *Clear Progress*, in which both a counter appears, and the button fades when pressed. On the right, in *Unclear Progress*, there is no visual difference between pressed and unpressed buttons, and no counter.

In addition to these differences, in the *clear progress* case, the following text was included in the initial instructions:

```
The number at the top of the screen will indicate how
many red buttons you have pressed at any point in the
game.
```

Clear vs. Unclear Rules

Subjects in the *clear rules* version were given more details about the game, and told how many buttons there were to find. The extra details explain that not all buttons are readily visible or accessible at first, that pressing a button causes parts of the wall to open, that somewhere in the game there is a one-way teleporter, and that nothing in the game is dangerous. In the *clear rules* conditions this text was added to the intro screen:

There are eight red buttons scattered throughout the world map, and your task is to find and press all of them before returning to the starting area and pressing the large blue button. Each button you press will cause a wall somewhere on the map to open, making it possible to access different areas as you progress. Not all of the buttons are readily visible, and some may be impossible to reach initially. A one-way teleporter has been added to aid your exploration. Nothing in the game can hurt you, so feel free to investigate everything.

In the *unclear rules* case, this text was used instead:

There are a number of red buttons scattered throughout the world map, and your task is to find and press all of them before returning to the starting area and pressing the large blue button.

Notice how even in the unclear case, some task information is given (thus we never have completely unstructured problems).

4.3.2 Experimental Implementation

The game was hosted on a professional server, and no personally identifiable information was recorded from users. Approval from the IRB was obtained under the "Exempt" category by including only subjects over the age of 18 who spoke fluent English, and were kept anonymous. All participants found the experiment on Amazon Mechanical Turk without any efforts to advertise or recruit for it. The task was identified with the description "Play a short video game (< 10 minutes) and take a brief survey after" along with the keywords "game, problem solving, exploring, cognition" and made available to anyone with a high (95% or more) rate of approval, indicating that they had satisfactorily completed other tasks in the past. Users were presented with a research waiver to agree to before being allowed to enter the game, and once finished, they were given a short, multiple-choice, survey to fill out.[45] After completing the survey, users were given a randomly generated code to enter into Mechanical Turk to prove completion and receive credit. Once approved, participants were compensated with $0.70 each.

The pilot study was run with 229 participants. The second study reported here was run with 200 subjects. Of these, 196 successfully completed the game along with the survey. Users were not allowed to participate more than once. The game would time out if there was no activity for thirty seconds, requiring the player to start over (logs for players who timed out were excluded from the analysis). Of the remaining 196, 9 were

---

[45] The survey consisted of nineteen questions largely based on (Wiebe, Lamb, Hardy, & Sharek, 2014) meant to test for engagement, and included two "gotcha" questions to ensure that subjects were paying attention to how they answered the questions.

removed due to suspicions of inaccurate logs, along with a further 8 that were troubling (such as spending 10 minutes pacing back and forth and only pressing a single button). Ultimately 180 complete logs were retained.

4.3.3 Logging Player Data

A script was written in the server-side scripting language PHP to log player data. The player's position in pixel coordinates (a range from 0 to 639), and which keyboard keys were held down was saved approximately ten times per second (varying somewhat due to processor speed), and the accumulated data was appended to a log file on the server every ten seconds.

To visually inspect the results of a single trial, a player's movement was plotted on top of the map of the game world, a color gradient was added to indicate the order of motion, the location of each red button was indicated with an outline, those buttons that a player had pressed were filled in, and red marks were placed at the edges of each path (see Figure 16) to indicate when an agent had changed directions or stopped an started again.

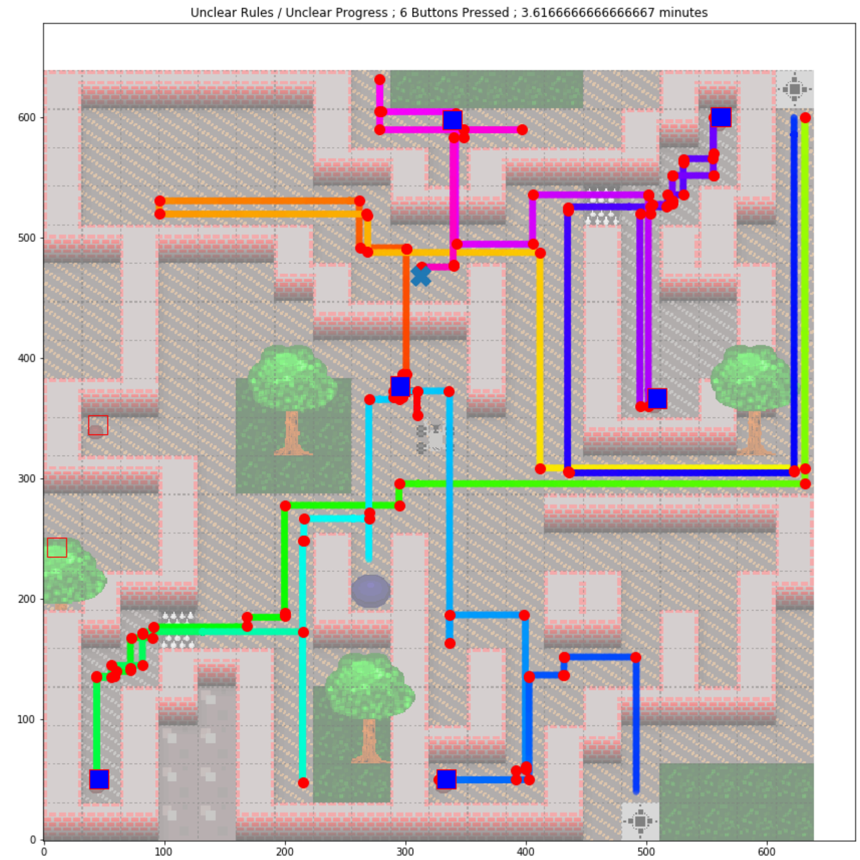

Figure 16: A detailed plot of one's activity in the game. Red squares indicate the location of red buttons, and those that the player pressed are filled in with blue. The colored lines correspond to the order in which the player moved, beginning with red and ending with violet.

Along with drawing their route directly, plots were also created showing information about each tile in the game and how long the player had spent in each tile (Figure 17). On the left, the grid shows the total amount of time was spent in each tile (in ticks of the game clock, so the number shown divided by ten is roughly the number of seconds), and on the right, the grid shows the number of times each tile was visited.

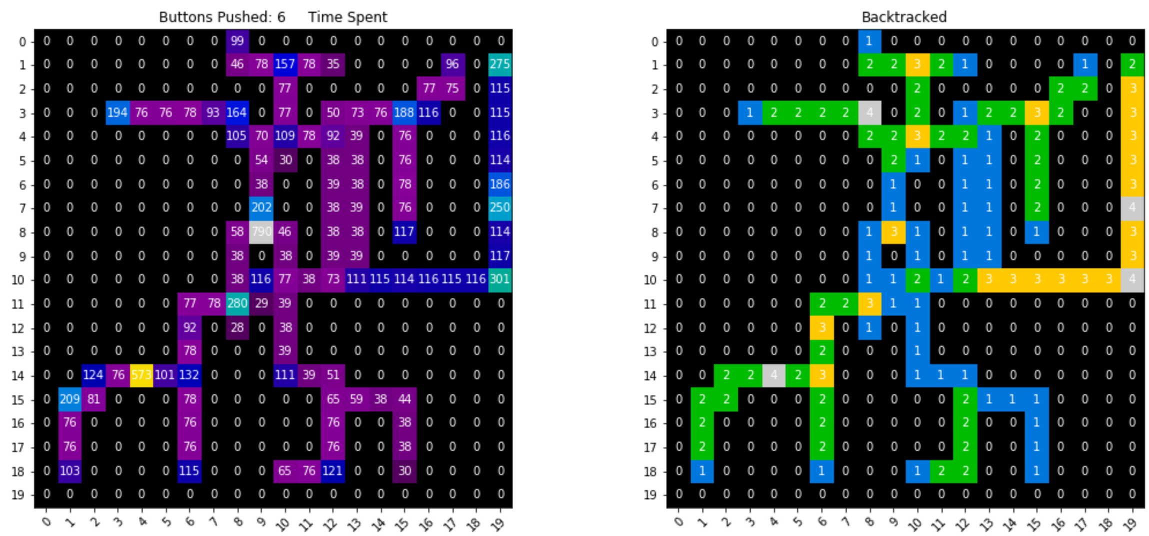

Figure 17: Plots of the game grid, showing the amount of time spend in each tile (left panel), and number of visits to each tile (right panel).

With these three plots, it possible to see where a player went, how many times an area was visited and how much time was spent there, how much of the board was covered, which buttons were found (and how many times they were pressed), and how often they stopped or changed directions.

Logging was implemented using a system that would build a buffer of the user's activity to be uploaded to the server every ten seconds.[46] Only what was determined to be useful, non-reconstructable data was logged, which consisted of the player's coordinates within the game, a timestamp provided by the game's internal clock, and a list of which keyboard keys were being held down at that moment.

---

[46] This was also a compromise between rate of recording and uploading. JavaScript doesn't have access to the file system of the client, so storing the entire log in memory before uploading it at once would slow down the game. Logging every moment directly to the server, on the other hand, created far too many server calls. This buffer-and-save system balances these two concerns.

4.4 Dependent Variables

All depended measures were derived from the logs describe above using Python scripts and then analyzed using Python and R scripts. The dependent measures used to test the impact of the four conditions on player behaviors, and to test the hypotheses described in section 4.1 are described in Table 5.

Table 5: Dependent Variable Descriptions

Dependent Variable	**Description**
Time Spent	The amount of time spent in the game from start to finish, in minutes.
Buttons Pushed	The final number of buttons ultimately found and pressed at the end of the game.
Number of Paths	The number of straight line movements between changes in direction or starts and stops.
Percent Visited	The percent of the entire map that was ultimately visited. The greater the number, the more of the map had been explored.
Blue Visits	The number of times the player went to press the blue button which ends the game. If the value is 1, that means they only went once, when they actually did finish. Anything higher than 1 indicates attempts to leave early. A 0 means they never went to the button, which should only apply if the player stayed for 10 minutes and was offered a chance to leave that way.
Average Button Revisits	Considering only the buttons that were pressed, this is the average number of times each button was pressed.

The *information-based* model predicts that players with more information should take the least amount of time, find the most buttons, take fewer paths (indicating more direct movement and confidence), visit as little of the map as necessary in order to find everything, press red buttons only one each, and only proceed to the blue button when they have completed everything. Those with the least information should take longer to finish the game, not necessarily find all the buttons, have many more paths (less direct motion), visit more of the board because they don't know where to go or not to, press the same red buttons multiple times either because they don't know if they have pressed them already, or because they aren't sure if it had an effect, and try to finish the game by going to the blue button multiple times.

The *engagement-based model* predicts that more engaged players will play the game longer, press more buttons, make less paths (more direct movements), visit most of the map, and wait longer to press the blue button. Note that most of these predictions are consistent with the information-based view except for time spent and amount of the map visited.

4.5 Results

The data was analyzed by using linear models[47] to look for significant relationships between the four conditions and player behavior as measured by the six dependent variables described in section 4.3.1. Five of the six models included a statistically significant relationship between the conditions and the behaviors, as can be seen in Table 6. The exception, Time Spent, did reveal a significant relationship to the conditions when studied more carefully. The table here is based on standardized data, but the subsequent graphs remain unstandardized for ease of interpretation.

Broadly speaking, the information-based hypothesis is confirmed by the data. Four of the six dependent measures support the model significantly. With more information, players take fewer paths, re-press red buttons less often, find more of the red buttons and don't go to the blue button as many times. The remaining two measures, Time Spent and Percent Visited, did not support the information-based model, but did support the engagement-based model by showing that players with clear progress take more time in the game and explore a larger portion of the map. Only Time Spent conformed entirely to the engagement-based theory, although Percent Visited and Number of Paths were somewhat divided between the models.

[47] It would have been natural to use ANOVA instead, since it is traditional in contexts like this where a 2x2 design is used. Linear regression models, which are formally equivalent in the case of categorical predictor variables were used instead for several reasons, including ease of interpretation of the model parameters and ease of adding covariates.

Table 6: Summary of the six main linear regression models reported in this section. Note that in every case but Time Spent that problem type was a significant predictor of behavior.

	Time Spent				Percent Visited				Number of Paths			
Dependent Measures	β	t	p		β	t	p		β	t	p	
Progress	0.02087	0.572	0.568		0.083561	2.68	0.00806	**	0.01639	0.406	0.685	
Rules	-0.00353	-0.097	0.923		-0.001973	-0.064	0.94935		-0.07989	-1.991	0.0481	*
Progress:Rules	0.03431	0.688	0.493		-0.054458	-1.278	0.20303		0.01612	0.292	0.7704	
	R^2				R^2				R^2			
	0.001815				0.03843				0.02478			

	Average Button Revisits				Buttons Pushed				Blue Visits			
Dependent Measures	β	t	p		β	t	p		β	t	p	
Progress	-0.15833	-3.081	0.002	**	0.10303	2.469	0.0145	*	-0.038068	-0.854	0.3941	
Rules	-0.15463	-3.025	0.003	**	0.06111	1.472	0.1427		-0.10625	-2.397	0.0176	*
Progress:Rules	0.14181	2.019	0.045	*	-0.05645	-0.99	0.3237		0.009222	0.151	0.8798	
	R^2				R^2				R^2			
	0.06197				0.03123				0.04974			

Table 6 summarizes the results of the six linear models which were run in R using standardized data. Each column corresponds to one of the dependent measures used to analyze player behavior, and gives the coefficients β, t values and p values to the side of the dependent measures, clear progress and clear rules. The asterisks indicate statistical significance, with a single star showing a p value of less than 0.05, and two stars meaning a p value below 0.01.

Time Spent

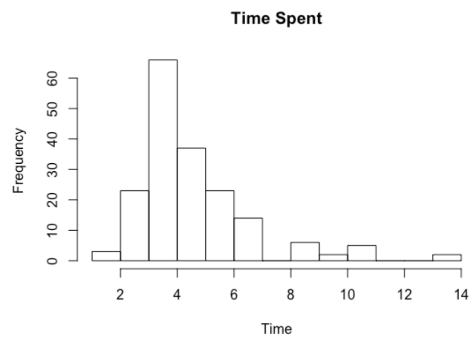

Figure 18: The length of time players spent in the game. The mean value for time spent across all participants was 4.6 minutes.

The average participant stayed in the game for 4.6 minutes (Figure 18 shows the histogram of time spent). The players that had clear progress indications tended to spend about 4.8 minutes, and those with unclear progress only stayed for about 4.3 minutes, suggesting that awareness of progress motivated players to remain longer.

The linear model for time spent did not show any significant relationships between the amount of time players spend in the game and whether they had been given clear rules or progress, as can be seen in Figure 19. Progress and rules did not significantly predict how much time players spent in the game. According to the information-based model, it was expected that players with more information available to them would take less time to complete the game, as they would know how many buttons they needed to press and/or how many they had found, making it easier for them to know when they were done.

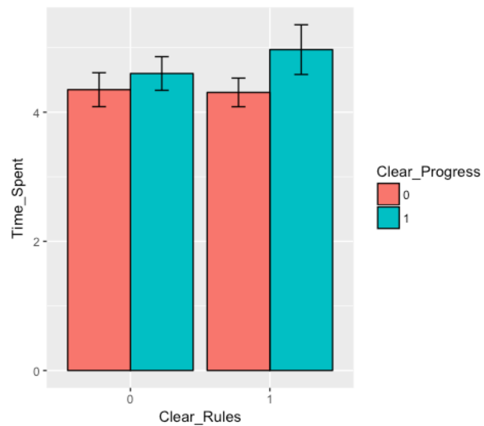

Figure 19: Time Spent for all subjects. The left two bars represent unclear rules, while the right two bars are the clear rules conditions. Unclear progress is shown in red, and clear progress is in teal. This graph shows all four experimental conditions, from left to right: Uninformed, Progress Only, Rules Only, Fully Informed. This same style is used for the following plots.

However, when plotted, it does visually appear as though progress had an impact on how long a player spent. Clear progress should lead to less time spent in the game under the information-based hypothesis, but the engagement-based account predicts that a progress indicator leads to players feeling more interested in the game, and spending a longer time as a result. Since this effect looked like it might be present in the data, a closer look seemed warranted.

The results reported in Table 6 and shown above correspond to all the participants, regardless of performance otherwise. Other problem-solving studies tend to report how long it takes for subjects to complete a task, rather than how long they decide to try, so with that in mind, data was restricted to only those players who actually succeeded in finding all eight red buttons. When considering only those players who actually completed the game by finding all eight buttons, a main effect for time spent does appear. Within this subset of the data, clear progress was found to be a significant predictor of time spent $\beta = 0.83563$, $t(89) = 0.0465$, $p < 0.05$.[48] People spent more time in the game when they had a progress indicator. Figure 20 shows this relationship.

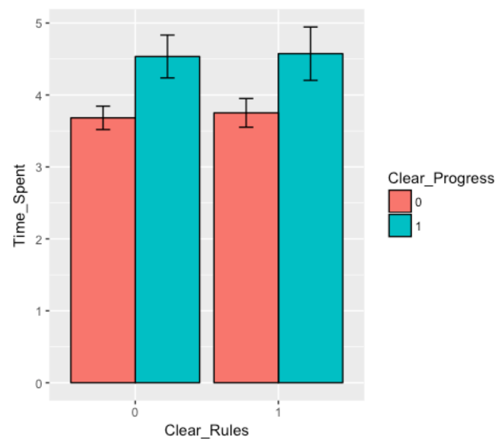

Figure 20: Time Spent for all four conditions using only the subjects who found all eight red buttons. Clear progress was a significant predictor of time spent, with subjects spending longer periods of time in the game when given clear indications of progress.

[48] The other linear models were also run on this restricted subset of subjects who found all eight buttons. *Buttons pressed* was excluded for obvious reasons. In three of the four cases the effect stayed the same. In the case of percent visited, the results were interesting and are discussed in that section. It should be noted that the effects found for the case of Time Spent and Percent Visited, when run on the restricted subset of the data, did not persist when the interaction term was included. This appears to be because the interaction term draws variance away from the main effects in both cases, but especially in the case of Figure 20 the effect seems real. Further study is obviously needed before any firm conclusions can be drawn from this data.

Percent Visited

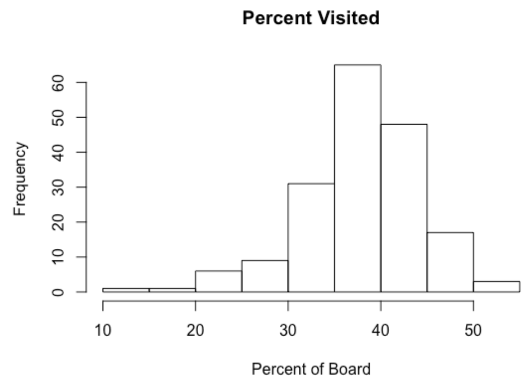

Figure 21: A histogram of the total percent of the game board visited by players. The mean value for board coverage was 37.9%.

Players covered 38% of the game board on average (see Figure 17 and Figure 21). Those with clear progress averaged 39% while those with unclear progress tended to visit closer to 36% of the map.

In this game the board is literally a search space. If the intent of the subject is to finish quickly and efficiently, as the information-based model predicts, then it should be the case that they would only explore as much of the map as necessary. Hence, the greater the amount of information available, the less of the board should ultimately be covered. The reverse should also be true; the less you know, the more you'll explore. If you know that you need to find *all* of the buttons, but not how many there are, you are more likely to wander into every corner and scour the map to be sure you didn't miss one.[49]

[49] It may be harder to apply the idea of minimal search in an environment like this, since the buttons are scattered throughout the board, and consequently a certain amount of coverage is required. Also, the small

In contrast, the engagement hypothesis predicts that players will go anywhere that looks interesting, exploring as much as they can, as long as they feel drawn in.[50] Feeling more engaged on this account, should result in a higher motivation to be thorough, leading to more of the board being visited, and more time being taken by the player.

The data appears to agree with the engagement-based model for how much of the map is visited, with a main effect for progress of $\beta = 0.083561$, $t(177) = 2.68$, $p < 0.01$ (see Figure 22). Players that were given a clear sense of progress did more exploring than those without, ultimately covering a greater percentage of the map.

Somewhat contrary to both theories, however, rules did not produce an effect, with $\beta = -0.001973$, $t(177) = -0.064$, $p = 0.94935$, nor did the interaction with $\beta = -0.054458$, $t(177) = -1.278$, $p = 0.20303$. As can be seen in the graph, subjects who lacked clear progress visited approximately the same amount of the map on average, regardless of whether they had been given clear rules or not.[51] Those who knew how many buttons they had found, but not how many they had left to find, tended to cover the greatest percentage of the board.

field of view limits a player's ability to plan their route ahead of time. However, it would still make sense that anyone knowing that a button isn't invisible would turn away as soon as they see walls on all sides.

[50] This actually occurs quite frequently in video games, where players often try to follow every route they can find, just in case there is something tucked away that could aid them or boost their score. Game designers have, accordingly, been putting prizes in remote areas to reward players for such exploration. These are not usually critical to the game, but rather extras that are included for players who are more engaged.

[51] The differences show that without clear progress, players only visited about .2% more of the board when they had the rules. By contrast, the difference in coverage was about 5% among those with clear progress.

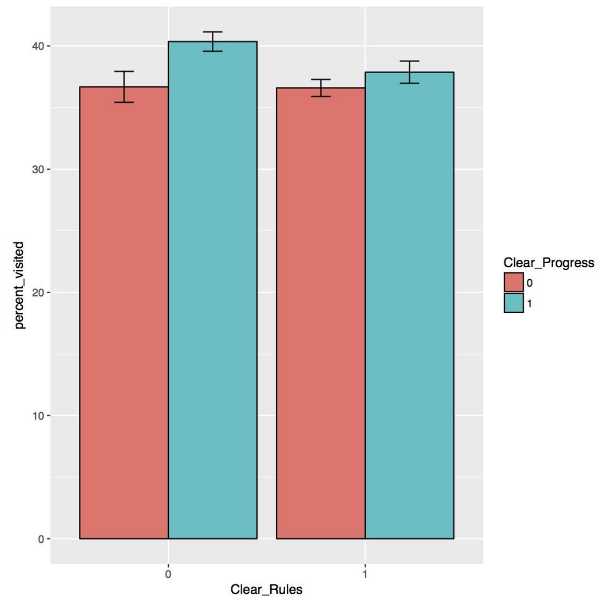

Figure 22: Percent of the game board visited by participants. Clear progress was a significant predictor of coverage, with subjects given clear progress indicators tending to visit more of the map.

As was done for the Time Spent, an additional analysis was performed for the percent of the board visited by only those subjects who succeeded in finding all eight red buttons (Figure 23). In this subset the opposite effect is found, with rules as the significant predictor of coverage $\beta = -2.866$, $t(177) = -2.734$, $p < 0.01$, while progress does not significantly predict how much of the board is visited $\beta = 1.094$, $t(0) = 1.04$, $p = 0.30106$. This suggests that for those who complete the task of finding all the red buttons, the availability of clear rules reduced the percent of the map covered, as the information-based hypothesis predicted. It is interesting to note that knowledge of the game's rules led those who were most successful to visit less of the board, while at the same time the availability of progress indication generally resulted in a higher percent of coverage for participants. Perhaps in this instance there is some friction between the two dimensions that is less evident elsewhere. Further study of this would be needed in order to tell more.

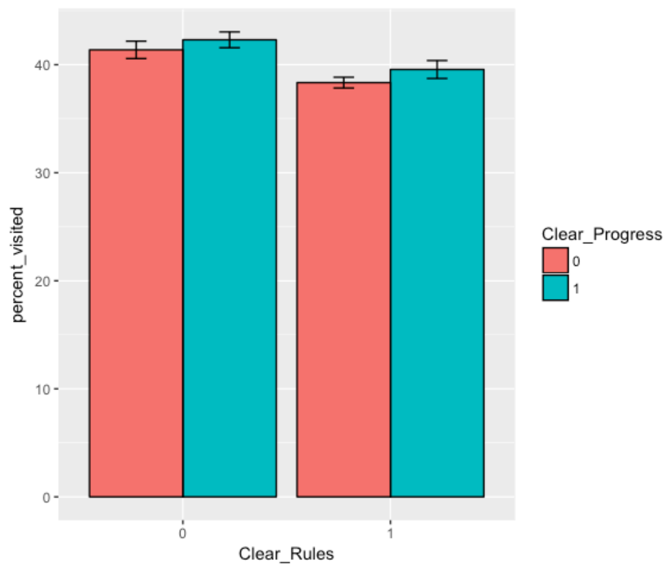

Figure 23: Percent of the game board visited by subjects who found all red buttons. For this subset, the presence of clear rules was a significant predictor, resulting in less of the board getting covered when the rules of the game were known.

Number of Paths

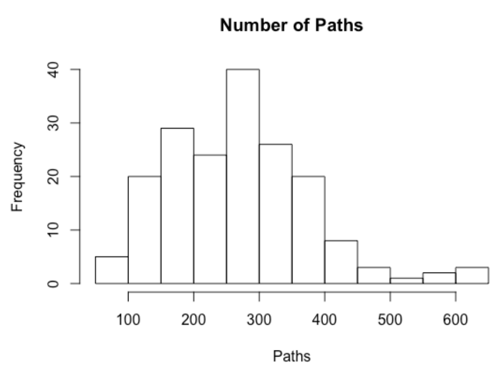

Figure 24: Histogram of the number of paths taken by players. The mean number of paths taken by subjects was 271.

A *path* is defined as the number of straight line movements between changes in direction or starts and stops (see Figure 25). This means that any time a player stopped moving, or changed directions, a new path was formed. The average player took 271 paths throughout the course of the game (see Figure 24), but those with clear rules tended to produce fewer paths than those with unclear rules did.

Figure 25: Any time a player changed direction or stopped and started again, it was used to demarcate the end of a path. In this way it is possible to see how deliberately the player moves, whether via long straight paths or shorter more hesitant steps (Compare Figure 16).

Paths were used to indicate how often the player changed direction, or started and stopped moving. The information-based approach suggests that players who know the rules of the game (specifically that walls open on button presses, that there is no danger, and that buttons might be hidden), should, on this account, be more confident as they travel the board, moving in broad strokes and producing fewer paths. A lack of such information should result in a more cautious and hesitant style, consisting of shorter paths with more turns as the player tries to pay more attention to their environment.[52] Köhler's

[52] Imagine having to walk across a room in the dark. Someone who is familiar with the room's layout and content would likely move much more swiftly to the other side than someone with no idea what might be in their way, proceeding carefully so they don't step on anything, etc. Here again, more information is associated with fewer paths.

studies with dogs and chimps lend credence to this view by demonstrating that animals wander aimlessly around when they are unable to obtain food, but transition to direct, fast, movement after overcoming an impasse and figuring out how to reach it (Köhler, 1924).

The effect of clear progress here seems like it would be less impactful than rules, and the data support this. Progress was not a significant predictor of the number of paths $\beta = 0.01639$, $t(177) = 0.406$, $p = 0.6850$ (see Figure 26), but having clear rules did significantly predict fewer paths being taken $\beta = -0.07989$, $t(177) = -1.991$, $p < 0.05$. On average, players produced 293 paths when they didn't have clear rules, and 251 when they did. This suggests that players who knew more about the game were able to take more direct routes with fewer stops than those who were uncertain what lay ahead, or how much they needed to do.

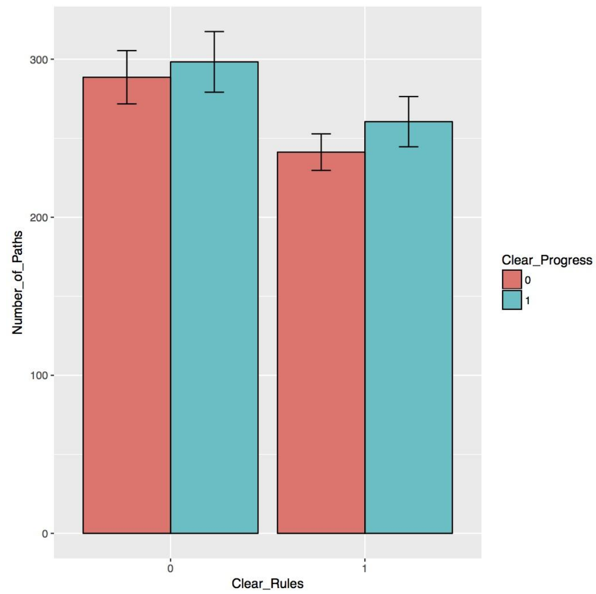

Figure 26: The number of paths taken by players in the course of playing the game. Clear rules was a significant predictor of paths, resulting in fewer paths taken when rules were known.

Average Button Revisits

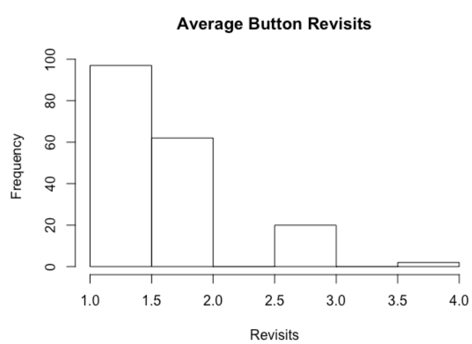

Figure 27: Histogram of players pressing buttons multiple times. The mean for the number of times each red button was pressed was 1.6.

Although pressing a red button more than once did not have an effect in the game, many players returned to the same buttons on multiple occasions anyway. On average, each button was pressed 1.6 times, but this number decreased when players had either clear rules or clear progress to guide them. The average for both clear progress and for clear rules was about 1.5 times, while unclear for each was about 1.7 (see Figure 27).

The information-based model predicts that less information will result in players revisiting the same buttons more frequently. Lack of information about which buttons have been pressed should make it likely that subjects will press a button several times, either because they don't remember whether they have pressed it already, or because they want to be sure it got counted. Knowing the rules also seems like it should make a difference to the number of times the buttons are pressed. A player who doesn't know the rules about the buttons may guess at some of the rules, speculating that buttons might need to be pressed in sequence, for example, to open a wall, or that buttons should be pressed repeatedly in order to take effect.

In cases of clear progress (either Fully Informed or Progress Only), buttons become translucent once pressed, and a number at the top of the screen increases. This makes it very clear that a button has been pressed and that re-pressing it has no effect, so subjects in these cases should not be revisiting buttons they have already been to.

Table 6 shows significant main effects for both progress and rules. However, the main effects are illusory. The effect is clearly driven by the interaction between progress and rules, as can be seen in Figure 28. Having both the rules and the progress indicators together strongly predicts the number of times buttons would be re-pressed. Specifically, a lack of both rules and progress leaves players with too little to go on that would keep them from pressing the same buttons repeatedly. Either rules or progress would give them a reason not to revisit buttons, but when both are missing, the effect is clear.

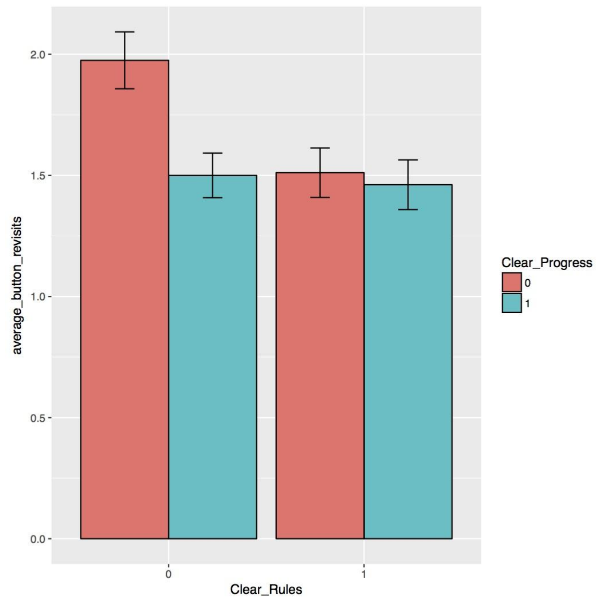

Figure 28: Average button revisits by subjects yields main effects as well as an interaction effect from having both unclear rules and unclear progress. Having both rules and progress indicators leads to people revisiting the buttons significantly more times than in any other condition.

Buttons Pushed

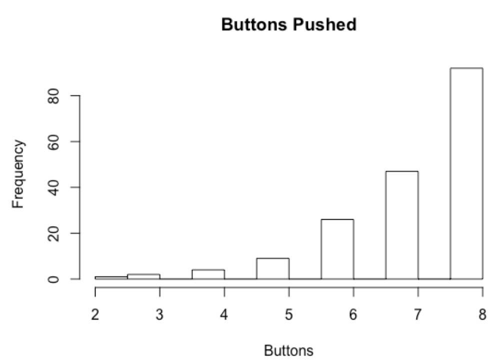

Figure 29: Histogram of how many buttons players were able to find and press. Mean number of buttons pressed was 7.1.

Since the task given to all participants was to find and press all red buttons on the map, the final count of how many they reached works as a performance score. The information-based model holds that the more information a player has, the more buttons they will find, and the less information, the fewer the number of buttons pressed. On average, players found 7.1 of the 8 buttons, though without a progress indicator this number dips to 6.8 (Figure 29).

The interaction-based hypothesis expects main effects for both rules and progress. If a player knows how many buttons they have found (clear progress), they should try to find as many as they can, rather than hitting the same ones multiple times. Similarly, if the number of buttons that can be found is known, players should keep playing until they have found that number (to the best of their knowledge).

The engagement-based hypothesis suggests that players will stop finding buttons when they're no longer motivated to do so, and will continue as long as they are. A counter may work as a motivational element that makes players want to increase the number, but without it, they may not be engaged enough to bother finding everything.

The resulting data did find a main effect for progress, which supports both the information and engagement theories.[53] In other words, players given a way to tell how much progress they were making tended to find more of the red buttons than those without such indicators.

---

[53] It may be worth noting that buttons pushed is significantly related to how taxing and frustrating the process was reported to be, and that this has some interaction with the main effect (post game survey data are presented in Appendix A). When including reported frustration as a covariate, progress becomes more significant, with $\beta = 0.47528$, $t(177) = 2.849$, $p < 0.01$. When restricting attention to participants who did not find the task taxing, the relation to the progress bar also became significant with $\beta = 0.21917$, $t(47) = 3.323$, $p < 0.01$.

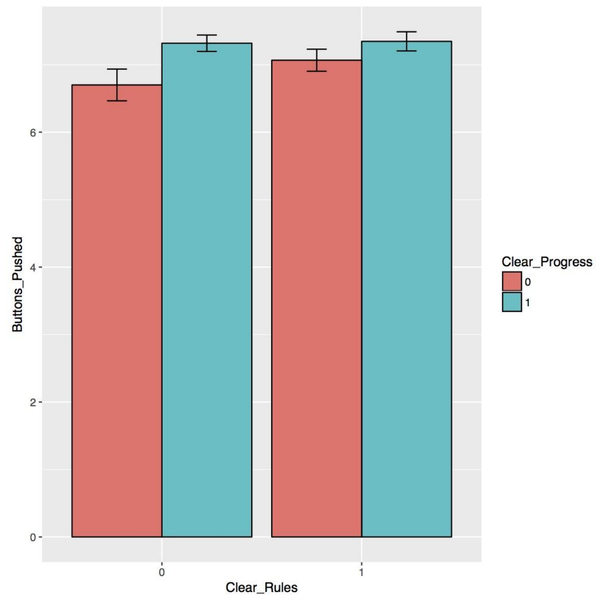

Figure 30: The total number of red buttons found by participants. Clear progress was a significant predictor of the number of buttons pressed, with the presence of the progress indicator resulting in a higher number of button presses than without.

Blue Visits

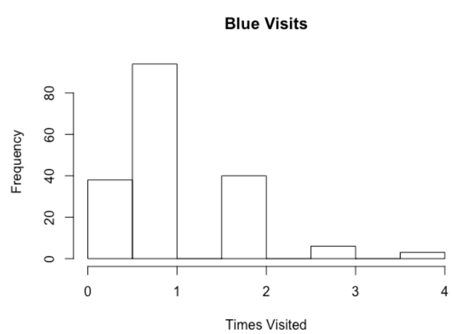

Figure 31: Histogram of the number of times players tried to end the game. Mean number of visits to the blue button was 1.12.

Players visited the blue button to end the game an average of 1.12 times. Those who had been given clear rules, however, only visited it 0.9 times while those with unclear rules pressed the blue button 1.3 times on average (Figure 31).

The blue button was used as a way for players to who didn't know how many buttons there were to check whether they had found everything. This is related to the kinds of unnecessary actions described in section 4.1, since pressing the blue button is not necessary to the player until they have completed the game by finding every red button. As a way to keep players from exiting too early, the blue button would send players back to the game unless they had found all eight buttons, or spent at least three minutes trying.

The information-based hypothesis seems quite straightforward for this measure; the more information you have, the fewer times you'll try to end the game, and vice versa. Players that know how many red buttons there are to find should avoid visiting the blue button until they believe they have found them all, resulting in minimal attempts to end the game prematurely. Those without a sense of how many buttons they have yet to find (by not knowing how many there are), are more likely to press the blue button in an attempt to see if they're done. If told that there is more to do, they may wander around until they find another red button, then return to the blue button to check if there are any more.

The data confirmed the hypothesis that those who had been told the rules of the game (specifically the number of red buttons they needed find), ended up pressing the blue button fewer times, as expected.

One way to interpret this result is that people without clear rules will attempt to find their own with what's available, such as using the exit button as a means to check for completion. The engagement-based hypothesis suggested that people would try to leave when they stopped feeling interested enough to continue, which would also make sense in the condition where there is minimal knowledge.

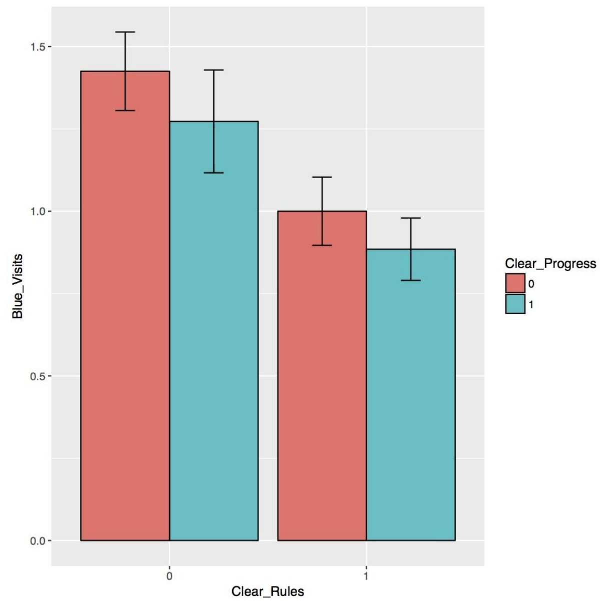

Figure 32: The number of times players tried to end the game by pressing the blue button. Clear rules was a significant indicator of blue button visits, with those subjects having been given a clear description of the game's rules pressing the blue button fewer times than those with unclear rules.

Chapter 5: Conclusion

This dissertation has presented a model of the *problem cycle* that places problem solving in the context of a larger process which includes the discovery and framing of problems. Working from the existing literature on problem solving, it has been shown that not all problems are equivalent in terms of what makes them difficult and what is necessary in order to solve them. Since most problems that are encountered are not immediately well-defined and neatly presented, it has been argued that problem finding is a critical component of the problem cycle, and one that bears greater investigation.

To further support the idea that problem finding is an important part of (and precursor to) problem solving, the recent history of study in this area was examined, laying out the observed phenomena involved with solving problems and the efforts to clarify whether and in what way problems might be distinguished from one another by type. Using the literature on well and ill-defined problems as a starting point, a number of distinct problems were considered and each was laid out in terms of what each supplied in terms of structure, what was required for a solution, and what made each problem hard. Common features among these aspects were identified and a table of four components was constructed that yielded sixteen types of problems.

Finally, to test this model of different problem types, an experiment was conducted that began with a single source problem that was then framed in four different ways, corresponding to four different problem types. The results of this study show that behavior in problem solving does indeed rely to a significant extent on the type of problem involved, as demonstrated by changes in performance when the problem type is altered.

There are clearly ways this study could be improved in follow-up work. The game could be made simpler by removing extraneous features like the spikes and teleporter. Alternatively, these could be made more relevant by making them necessary for completion, and having that information selectively presented in the instructions. Implementation issues with the data logging also need to be addressed. More generally the study as a whole was exploratory, and based on the experience of this study, a future version could be more explicit in expectation. Multiple dependent variables and subsets of the data were considered, increasing the likelihood of type 1 errors. A natural next step would be to run an improved version of the experiment using explicitly pre-registered hypotheses, and on more subjects to increase statistical power.

There is clearly more work to be done in the area of problem definition and classification, and I am interested in continuing to research the nature of problems, in particular problem finding. This research has provided evidence that problem finding exists and makes a difference to how problems are solved, but deeper questions about how problem finding truly works have not yet been investigated. In the experiment reported in Chapter 4, problem determination was done on behalf of subjects by giving them different information to work with. A possible next step would be to present an ambiguous problem that could be framed in multiple ways to subjects. In that case subjects could carry out the problem determination on their own, using observation of what they attend to (as in Spivey & Dale, 2011), or behavioral differences like those in the present study, to identify which interpretation gets chosen. Ultimately I'd like to be able to develop a computer simulation of problem solving, and to use this simulation to formulate and test more precise hypotheses about how problem finding works in humans.

Appendix A: Survey Questions

The following survey was given to all participants of the experiment after completing the game, and before they were given a code to enter into Mechanical Turk for credit. The first 13 questions were phrased as statements with a Likert scale used to indicate how well the subject agreed. The possible answers were:

Strongly Disagree, Disagree, Neutral, Agree, Strongly Agree

These were the questions asked. In italics before each question is the shorter word we used to reference the answers in our analyses, which was not shown to subjects.

1. *Confident* I was confident that I completed the task.
2. *Remember* During the course of the game it was easy to remember where I had already been.
3. *Last* It was not very difficult to find the last button that I pressed.
4. *Smooth* I found that the game ran smoothly.
5. *Wander* I found myself wandering aimlessly through the game.
6. *Surprised* I was rarely surprised during the game.
7. *Practice* I think that with practice, I could become very proficient with this game.
8. *Drawn* I was really drawn into my gaming task.
9. *Taxing* Playing the game was mentally taxing.
10. *Frustrated* I felt frustrated while playing the game.

11. *Worthwhile* Playing the game was worthwhile.

12. *Rewarding* My game experience was rewarding.

13. *Fun* This gaming experience was fun.

We also asked a few demographic questions, along with a pair of catch questions to make sure subjects were paying attention. These were

14. How old are you?
15. What is your gender?
16. Is English your first language?
17. How many letters are in the English alphabet?
18. Growing up, how many hours per week did you play video games?
19. How many hours per week do you play video games now?
20. Select the number 17

Adamson, R. E. (1952). Functional fixedness as related to problem solving: A repetition of three experiments. *Journal of Experimental Psychology*, *44*(4), 288–291.

Amarel, S. (1981). On representations of problems of reasoning about actions. In *Readings in artificial intelligence* (pp. 2–22). Elsevier.

Arlin, P. K. (1975). A cognitive process model of problem finding. *Educational Horizons*, 99–106.

Birch, H. G., & Rabinowitz, H. S. (1951). The negative effect of previous experience on productive thinking. *Journal of Experimental Psychology*, *41*(2), 121.

Davey, R. (2018). *phaser: Phaser is a fun, free and fast 2D game framework for making HTML5 games for desktop and mobile web browsers, supporting Canvas and WebGL rendering*. JavaScript. Retrieved from https://github.com/photonstorm/phaser (Original work published 2013)

Davidson, J. E., & Sternberg, R. J. (2003). *The psychology of problem solving*. Cambridge University Press.

De Groot, A. D. (1946). *Thought and choice in chess*. Ishi Press International.

Dieciuc, M., Roque, N., & Folstein, J. (2016). Category learning causes a stable advantage for category-relevant shape dimensions during a task requiring attention to all dimensions: ERP evidence. *Journal of Vision*, *16*(12), 257–257.

Duncker, K. (1945). On problem-solving. *Psychological Monographs*, *58*(5), i.

Durso, F. T., Rea, C. B., & Dayton, T. (1994). Graph-theoretic confirmation of restructuring during insight. *Psychological Science*, *5*(2), 94–98.

Gardner, H. (1990). *Varieties of Thinking: Essays From Harvard's Philosophy of Education Research Center*. (V. A. Howard, Ed.) (1 edition). New York: Routledge.

Gershman, S. J., Horvitz, E. J., & Tenenbaum, J. B. (2015). Computational rationality: A converging paradigm for intelligence in brains, minds, and machines. *Science*, *349*(6245), 273–278. https://doi.org/10.1126/science.aac6076

Gick, M. L., & Holyoak, K. J. (1980). Analogical problem solving. *Cognitive Psychology*, *12*(3), 306–355. https://doi.org/10.1016/0010-0285(80)90013-4

Gilhooly, K. J. (2012). *Human and machine problem solving*. Springer Science & Business Media.

Goel, V., & Pirolli, P. (1992). The structure of design problem spaces. *Cognitive Science*, *16*(3), 395–429.

Grant, E. R., & Spivey, M. J. (2003). Eye movements and problem solving: Guiding attention guides thought. *Psychological Science*, *14*(5), 462–466.

Greeno, J. G. (1974). Hobbits and orcs: Acquisition of a sequential concept. *Cognitive Psychology*, *6*(2), 270–292.

Greeno, J. G., & Simon, H. A. (1988). *Problem Solving and Reasoning*. DTIC Document.

Guidone, M. (2018). *Interpretation of Rope Problem* [Microsoft Paint].

Jeff Lait. (n.d.). Powder (Version Build 111) [Game Boy Advance].

Jonassen, D. H. (2000). Toward a design theory of problem solving. *Educational Technology Research and Development*, *48*(4), 63–85.

Kaufmann, G. (1990). Imagery effects on problem solving.

Kershaw, T. C., & Ohlsson, S. (2004). Multiple causes of difficulty in insight: the case of the nine-dot problem. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *30*(1), 3.

Kohler, W. (1924). The mentality of apes. Retrieved from http://doi.apa.org/?uid=1924-15026-000

Langley, P. (2000). The computational support of scientific discovery. *International Journal of Human-Computer Studies*, *53*(3), 393–410.

Lindeijer, T. (2018). *tiled: A flexible level editor*. C++. Retrieved from https://github.com/bjorn/tiled (Original work published 2011)

Lovett, M. C., & Anderson, J. R. (1996). History of success and current context in problem solving: Combined influences on operator selection. *Cognitive Psychology*, *31*(2), 168–217.

Luchins, A. S. (1942). Mechanization in problem solving: the effect of Einstellung. *Psychological Monographs*, *54*(6), i.

Mackworth, N. H. (1965). Originality. *American Psychologist*, *20*(1), 51.

Maier, N. R. F. (1931). Reasoning in humans. II. The solution of a problem and its appearance in consciousness. *Journal of Comparative Psychology*, *12*(2), 181–194. http://dx.doi.org/10.1037/h0071361

Metcalfe, J., & Wiebe, D. (1987). Intuition in insight and noninsight problem solving. *Memory & Cognition*, *15*(3), 238–246.

Middleton, H. E. (1998). The Role of Visual Mental Imagery in Solving Complex Problems in Design.

Newell, A. (1980). Duncker on thinking: An inquiry into progress in cognition.

Newell, A., & Simon, H. A. (1972). *Human problem solving* (Vol. 104). Prentice-Hall Englewood Cliffs, NJ.

Novick, L. R., & Bassok, M. (2005). Problem solving.

Novick, L. R., & Sherman, S. J. (2003). On the nature of insight solutions: Evidence from skill differences in anagram solution. *The Quarterly Journal of Experimental Psychology: Section A*, *56*(2), 351–382.

Ohlsson, S. (1984). Restructuring revisited. *Scandinavian Journal of Psychology*, *25*(1), 65–78.

Pedone, R., Hummel, J. E., & Holyoak, K. J. (2001). The use of diagrams in analogical problem solving. *Memory & Cognition*, *29*(2), 214–221.

Reber, R., Ruch-Monachon, M.-A., & Perrig, W. J. (2007). Decomposing intuitive components in a conceptual problem solving task. *Consciousness and Cognition*, *16*(2), 294–309. https://doi.org/10.1016/j.concog.2006.05.004

Sharek, D., & Wiebe, E. (2015). Investigating real-time predictors of engagement: Implications for adaptive videogames and online training. *International Journal of Gaming and Computer-Mediated Simulations (IJGCMS)*, *7*(1), 20–37.

Spivey, M. J., & Dale, R. (2011). Eye movements both reveal and influence problem solving. *Eye*, *2*, 1.

Stephen, D. G., Dixon, J. A., & Isenhower, R. W. (2009). Dynamics of representational change: entropy, action, and cognition. *Journal of Experimental Psychology: Human Perception and Performance*, *35*(6), 1811.

Thomas, L. E., & Lleras, A. (2007). Moving eyes and moving thought: On the spatial compatibility between eye movements and cognition. *Psychonomic Bulletin & Review*, *14*(4), 663–668.

VanLehn, K. (1989). Problem solving and cognitive skill acquisition. In *Foundations of cognitive science* (pp. 527–579). Cambridge, MA, US: The MIT Press.

Weisberg, R. W., & Alba, J. W. (1981). An examination of the alleged role of" fixation" in the solution of several" insight" problems. *Journal of Experimental Psychology: General*, *110*(2), 169.

Wertheimer, Max. (1959). *Productive thinking*. (Michael Wertheimer, Ed.). Harper New York.

Wiebe, E. N., Lamb, A., Hardy, M., & Sharek, D. (2014). Measuring engagement in video game-based environments: Investigation of the User Engagement Scale. *Computers in Human Behavior*, *32*, 123–132. https://doi.org/10.1016/j.chb.2013.12.001

Zhang, X., Chen, X., Wang, W., Yang, J., Lantz, V., & Wang, K. (2009). Hand gesture recognition and virtual game control based on 3D accelerometer and EMG sensors. In *Proceedings of the 14th international conference on Intelligent user interfaces* (pp. 401–406). ACM.