Insight Problem Solving: A Critical Examination of the Possibility of Formal Theory

William H. Batchelder[1] and Gregory E. Alexander[1]

Abstract

This paper provides a critical examination of the current state and future possibility of formal cognitive theory for insight problem solving and its associated "aha!" experience. Insight problems are contrasted with move problems, which have been formally defined and studied extensively by cognitive psychologists since the pioneering work of Alan Newell and Herbert Simon. To facilitate our discussion, a number of classical brainteasers are presented along with their solutions and some conclusions derived from observing the behavior of many students trying to solve them. Some of these problems are interesting in their own right, and many of them have not been discussed before in the psychological literature. The main purpose of presenting the brainteasers is to assist in discussing the status of formal cognitive theory for insight problem solving, which is argued to be considerably weaker than that found in other areas of higher cognition such as human memory, decision-making, categorization, and perception. We discuss theoretical barriers that have plagued the development of successful formal theory for insight problem solving. A few suggestions are made that might serve to advance the field.

Keywords

Insight problems, move problems, modularity, problem representation

---

[1] Department of Cognitive Sciences, University of California Irvine

1. Introduction

This paper discusses the current state and a possible future of formal cognitive theory for insight problem solving and its associated "aha!" experience. Insight problems are contrasted with so-called move problems defined and studied extensively by Alan Newell and Herbert Simon (1972). These authors provided a formal, computational theory for such problems called the General Problem Solver (GPS), and this theory was one of the first formal information processing theories to be developed in cognitive psychology. A move problem is posed to solvers in terms of a clearly defined representation consisting of a starting state, a description of the goal state(s), and operators that allow transitions from one problem state to another, as in Newell and Simon (1972) and Mayer (1992). A solution to a move problem involves applying operators successively to generate a sequence of transitions (moves) from the starting state through intermediate problem states and finally to a goal state. Move problems will be discussed more extensively in Section 4.6.

In solving move problems, insight may be required for selecting productive moves at various states in the problem space; however, for our purposes we are interested in the sorts of problems that are described often as insight problems. Unlike Newell and Simon's formal definition of move problems, there has not been a generally agreed upon definition of an insight problem (Ash, Jee, and Wiley, 2012; Chronicle, MacGregor, and Ormerod, 2004; Chu and MacGregor, 2011). It is our view that it is not productive to attempt a precise logical definition of an insight problem, and instead we offer a set of shared defining characteristics in the spirit of Wittgenstein's (1958) definition of 'game' in terms of family resemblances. Problems that we will treat as insight problems share many of the following defining characteristics: (1) They are posed in such a way as to admit several possible problem representations, each with an associated solution search space. (2) Likely initial representations are inadequate in that they fail to allow the possibility of discovering a problem solution. (3) In order to overcome such a failure, it is necessary to find an alternative productive representation of the problem. (4) Finding a productive problem representation may be facilitated by a period of non-solving activity called incubation, and also it may be potentiated by well-chosen hints. (5) Once obtained, a productive representation leads quite directly and quickly to a solution. (6) The solution involves the use of knowledge that is well known to the solver. (7) Once the solution is obtained, it is accompanied by a so-called "aha!" experience. (8) When a solution is revealed to a non-solver, it is grasped quickly, often with a feeling of surprise at its simplicity, akin to an "aha!" experience.

It is our position that very little is known empirically or theoretically about the cognitive processes involved in solving insight problems. Furthermore, this lack of knowledge stands in stark contrast with other areas of cognition such as human memory, decision-making, categorization, and perception. These areas of cognition have a large number of replicable empirical facts, and many formal theories and computational models exist that attempt to explain these facts in terms of underlying cognitive processes. The main goal

of this paper is to explain the reasons why it has been so difficult to achieve a scientific understanding of the cognitive processes involved in insight problem solving.

There have been many scientific books and papers on insight problem solving, starting with the seminal work of the Gestalt psychologists Köhler (1925), Duncker (1945), and Wertheimer (1954), as well as the English social psychologist, Wallas (1926). Since the contributions of the early Gestalt psychologists, there have been many journal articles, a few scientific books, such as those by Sternberg and Davidson (1996) and Chu (2009), and a large number of books on the subject by laypersons. Most recently, two excellent critical reviews of insight problem solving have appeared: Ash, Cushen, and Wiley (2009) and Chu and MacGregor (2011).

The approach in this paper is to discuss, at a general level, the nature of several fundamental barriers to the scientific study of insight problem solving. Rather than criticizing particular experimental studies or specific theories in detail, we try to step back and take a look at the area itself. In this effort, we attempt to identify principled reasons why the area of insight problem solving is so resistant to scientific progress. To assist in this approach we discuss and informally analyze eighteen classical brainteasers in the main sections of the paper. These problems are among many that have been posed to hundreds of upper divisional undergraduate students in a course titled "Human Problem Solving" taught for many years by the senior author. Only the first two of these problems can be regarded strictly as move problems in the sense of Newell and Simon, and most of the rest share many of the characteristics of insight problems as described earlier.

The paper is divided into five main sections. After the Introduction, Section 2 describes the nature of the problem solving class. Section 3 poses the eighteen brainteasers that will be discussed in later sections of the paper. The reader is invited to try to solve these problems before checking out the solutions in the Appendix. Section 4 lays out six major barriers to developing a deep scientific theory of insight problem solving that we believe are endemic to the field. We argue that these barriers are not present in other, more theoretically advanced areas of higher cognition such as human memory, decision-making, categorization, and perception. These barriers include the lack of many experimental paradigms (4.1), the lack of a large, well-classified set of stimulus material (4.2), and the lack of many informative behavioral measures (4.3). In addition, it is argued that insight problem solving is difficult to study because it is non-modular, both in the sense of Fodor (1983) but more importantly in several weaker senses of modularity that admit other areas of higher cognition (4.4), the lack of theoretical generalizations about insight problem solving from experiments with particular insight problems (4.5), and the lack of computational theories of human insight (4.6). Finally, in Section 5, we suggest several avenues that may help overcome some of the barriers described in Section 4. These include suggestions for useful classes of insight problems (5.1), suggestions for experimental work with expert problem solvers (5.2), and some possibilities for a computational theory of insight.

2. Batchelder's Human Problem Solving Class

The senior author, William Batchelder, has taught an Upper Divisional Undergraduate course called 'Human Problem Solving" for over twenty-five years to classes ranging in size from 75 to 100 students. By way of background, his active research is in other areas of the cognitive sciences; however, he maintains a long-term hobby of studying classical brainteasers. In the area of complex games, he achieved the title of Senior Master from the United States Chess Federation, he was an active duplicate bridge player throughout undergraduate and graduate school, and he also achieved a reasonable level of skill in the game of Go.

The content of the problem-solving course is split into two main topics. The first topic involves encouraging students to try their hand at solving a number of famous brainteasers drawn from the sizeable folklore of insight problems, especially the work of Martin Gardner (1978, 1982), Sam Loyd (1914), and Raymond Smullyan (1978). In addition, games like chess, bridge, and Go are discussed. The second topic involves presenting the psychological theory of thinking and problem solving, and in most cases the material is organized around developments in topics that are covered in the first eight chapters of Mayer (1992). These topics include work of the Gestalt psychologists on problem solving, discussion of experiments and theories concerning induction and deduction, presenting the work on move problems, including the General Problem Solver (Newell & Simon, 1972), showing how response time studies can reveal mental architectures, and describing theories of memory representation and question answering.

Despite efforts, the structure of the course does not reflect a close overlap between its two main topics. The principal reason for this is that in our view the level of theoretical and empirical work on insight problem solving is at a substantially lower level than is the work in almost any other area of cognition dealing with higher processes. The main goal of this paper is to explain our reasons for this pessimistic view. To assist in this goal, it is helpful to get some classical brainteasers on the table. While most of these problems have not been used in experimental studies, the senior author has experienced the solution efforts and post solution discussions of over 2,000 students who have grappled with these problems in class.

3. Some Classic Brainteasers

In this section we present eighteen classical brainteasers from the folklore of problem solving that will be discussed in the remainder of the paper. These problems have delighted brainteaser connoisseurs for years, and most are capable of giving the solver a large dose of the "aha!" experience. There are numerous collections of these problems in books, and many collections of them are accessible through the Internet. We have selected these problems because they, and others like them, pose a real challenge to any effort to

develop a deep and general formal theory of human or machine insight problem solving. With the exception of Problems 3.1 and 3.2, and arguably 3.6, the problems are different in important respects from so-called move problems of Newell and Simon (1972) described earlier and in Section 4.6.

Most of the problems posed in this section share many of the defining characteristics of insight problems described in Section 1. In particular, they do not involve multiple steps, they require at most a very minimal amount of technical knowledge, and most of them can be solved by one or two fairly simple insights, albeit insights that are rarely achieved in real time by problem solvers. What makes these problems interesting is that they are posed in such a way as to induce solvers to represent the problem information in an unproductive way. Then the main barrier to finding a solution to one of these problems is to overcome a poor initial problem representation. This may involve such things as a re-representation of the problem, the dropping of an implicit constraint on the solution space, or seeing a parallel to some other similar problem. If the solver finds a productive way of viewing the problem, the solution generally follows rapidly and comes with burst of insight, namely the "aha!" experience. In addition, when non-solvers are given the solution they too may experience a burst of insight.

What follows next are statements of the eighteen brainteasers. The solutions are presented in the Appendix, and we recommend that after whatever problem solving activity a reader wishes to engage in, that the Appendix is studied before reading the remaining two sections of the paper. As we discuss each problem in the paper, we provide authorship information where authorship is known. In addition, we rephrased some of the problems from their original sources.

Problem 3.1. Imagine you have an 8-inch by 8-inch array of 1-inch by 1-inch little squares. You also have a large box of 2-inch by 1-inch rectangular shaped dominoes. Of course it is easy to *tile* the 64 little squares with dominoes in the sense that every square is covered exactly once by a domino and no domino is hanging off the array. Now suppose the upper right and lower left corner squares are cut off the array. Is it possible to tile the new configuration of 62 little squares with dominoes allowing no overlaps and no overhangs?

Problem 3.2. A 3-inch by 3-inch by 3-inch cheese cube is made of 27 little 1-inch cheese cubes of different flavors so that it is configured like a Rubik's cube. A cheese-eating worm devours one of the top corner cubes. After eating any little cube, the worm can go on to eat any adjacent little cube (one that shares a wall). The middlemost little cube is by far the tastiest, so our worm wants to eat through all the little cubes finishing last with the middlemost cube. Is it possible for the worm to accomplish this goal? Could he start with eating any other little cube and finish last with the middlemost cube as the 27th?

Figure 1. The cheese eating worm problem.
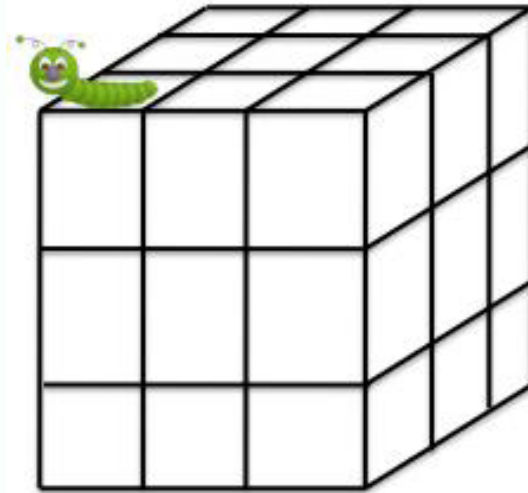
Problem 3.3. You have ten volumes of an encyclopedia numbered 1, . . . ,10 and shelved in a bookcase in sequence in the ordinary way. Each volume has 100 pages, and to simplify suppose the front cover of each volume is page 1 and numbering is consecutive through page 100, which is the back cover. You go to sleep and in the middle of the night a bookworm crawls onto the bookcase. It eats through the first page of the first volume and eats continuously onwards, stopping after eating the last page of the tenth volume. How many pieces of paper did the bookworm eat through?

*Figure 2.* Bookcase setup for the Bookworm Problem.
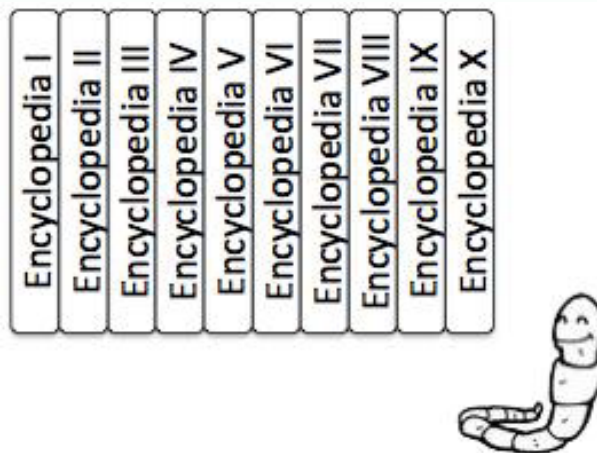
Problem 3.4. Suppose the earth is a perfect sphere, and an angel fits a tight gold belt around the equator so there is no room to slip anything under the belt. The angel has second thoughts and adds an inch to the belt, and fits it evenly around the equator. Could you slip a dime under the belt?

Problem 3.5. Consider the cube in Figure 1 and suppose the top and bottom surfaces are painted red and the other four sides are painted blue. How many little cubes have at least one red and at least one blue side?

Problem 3.6. Look at the nine dots in Figure 3. Your job is to take a pencil and connect them using only three straight lines. Retracing a line is not allowed and removing your pencil from the paper as you draw is not allowed. Note the usual nine-dot problem requires you to do it with four lines; you may want to try that stipulation as well.

Figure 3. The setup for the Nine-Dot Problem.
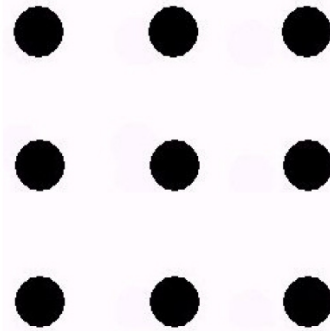
Problem 3.7. You are standing outside a light-tight, well-insulated closet with one door, which is closed. The closet contains three light sockets each containing a working light bulb. Outside the closet, there are three on/off light switches, each of which controls a different one of the sockets in the closet. All switches are off. Your task is to identify which switch operates which light bulb. You can turn the switches off and on and leave them in any position, but once you open the closet door you cannot change the setting of any switch. Your task is to figure out which switch controls which light bulb while you are only allowed to open the door once.

*Figure 4.* The setup of the Light Bulb Problem.
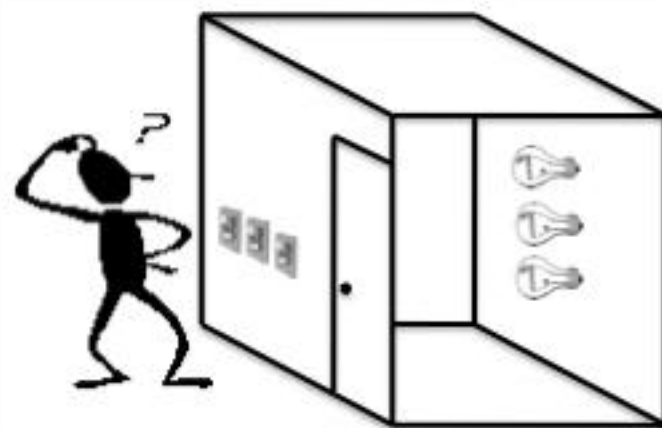
Problem 3.8. We know that any finite string of symbols can be extended in infinitely many ways depending on the inductive (recursive) rule; however, many of these ways are not 'reasonable' from a human perspective. With this in mind, find a reasonable rule to continue the following series:

$$A_{BCD}E_FG^{HI}_JKLM............$$

Problem 3.9. You have two quart-size beakers labeled A and B. Beaker A has a pint of coffee in it and beaker B has a pint of cream in it. First you take a tablespoon of coffee from A and pour it in B. After mixing the contents of B thoroughly you take a tablespoon of the mixture in B and pour it back into A, again mixing thoroughly. After the two transfers, which beaker, if either, has a less diluted (more pure) content of its original substance - coffee in A or cream in B? (Forget any issues of chemistry such as miscibility).

Figure 5. The setup of the Coffee and Cream Problem.
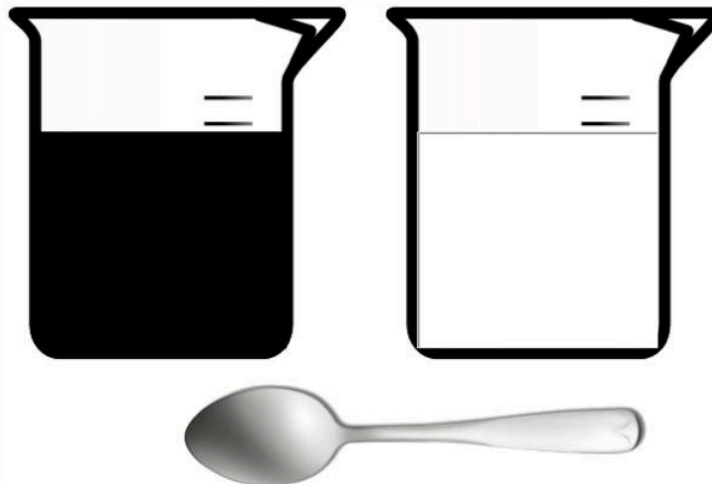
Problem 3.10. There are two large jars, A and B. Jar A is filled with a large number of blue beads, and Jar B is filled with the same number of red beads. Five beads from Jar A are scooped out and transferred to Jar B. Someone then puts a hand in Jar B and randomly grabs five beads from it and places them in Jar A. Under what conditions after the second transfer would there be the same number of red beads in Jar A as there are blue beads in Jar B.

Problem 3.11. Two trains A and B leave their train stations at exactly the same time, and, unaware of each other, head toward each other on a straight 100-mile track between the two stations. Each is going exactly 50 mph, and they are destined to crash. At the time the trains leave their stations, a SUPERFLY takes off from the engine of train A and flies directly toward train B at 100 mph. When he reaches train B, he turns around instantly,

continuing at 100 mph toward train A. The SUPERFLY continues in this way until the trains crash head-on, and on the very last moment he slips out to live another day. How many miles does the SUPERFLY travel on his zigzag route by the time the trains collide?

Problem 3.12. George lives at the foot of a mountain, and there is a single narrow trail from his house to a campsite on the top of the mountain. At exactly 6 a.m. on Saturday he starts up the trail, and without stopping or backtracking arrives at the top before 6 p.m. He pitches his tent, stays the night, and the next morning, on Sunday, at exactly 6 a.m., he starts down the trail, hiking continuously without backtracking, and reaches his house before 6 p.m. Must there be a time of day on Sunday where he was exactly at the same place on the trail as he was at that time on Saturday? Could there be more than one such place?

Problem 3.13. You are driving up and down a mountain that is 20 miles up and 20 miles down. You average 30 mph going up; how fast would you have to go coming down the mountain to average 60 mph for the entire trip?

Problem 3.14. During a recent census, a man told the census taker that he had three children. The census taker said that he needed to know their ages, and the man replied that the product of their ages was 36. The census taker, slightly miffed, said he needed to know each of their ages. The man said, "Well the sum of their ages is the same as my house number." The census taker looked at the house number and complained, "I still can't tell their ages." The man said, "Oh, that's right, the oldest one taught the younger ones to play chess." The census taker promptly wrote down the ages of the three children. How did he know, and what were the ages?

Problem 3.15. A closet has two red hats and three white hats. Three participants and a Gamesmaster know that these are the only hats in play. Man A has two good eyes, man B only one good eye, and man C is blind. The three men sit on chairs facing each other, and the Gamesmaster places a hat on each man's head, in such a way that no man can see the color of his own hat. The Gamesmaster offers a deal, namely if any man correctly states the color of his hat, he will get $50,000; however, if he is in error, then he has to serve the rest of his life as an indentured servant to the Gamesmaster. Man A looks around and says, "I am not going to guess." Then Man B looks around and says, "I am not going to guess." Finally Man C says, " From what my friends with eyes have said, I can clearly see that my hat is _____". He wins the $50,000, and your task is to fill in the blank and explain how the blind man knew the color of his hat.

Problem 3.16. A king dies and leaves an estate, including 17 horses, to his three daughters. According to his will, everything is to be divided among his daughters as follows: 1/2 to the oldest daughter, 1/3 to the middle daughter, and 1/9 to the youngest daughter. The three heirs are puzzled as to how to divide the horses among themselves, when a probate lawyer rides up on his horse and offers to assist. He adds his horse to the kings' horses, so there will be 18 horses. Then he proceeds to divide the horses among

the daughters. The oldest gets ½ of the horses, which is 9; the middle daughter gets 6 horses which is 1/3rd of the horses, and the youngest gets 2 horses, 1/9th of the lot. That's 17 horses, so the lawyer gets on his own horse and rides off with a nice commission. How was it possible for the lawyer to solve the heirs' problem and still retain his own horse?

Problem 3.17. A logical wizard offers you the opportunity to make one statement: if it is false, he will give you exactly ten dollars, and if it is true, he will give you an amount of money other than ten dollars. Give an example of a statement that would be sure to make you rich.

Problem 3.18. Discover an interesting sense of the claim that it is in principle impossible to draw a perfect map of England while standing in a London flat; however, it is not in principle impossible to do so while living in a New York City Pad.

4. Barriers to a Theory of Insight Problem Solving

As mentioned earlier, our view is that there are a number of theoretical barriers that make it difficult to develop a satisfactory formal theory of the cognitive processes in play when humans solve classical brainteasers of the sort posed in Section 3. Further these barriers seem almost unique to insight problem solving in comparison with the more fully developed higher process areas of the cognitive sciences such as human memory, decision-making, categorization, and perception. Indeed it seems uncontroversial to us that neither human nor machine insight problem solving is well understood, and compared to other higher process areas in psychology, it is the least developed area both empirically and theoretically.

There are two recent comprehensive critical reviews concerning insight problem solving by Ash, Cushen, and Wiley (2009) and Chu and MacGregor (2011). These articles describe the current state of empirical and theoretical work on insight problem solving, with a focus on experimental studies and theories of problem restructuring. In our view, both reviews are consistent with our belief that there has been very little sustainable progress in achieving a general scientific understanding of insight. Particularly striking is that are no established general, formal theories or models of insight problem solving. By a general formal model of insight problem solving we mean a set of clearly formulated assumptions that lead formally or logically to precise behavioral predictions over a wide range of insight problems. Such a formal model could be posed in terms of a number of formal languages including information processing assumptions, neural networks, computer simulation, stochastic assumptions, or Bayesian assumptions.

Since the groundbreaking work by the Gestalt psychologists on insight problem solving, there have been theoretical ideas that have been helpful in explaining the cognitive processes at play in solving certain selected insight problems. Among the earlier ideas are Luchins' concept of *einstellung* (blind spot) and Duncker's *functional fixedness*,

as in Maher (1992). More recently, there have been two developed theoretical ideas: (1) Criterion for Satisfactory Progress theory (Chu, Dewald, & Chronicle, 2007; MacGregor, Ormerod, & Chronicle, 2001), and (2) Representational Change Theory (Knoblich, Ohlsson, Haider, & Rhenius, 1999). We will discuss these theories in more detail in Section 4. While it is arguable that these theoretical ideas have done good work in understanding in detail a few selected insight problems, we argue that it is not at all clear how these ideas can be generalized to constitute a formal theory of insight problem solving at anywhere near the level of generality that has been achieved by formal theories in other areas of higher process cognition.

The dearth of formal theories of insight problem solving is in stark contrast with other areas of problem solving discussed in Section 4.6, for example move problems discussed earlier and the more recent work on combinatorial optimization problems such as the two dimensional traveling salesman problem (MacGregor and Chu, 2011). In addition, most other higher process areas of cognition are replete with a variety of formal theories and models. For example, in the area of human memory there are currently a very large number of formal, information processing models, many of which have evolved from earlier mathematical models, as in Norman (1970). In the area of categorization, there are currently several major formal theories along with many variations that stem from earlier theories discussed in Ashby (1992) and Estes (1996). In areas ranging from psycholinguistics to perception, there are a number of formal models based on brain-style computation stemming from Rumelhart, McClelland, and PDP Research Group's (1987) classic two-volume book on parallel distributed processing. Since Daniel Kahneman's 2002 Nobel Memorial Prize in the Economic Sciences for work jointly with Amos Tversky developing *prospect theory*, as in Kahneman and Tversky (1979), psychologically based formal models of human decision-making is a major theoretical area in cognitive psychology today. In our view, there is nothing in the area of insight problem solving that approaches the depth and breadth of formal models seen in the areas mentioned above.

In the following subsections, we will discuss some of the barriers that have prevented the development of a satisfactory theory of insight problem solving. Some of the barriers will be illustrated with references to the problems in Section 3. Then, in Section 5 we will assuage our pessimism a bit by suggesting how some of these barriers might be removed in future work to facilitate the development of an adequate theory of insight problem solving.

4.1 Lack of Many Experimental Paradigms

There are not many distinct experimental paradigms to study insight problem solving. The standard paradigm is to pick a particular problem, such as one of the ones in Section 3, and present it to several groups of subjects, perhaps in different ways. For example, groups may differ in the way a hint is presented, a diagram is provided, or an instruction

is given on how to represent the problem. This is the approach initiated originally by the Gestalt psychologists, as in Luchins (1942), Duncker (1945), Birch and Rabinowitz (1951), and it has continued to the present day as the common experimental paradigm for studying insight problem solving, as in MacGregor, Ormerod, and Chronicle (2001), Ormerod, MacGregor, and Chronicle (2002) and Chu, Dewald, and Chronicle (2007). Differences in the proportion of solvers may then be traced to the differences in the way that a problem is presented and therefore processed. The barrier for insight problem solving theory that we see in the usual paradigm is that in such studies it is difficult to generalize from one insight problem to other insight problems. As a consequence any general theory accrued from such experiments tends to take the form of simple, problem specific ideas like Luchins' *einstellung*, Duncker's *functional fixedness*, or the popular term *thinking outside of the box* (Adams, 1974). This concern is discussed in more detail in Section 4.5.

This lack of a rich collection of problem solving paradigms for insight problems contrasts markedly with other successful higher process areas of psychology. For example, consider human memory. There are well explored paradigms for studying iconic memory, echoic memory, short term memory, recognition memory, recall memory, serial recall, source monitoring, semantic memory, procedural memory, and prospective memory (Baddeley, Eysenck, & Anderson, 2008), just to name a few of the standard experimental settings. Numerous experimental studies in each of these paradigms have led to many empirical generalizations as well as a multiplicity of computational theories, each open to experimental testing. Further, some of these theories interrelate the different types of human memory, and cognitive neuroscientists are well on their way to discovering the brain processes behind different types of memory phenomena.

The higher process areas of decision-making and perception share with the area of human memory the proliferation of experimental paradigms. In these areas, unlike the area of human memory, there are related scientific disciplines that provide normative concepts to assist in theory formulation and associated experimental paradigms. For example in decision-making, the fields of economics and statistics have supplied many theoretical concepts such as various senses of rationality, subjective probability, and the differential effects of gains and losses, as in Luce and Raiffa (1957) and Kahneman and Tversky (1979). This has led to a variety of human decision-making paradigms in the area of experimental economics, most notably ones connected to a variety of game situations, as in Gintis (2009). In perception, empirical studies use results from physics to construct stimuli, as well as knowledge from the neurosciences to understand the role of brain systems in perception. The closest allied field to insight problem solving would be logic. While it is clear that ideas from the field of logic do assist theory formation in studies of reasoning, we are unaware of any substantial role of logic in aiding in the proliferation of experimental paradigms for insight problem solving.

4.2 The Lack of a Well-classified Set of Stimulus Material

Unfortunately, there are only a limited number of classical insight problems, and those that do exist are often quite different from one another. The lack of a large collection of well-graded, homogeneous classes of insight problems prevents one from running the usual sorts of within and between subject experiments that are based on setting up several experimental and control groups with different problems that are of comparable difficulty. The unavailability of adequate numbers of well-graded insight problems is pointed out in Chu and MacGregor (2011). In addition, what insight problems that do exist have not been classified beyond usual coarse categories such as verbal, mathematical, and spatial (Dow & Mayer, 2004), and within these categories, there are no carefully formulated metrics of similarity between problems or graded levels of difficulty. Indeed, what makes most classical insight problems so endearing is that they are usually one-of-a-kind.

Other more successful higher process areas of cognition do not suffer from such a shortage of large sets of well-classified stimulus material. For example, in the area of human memory there are a variety of norms for verbal material such as concreteness ratings, frequency of usage norms, meaningfulness, affective categories, and association strength between pairs of words (Friendly, Franklin, Hoffman, & Rubin, 1982; Bradley & Lang, 1999), and in the area of categorization, norms exist for category prototypicality, instance similarity, and the degree of category membership (Van Overschelde, Rawson, & Dunlosky, 2004). In decision-making, among the many classes of stimuli are gambles, and members of this class are well calibrated because they are defined by probability distributions over a set of valued outcomes (von Neumann & Morgenstern, 1944). Much of the current cognitive theory in this area rests on extensive studies of choice among gambles. There is hope that the lack of a large, systematic body of stimulus material for insight problem solving may be overcome, and we will expand on a suggestion of Chu and MacGregor (2011) for how this might happen in Section 5.1.

4.3 The Lack of Many Behavioral Measures

In addition to problems of the lack of stimulus material discussed in the previous section, there are problems on the response side as well. In particular, there is a paucity of useful behavioral measures that can be obtained while subjects are in the process of trying to solve an insight problem. For many insight problems like the ones in Section 3, only a very small number of subjects are able to solve the problem within a reasonable period of time. Without some sort of auxiliary data, all one has is the period of time between when the problem is presented and when either a solution occurs or when a deadline is reached. Apart from solution time (for solvers) and the proportion of solvers in each group, there is very little data to work with.

For example, consider Problem 3.3 involving the number of pieces of paper a worm will eat. This problem is presented to the Human Problem Solving class discussed in Sec-

tion 2 on the first day of class. In over 25 years, more than two thousand students have been given this problem, and so far only one student has solved it in the roughly ten minutes during which a solution is allowed. After presenting the problem and waiting for two minutes, students are asked to suggest answers. Apart from haphazard guesses, students have proposed only three reasonable possible answers: 500, 802, and 1,000. Of these three, 802 is very rare, 1,000 always comes up, but generally it is accompanied with a side comment that it is surely wrong, and 500 is the one that draws most interest from the class. Only one person (so far) seems to have put the two needed insights together and correctly concluded that 402 is the correct answer as described in the Appendix.

In our view, this is a great insight problem because it involves going past an initial "aha!" experience to finding a second one, and few insight problems have this property. Nevertheless, virtually everyone has been a non-solver, with only a small subset realizing one of the needed insights (leading to the responses 500 or 802), and this behavioral evidence carries the empirical basis of our theoretical views about this problem. Indeed, in any real time study involving most of the insight problems in Section 3, the norm would be a non-solver, and even if there were solvers, one would only have the proportion of solvers, and their time-to-solution as the empirical measures. Despite the prevalence of failure, almost everyone understands the solution when it is given, and with it, they experience some amount of the "aha!" experience.

There have been efforts to augment the available behavioral measures for insight problems, such as, for example, collecting running verbal protocols from subjects during the course of solution (Ericsson, 2003). This procedure was a mainstay of the work on move problems (Newell & Simon, 1972); however, in the case of insight problems there is little evidence that this addition to the database has led to deep theoretical insights. Another type of behavioral data that has been collected during problem solving is feeling-of-warmth (FOW) ratings, as in Metcalfe (1986). She found that subjects could indicate FOW ratings that were predictive of solution announcements; however, announced solutions were often wrong leaving open the question about whether or not the FOW was tracking the occurrence of an insight experience, as in Weisberg, R. W. (1992). In any event, efforts to incorporate metacognitive measures during problem solving have yet to contribute to a general theory of insight problem solving, and this is significant because in the area of human memory, metamemory measures have played a significant role in theory building (Dunlosky & Bjork, 2008)

One promising measure for some types of insight problems is data obtained from an eye movement camera that track across time which aspects of a stimulus a subject is inspecting. For example, Knoblich, Ohlsson, and Raney (2001) used this measure to study how subjects solved matchstick arithmetic problems described in Section 5.1. This is very useful for problems where parts have to be realigned because eye movements reveal what parts of the problem are being inspected across a time horizon. Another successful

utilization of eye movements is Jones (2003) who used an eye movement camera during solution of the so-called *car park game*, where vertically and horizontally placed cars must be moved to allow a certain car to move out of the car park. Of course, solutions to most insight problems do not involve long periods of visual inspection of stimulus material, so this measure is only available for special problem types. The next section discusses some brain recording measures that have been used to attempt to find areas of the brain that are involved during problem solving, and upon close analysis we suggest that little useful information about insight problem solving has come from utilizing these measures.

4.4 The Lack of Problem Solving Modules

Jerry Fodor (1983) wrote an influential book, *The Modularity of Mind*, in which he argued that cognitive science is unlikely to make substantial theoretical progress in studying cognitive systems that are not *modular*. By a modular system, Fodor meant one that is characterized by special computational devices that are domain specific and operate only on specially defined inputs. He postulated further that such systems are innate (hardwired), domain specific, have fixed neural locations, are inaccessible to consciousness, and are characterized by specific breakdown patterns. Examples of modular systems in Fodor's book are mechanisms of color vision, face recognition, speech recognition, and mechanisms that associate grammatical descriptions to utterances. In particular, reasoning and problem solving, as well as other areas of higher processes like memory and decision-making, are non-modular in Fodor's sense because they involve combining the outputs from a variety of cognitive processes operating on various types of information drawn from several sources. The proposition that only modular systems in Fodor's sense could be productively studied led to most psychologists, and eventually even Fodor (2000) himself, to accept that this strict definition of modularity would not work to delimit the productive study of the mind.

However, some cognitive theorists advanced a weakened and potentially more productive concept of modularity known as the massive modularity thesis (Sperber, 1994; Samuels, 1998). This concept views minds as structured by a computational architecture consisting of many innate modules, each based on functional specialization developed as evolutionary adaptation by natural selection. In a nutshell, the view is that human minds have evolved to adapt to life's challenges by developing a number of specialized computational devices rather than a general problem solving system. This view allows some information-processes mechanisms to be modular, such as ones that Fodor referred to as non-modular "central" processes: certain types of reasoning, different types of memory, and decision-making (Barrett & Kurzban, 2006; Cosimides, & Tooby, 1994). According to Barrett and Kurzban (2006), the signal properties of modularity can be characterized by functionally specialized mechanisms that contain properly defined informational inputs. Furthermore, modules evolved to process information to help the organisms' fitness.

With this definition of modularity, it is easy to see that a number of specific problem solving mechanisms could be functionally useful for an organism's survival; however, it seems implausible that a general insight problem solving system would qualify as a modular system even under this weaker sense. For example according to evolutionary psychology, we have the same mental processes as the Neolithic Stone Age men of 12,000 years ago, and it took centuries for them to achieve useful insights that once understood can be easily taught to others, for example fire making technology and specially designed uses of wheels. Regardless of the definition of modularity, two conclusions seem to be justified. First, theoretical advances are more likely to occur in modular rather than non-modular systems, and second; insight problem solving, in comparison with other higher process areas of cognition, is far from being modular in any sense of the term.

By any definition, modular systems should manifest themselves in specific neural locations. Currently there are imaging studies in every higher process area of cognition, including a few in the area of insight problem solving. For example, there are several studies in the literature that have attempted to look for a link between the act of insight (the "aha!" experience) and brain regions. A few of the brain regions that have been linked to insight are bilateral insula, right prefrontal cortex, anterior cingulate (Aziz-Zadeh, Kaplan, & Iacoboni, 2009), hippocampus (Luo & Niki, 2003), precuneus, left inferior/middle frontal gyrus, inferior occipital gyrus, cerebellum (Qiu et al., 2010) and right anterior superior temporal gyrus (Jung-Beeman et al., 2004; Bowden, Jung-Beeman, Fleck, & Kounios, 2005; Kounios & Jung-Beeman, 2009).

It is evident from these few studies that there seems to be no consensus on a neural correlate of insight, and instead it is the case that neural activities associated with insight are spread across several brain regions. Further, many of the brain regions that have been proposed seem to be correlated with other functions that may explain the observed research results. For example, Kounios and Jung-Beeman (2009) suggested that the right anterior superior temporal gyrus is a brain region for insight by using compound remote association problems (Bowden & Jung-Beeman, 2003). These problems test a subject's ability to produce a single word that would be a compound remote associate with three presented words, for example, finding the word 'dog', given the three words: 'patch', 'tired', 'pound'. This type of problem requires that any possible word that may be used must be in semantic memory, because a person is not allowed to make up a new word. Grasby et al. (1993) have suggested that the superior temporal gyrus is part of a system for auditory-verbal memory functions, and these functions must surely be a part of a cognitive activity that solves a compound remote association problem.

In another example, Luo and Niki (2003) had subjects solve Japanese riddles such as "The thing that can move heavy logs, but cannot move a small nail," and the answer is "a river." Each subject was given each riddle for a maximum of 3 minutes before the fMRI

scanning started. Once they were in the fMRI chamber, each subject was given the same riddle and instructed to indicate whether they knew the answer. When a subject did not indicate that they knew the answer, the experimenters presented it. Their conclusion was that the hippocampus is a brain region for insight. The hippocampus has also been linked to stimulus novelty (Tulving et al., 1994; Tulving, Markowitsch, Craik, Habib, & Houle, 1996; Knight, 1996; Grunwald, Lehnertz, Heinze, Helmstaedter, & Elger, 1998), and Luo and Niki (2003) addressed this by stating that because the solutions to the riddles were everyday concepts it is unlikely that the hippocampus showed greater blood-oxygen level-dependent (BOLD) response due to stimulus novelty. However, Yassa and Stark (2008) have found that the hippocampus is linked to stimulus novelty, and what makes their study relevant in this situation is that they used familiar images such as a sailboat or the Mona Lisa by Leonardo Da Vinci. This suggested that although a subject may have seen these images before in their lifetime, the first presentation of a stimulus in an experimental condition elicits a BOLD response in the hippocampus.

The other neuroanatomy associated with insight such as the prefrontal cortex and anterior cingulate also have been associated with emotions (Greene, Sommerville, Nystrom, Darley, & Cohen, 2001). These studies consist of self-reports of an insight experience so the data may be somewhat biased. Also, it is difficult to create experiments that would precipitate the "aha!" moment, and many of these experiments involve the experimenter presenting the solution and asking the subject after viewing the answer if they experienced an "aha!" After reviewing the few studies on neural specificity of insight problem solving, we believe there is scant evidence available that indicates that there are specific brain regions that can be reliably associated with any common aspect of the process of insight problem solving, and this conclusion is consistent with our view that insight problem solving is non-modular.

4.5 The Lack of Theoretical Generalizations

While there is detailed and arguably adequate theoretical understanding of particular insight problems, the task of finding theoretical generalizations that apply across a wide class of insight problems has proven to be quite difficult. For example, a critical step in finding a solution to most classical insight problems is discovering a productive cognitive representation of the problem. Such a representation guides the search for a solution, and if the initial representation of a problem fails to include any route to a solution, then until the initial representation is modified or replaced, a solution is impossible. Many classical insight problems have been constructed to minimize the probability a solver will initially represent a problem in a productive way. In such cases, often the "aha!" experience along with the solution comes shortly after the solver sees a new, productive way to represent the problem. We believe that a crucial centerpiece in a deep theory of insight problem solving would be a theory of problem representation.

Unfortunately, to our knowledge, very little progress has been made in formulating such a theory. Ash, Cushen, and Wiley (2009) provide an interesting critical review of work on restructuring the problem representation while solving an insight problem. In their review, they discuss both experimental studies and theoretical ideas, and it appears from their review that little sustainable scientific progress has been made on understanding insight problem restructuring. In particular, they point out obstacles in interpreting experimental data during problem solving as supporting the occurrence or lack of occurrence of spontaneous (unaided) restructuring. Many of their concerns are echoed in our remarks in Sections 4.1, 4.2, and 4.3 concerning, respectively, the lack of a variety of experimental paradigms, the lack of a large set of well-classified insight problems, and the lack of a variety of behavioral measures. In this section we discuss a number of classical insight problems, mostly drawn from Section 3, that are specifically designed to thwart an initial productive problem representation. Our main conclusion is that while the representational issues of each of these problems may be able to be understood in some detail, there are not any clear generalizations that could be used to construct a formal theory of representation that applies to insight problem solving in general.

First consider Problem 3.4 concerning the band around the earth. This problem is related to the 'rope around the earth' problem credited to the mathematician and puzzle master Harry Langman (1934). Once, an informal study was conducted in the problem solving course described in Section 2, in which different groups of students were presented with this problem in different forms. Each group was given a version of the problem with a different sized sphere, and the question in each case was whether or not a dime could be slipped under the belt. In one group, the sphere was represented by the earth, in another group by a basketball, and in the third by a golf ball. The results were that the number of correct 'yes' answers significantly increased as the volume of the sphere decreased. Talking with the class after the experiment, it was clear that none of the students thought through the simple fact, learned at least by the 7th grade, that the circumference C of a circle is directly proportional to its radius R, $C = 2\pi R$. Thus increasing the circumference by 1" increases the radius by $1/2\pi \approx 0.16$" no matter what is the value of the original circumference. It is a reasonable conjecture that almost all students would have solved the problem easily if they were told directly that the angel added 0.16" to the radius, and in fact with that increase in the radius two dimes could be slipped under the belt. Instead, it seems clear that students were representing the problem with mental imagery, and imagery suggests that the addition of an inch to the circumference of an earth-sized sphere will not raise the band more than an infinitesimal amount. Thus for this problem using mental imagery seems to be leading to an unproductive representation of the problem.

Another example of how imagery can have a very strong, negative effect on problem representation is Problem 3.12 concerning the hiker. Given the description in Section 3, only a small number of students in the problem-solving class were able to see clearly

that there is exactly one place on the trail where the hiker has been at the same time of day while hiking up on Saturday and hiking down on Sunday. On the other hand, if the problem is presented as a friend walking up the trail starting at 6:00 a.m. on the same day as the hiker is walking down, everyone sees that they must pass each other at exactly one point somewhere on the trail, and in addition, most solvers see that when they pass it will be at the same time of day. Why does one description of the problem yield so few correct answers while the other one, logically equivalent, yields many correct answers? We would conjecture that the difference is that the second version, but not the first version, allows the events to take place in the same spatiotemporal setting, where mental imagery can track the hikers. Some conformation for this conjecture comes from the fact that students are more likely to solve the problem when they are given a hint to imagine that the Saturday hike up the mountain is taking place again on Sunday, with the down hikers 'ectoplasmic' body recreating his earlier Saturday hike. The work with the hiker problem has led to a possibly interesting theoretical insight about the relative ease of imagining and reasoning about events that take place in space and time.

Despite gaining an understanding of the barriers to solutions of Problems 3.4 and 3.12, it is difficult for us to imagine exactly what can be generalized from these discoveries that could be used in constructing a general theory of the role of imagery in insight problem solving. For example, consider Problem 3.5, where subjects are asked to figure out the number of little cubes that have at least one red and one blue face. This is a relatively easy problem, and many subjects solve this problem readily without the use of paper and pencil. The students that fail to solve the problem usually report 18 cubes rather than the correct answer of 16 cubes, probably because they erroneously include the middle top and bottom cubes in their count. It is very difficult for us to imagine how humans can solve this problem without using imagery. If that is the case, then we have learned that imagery sometimes helps solvers and at other times it fails them; however, exactly when it helps, when it fails to help, and why are the important theoretical issues that need to be addressed. The answers to these questions seem to be ad hoc, namely they depend on the particular problem in question.

An interesting feature of problem representation is that it is often possible to present a problem in several ways that are logically equivalent; however, different presentations strongly affect the likelihood that a solver will start with a productive representation. The different ways of posing Problems 3.4 and 3.12 discussed in this section are good examples of this. Another example is the relationship between Problem 3.9 and Problem 3.10. Martin Gardner (1982) has a nice discussion of problems that are logically equivalent to these two problems. Both have tended to stump almost all students in the problem solving class discussed in Section 2, even when the students have had a night to think about them. When the solution to Problem 3.10 is given, most students understand it quickly because they can see that every displaced red bead in the blue bead jar corresponds to a displaced

blue bead in the red bead jar. Indeed, many non-solvers achieve an "aha!" experience after that explanation. Problem 3.9, on the other hand, usually stumps everyone, and when the solution along the lines of the solution to Problem 3.10 is given, there are still many blank faces. Sometimes arguments break out, and in an effort to quell them the following proof using 7th grade algebra is presented.

Let *V* be the volume of liquid in the beakers and *T* the volume of a tablespoon. Then after the tablespoon of coffee is added to the cream, the purity of the cream is *V*/(*V*+*T*). Now when a tablespoon of the mixture is returned to the coffee, the amount of coffee is given by

$$(V - T) + T[T/(V + T)]$$

and the purity of the coffee is given by

$$\frac{(V - T) + T[T/V + T]}{V} = \frac{V}{V + T}$$

and that is the same purity as the cream. Unfortunately, some students remain stumped, even after the presentation of the proof. The only successful explanation of the solution to the students is to ask them to imagine that the coffee and cream separates into different regions of the beakers. Then it is relatively easy to convince them that the volume of coffee left in the cream beaker is the same as the (displaced) volume of cream that is now in the coffee beaker. This explanation is essentially the same as the explanation for the solution of Problem 3.10.

Why is Problem 3.10 easier to grasp and represent productively than is Problem 3.9? Our conjecture is that it has to do with the different linguistic status of the words, namely 'bead' versus 'coffee' and 'cream'. Bead is a count noun and coffee and cream are mass nouns; for example, you can have seven red beads but not seven creams (at least not in the sense of the problem). Thus 'purity' becomes a ratio of the pure to the total rather than an amount of transferred beads as in Problem 3.10. Further the cream and tablespoon of coffee are 'thoroughly mixed,' whereas for the bead problem there is no stipulation other than the number placed on the size of the transfers. As Gardner (1982) points out in his discussion of variations on these problems, it is not necessary to thoroughly mix the coffee in the cream jar. Furthermore, if it is the relative amount of cream in the coffee versus coffee in the cream, the size of the beakers does not matter so long as at the end of the transfer the amounts in each beaker are the same as at the beginning. Irrespective of whether or not our conjecture about count versus mass nouns is correct, it is difficult for us to imagine what can be learned from studying this problem that can inform a general theory of insight problem representation. Apart from a few general classes of insight problems, it seems to us that each insight problem has its own character that influences why an unproductive initial representation is likely.

An interesting subclass of insight problems involves a story setting where humans interact, and each person in the story knows some facts that the others do not know. Examples of these problems are Problem 3.14 and Problem 3.15. For versions of Problem 3.14, many solvers have difficulty seeing how the statement 'the sum of their ages is the same as my house number' can convey any information to the census taker since the number is not revealed to the solver. Those who fail to see that it is the census taker who does see the number and can reason from it often find the rest of the story baffling. In the case of Problem 3.15, the blind man has all the information he needs about the color of his hat from the fact that the others, who do see his hat, have failed to take the Gamesmaster up on her offer.

There is a collection of classical insight problems like Problems 3.14 and 3.15 that many would regard as the most difficult of all insight problems. These problems are usually called sum-product problems, and they derive from a problem posed originally by Freudenthal (1968). The setting involves two mathematicians who are trying to figure out which two positive integers have been selected by a third mathematician. The first mathematician knows the sum of the two numbers and the second knows their product. Then they have a remarkably brief conversation about what each knows that the other can't know, and then they both announce that they now have the answer. There are several variations on this problem type, and a recent discussion can be found in Born, Hurken, & Woeginger (2008).

These problems along with Problems 3.14 and 3.15 all share a representational requirement that is an essential aspect of expert ability in some multiperson games, like bridge, hearts, or gin rummy. In games such as these, each player sees their own cards but not the cards of the other player(s). During play, there is an exchange of overt information according to the rules of the game, such as bidding for the contract in bridge or requesting a card denomination from one's opponent in gin rummy. The overt communications coupled with knowledge of one's own cards and the cards that have been revealed in play allow a player to make inferences about the cards the other players hold, and in this way make better decisions about future actions during the game. Being able to simultaneously represent what other players know from their actions is a key to achieving expertise in these games. While we can imagine that a formal theory of representation during problem solving for such situations might be developed, it is difficult to imagine that this theory would have concepts in it that would generalize to representational issues for the other problems discussed so far in this section.

Problem 3.16 is an example of another variety of insight problem that resists a productive representation. This problem sometimes crops up in books on brainteasers for children, and, according to Petković (2009), it was initially posed by the Italian mathematician Niccolo Tartaglia (1500–1557). In the problem, we are given the rules to divide the estate of a king among his three daughters, along with an enthralling story. The story

starts by pointing out that it appears to be impossible to carry out the conditions of the estate, and then, as if by magic, a solution is found. The estate conditions have a crucial arithmetic flaw (see Appendix); however, the story is so intriguing that a solver is likely to miss the flaw and thereby incorrectly represent the problem. In a sense this type of insight problem is like a magic trick, in that the story is designed to divert the solver's attention from finding a productive representation of the problem.

A much simpler problem having similar magic trick characteristics is often used to fool very young children who are just learning arithmetic. The problem is as follows. Suppose John has three haystacks and Jim has five haystacks. They decide to put them all together. How many haystacks do they now have? Often a smart child will respond quickly with eight; however, if good-natured, they will smile when they are corrected with the correct answer, namely one. This problem as well as 3.16 suggest that one component of a general theory of problem representation will need to deal with how problem wording misdirects attention.

Problem 3.11 and Problem 3.13 both deal with computing average speed, and they are also examples of problems involving elementary arithmetic, where the storyline of the problem promotes an unproductive representation. Problem 3.11 about the SUPERFLY is a form of a famous problem that crops up in many collections of brainteasers. Our version specifies that the fly will have to travel on a zig-zag route between the two trains before they crash, and it invites an unwary solver to set up and sum an infinite series of the zig and zag lengths. Such a representational set will block the solver from solving the problem quickly by applying the simple rule that the product of average speed and total travel time gives the total distance traveled. In fact, MacRae (1992, p.10) in a humorous passage, states that John von Neumann, the famous mathematician and scientist, provided the answer to an equivalent problem immediately upon hearing the problem. When teased that he was not thinking like a mathematician, he is said to have replied that he in fact set up and summed the infinite series in his head in a couple of seconds.

Problem 3.13 is one of many brainteasers involving averaging speeds. Of course any college student knows that average speed is the ratio of distance traveled to travel time; however, in reasoning about averaging speeds, incorrect intuitions often trump calculations with this simple formula. When told that in the first half of a forty-mile trip the average speed was 30 mph, students in the problem solving class discussed in Section 2 often state that an average speed of 90 mph for the second half of the trip will result in a 60 mph average for the entire trip. Even mathematically literate students are often surprised to discover that it is impossible to travel fast enough for the second half of the trip to average 60 mph for the whole trip, and this is true regardless of the length of the trip. It seems plausible that a theory could be worked out for a subclass of problems like the four above involving wording of arithmetic information; however, as with the other problem types discussed in this section, it seems to us that generalizations that hold across other problem types are in short supply.

Perhaps the most famous brainteaser involving problem representation is the so-called nine-dot problem, a version of which is presented as Problem 3.6. This problem is usually presented as requiring four continuously drawn lines with no backtracking to cover all nine dots. A solution is illustrated in the left panel of Fig. 6. This problem satisfies some of the definitional requirements of a move problem in that there is a well defined starting state and allowable operations; however, the goal state is not explicitly depicted but instead defined in terms of a criterion that would admit an infinite number of possibilities. Subjects' attempting to solve the nine-dot problem often act unproductively as if the problem stipulation requires that solutions must involve lines that are confined to the area circumscribed by the convex hull of the nine dots, which in this case has the shape of a square box. This assumption leads to a problem representation that does not offer a solution path, so the key to solving this problem is to consider drawing lines outside of the box. Thus the nine-dot problem has become a prototype for the popular phrase 'thinking outside of the box.'

In a very creative discussion of the nine-dot problem, Adams (1974) shows how it is possible to find solutions that satisfy the stipulation with less than four lines. The solution on the right side of Figure 6 covers the case of three lines, and Adams (1974) goes on to provide solutions of fewer than three lines that nevertheless respect the given problem stipulation. These solutions importantly focus on the exact nature of a dot, which in ordinary language is not the same as a point in Euclidean geometry, which has zero area.

MacGregor, Ormerod, and Chronicle (2001) provide an interesting theoretical analysis of the nine-dot problem by explicitly viewing it as a move problem. Their analysis is based on an information processing account of the problem that utilizes some of the notions in Newell and Simon's (1972) computational model for move problem called the General Problem Solver (GPS) discussed in some detail in Section 4.6. They provide formal definitions and assumptions along with supporting data from five experiments involving adding dots to the problem configuration suggested by their theory. In a nutshell, their theory assumes that the subjects are employing the GPS concept of hill climbing to minimize the difference between the current state and the goal state, and only when they abandon this approach that it is possible to find a productive way of trying new approaches. Thus their theory emphasizes the fact that subjects monitor their solution progress, and they can be held back from productive ideas if they seem to be making satisfactory progress with moves that turn out to be unproductive. In particular for the nine-dot problem, initially trying lines that start and stop on a dot seem to result in good progress. Based on their experiments and theory, the authors argue that the barriers to solving the nine-dot problem are not, as some have argued, related to the inability of the subject to see that they can try lines that are outside the convex hull of the nine dots.

A recent theory of insight problem solving is the Representational Change Theory (RCT) by Knoblich et al. (1999). This theory deals directly with the case that an initial poor problem representation must be overcome to solve certain insight problems. The authors

proposed the theory in the context of so-called matchstick arithmetic problems. These problems present the solver with an invalid arithmetic formula made up of matchsticks depicting Roman numbers and arithmetic signs along with a stipulation to make a valid formula by moving a stated number of matchsticks (usually one). An example is to make a valid formula from the following invalid formula with the movement of one matchstick:

$$||| + | = |.$$

The solution is

$$||| - | = ||.$$

The RCT analyzes such problems by proposing that the initial representation sets unnecessary constraints on the solver that must be removed to achieve a solution. For example, solvers of the problem above may initially attempt to move only vertical lines and fail to see that lines in an arithmetic operator as well as in a Roman number are fair game to move. More generally the RCT seems well adapted to understanding representational issues in cases where a problem presentation can be described by a distribution of activity across pieces of information, and altering this distribution creates a new representation. In the case of the matchstick arithmetic above, seeing the plus sign as decomposable chunk is the key to the solution. While the RCT offers a plausible theoretical account of some subclasses of insight problems, it has a major disadvantage as a general theory of insight problem solving. The disadvantage is that for many insight problems such as the ones in Section 3, it is not at all clear how a theorist can determine possible activity distributions from problem statements. Instead, it appears that for most insight problems, the RCT might be a useful post hoc tool to discuss problem solving behavior rather than a theory that makes formal predictions across a wide class of insight problems.

4.6 The Lack of a Computational Theory

Most mental activities that humans engage in have a number of computational models that are able to simulate empirically some of the known aspects of human experimental data. Three popular computational approaches in higher cognition have been brain-style connectionist models (Rumelhart et al., 1987), models based on the ACT-R architecture (Anderson, 1990), and discrete state Multinomial Processing Tree models (Batchelder & Riefer, 1999; Erdfelder et al., 2009). In the area of problem solving, Newell and Simon (1972) developed and tested GPS as a major computer simulation theory of move problems, and some of the ideas in the GPS such as production systems are part of Anderson's ACT-R architecture. In contrast, the area of insight problem solving has seen very few efforts at computational modeling.

Because of its success, it is reasonable to examine Newell and Simon's GPS in some detail to see if it might offer suggestions on how a productive computational model of insight problem solving might be developed. As stated earlier, move problems are characterized by a well-defined starting state, a set of allowable operators that transform

one problem state into another, and one or more well defined goal states. Examples of move problems include cryptarithmetic problems, river crossing problems, and water jug problems, with several specific examples to be discussed later. The solution to a move problem is a sequence of allowable operators that apply first to the start state, and then successively to the resulting problem states until a goal state is reached. Thus, for move problems, unlike most of the problems in Section 3, the way the problem is posed already sets up a productive problem representation and its associated search space. Some theorists have elaborated on ideas like those in GPS for move problems (Greeno, 1974; Wickelgren, 1974) and some of the ideas developed by Newell and Simon are still playing a productive role in modern theories of thinking and problem solving, e.g. applications of ACT-R (Taatgen & Anderson, 2010) and the Criterion of Satisfactory Progress (CSP) theory discussed later in this section.

In solving a move problem, one is usually faced with several possible choices for the next move at each stage of the problem. While random choice at each stage may lead to a solution, both GPS and human solvers evidence selective strategies in their choice of next moves. The GPS strategy is called *means-ends* analysis. The idea is that when the program is in any current problem state, it sets an initial goal of reaching one of the goal states. If that goal is unachievable by a single move, as it usually is, it attempts to select a productive subgoal that is easier to achieve. GPS selects subgoals by identifying the obstacles that are blocking achieving the initial goal, and then it examines the allowable operators and selects one that leads to a new problem state that overcomes some of the obstacles. To enable GPS to work on a problem, information leading to a similarity metric between pairs of problem states must be imposed on the descriptions of the states. Basically, if the program is in a particular state, and another state is analyzed as being closer (more similar) to a goal state, then it looks for the difference between these states, and selects an operator that reduces this difference. This process leads to a succession of subgoals that are popped on and off a subgoal stack during the solution process (Newell & Simon, 1972; Mayer, 1992). Eventually all subgoals are achieved, and finally one of the goal states is popped off to empty the stack and solve the problem.

Newell and Simon regard GPS as a theory of human problem solving, and they test this notion by comparing the sequences of moves selected by humans with those selected by GPS, showing that, at least in some cases, there is a close correspondence. Sometimes verbal protocols are taken from solvers and used to argue that humans are attempting to engage in the sort of means-ends analysis that is built into GPS. The psychologist Wayne Wickelgren (1974) wrote an interesting book that discusses human strategies for solving move problems like those tackled by GPS. Some of these strategies like subgoaling, working backward, and hill climbing are part of the strategies used in GPS. Thus it seems fair to say that the GPS efforts have led to the development of plausible human computational models for solving move problems.

In addition to a valid problem representation, move problems have another advantage over classical insight problems in that for many move problems it is possible to construct a so-called *ideal observer* model of the problem. An ideal observer is a computational algorithm that can explore the entire problem space and determine sequences of moves that lead from the starting state toward a goal state that minimize a cost function such as the number of moves. The ideal observer is a valuable tool in comparing human performance with various senses of optimal behavior. For example, the so-called Orcs and Hobbits' river-crossing problem, due originally, in another equivalent form, to the Venetian mathematician, Niccoi Fontana Tartaglia (1499-1557), is a prototype of many other river crossing move problems. In this version, three orcs and three hobbits are stranded on the left bank of a river, and they must figure out a way to cross to the other side. The only available means to cross the river is in a two-seater boat. Orcs will eat hobbits if the hobbits are ever outnumbered, and orcs will not run away from the possibility of this opportunity in the future. Mayer (1992, Box 6.8) provides a 13-step sequence of moves that solves the problem, and a number of psychologists have done experiments that try to determine which steps in the solution are the most difficult to find (Greeno, 1974).

Another well-studied move problem is the so-called 'Tower of Hanoi' problem due originally to the French mathematician Edouard Lucus (1842-1891). For this problem, the solver is required to move three disks of different sizes stacked in order of size on an initial peg onto another peg subject to specific rules. It is a trivial matter to generate all the intermediate problem states, and they can be exhibited in a nice diagram (see Maher, 1992, Box 6.7). The availability of a complete solution space for this move problem has led to extensive experimentation with this problem, and in an equivalent setting (the Tower of London problem), it has become a tool in clinical assessment to examine possible deficits in a patient's executive functioning and planning (Shallice, 1982).

The presence of an ideal observer model is not always useful in move problems with huge and complex problem states. For example, in two-dimensional Euclidean versions of the traveling salesman problem, humans seem to approach optimality by applying simply understood strategies (Pizlo et al., 2006), whereas the problem itself is classified as an NP-Hard problem in combinatorial optimization theory (Applegate, Bixby, Chvátal, & Cook, 2006). There are other standard problems that can be regarded as move problems where the ideal observer has little to offer. For example jigsaw puzzles, Sudoku puzzles, and crossword puzzles can be regarded as move problems, but humans solving these problems use, respectively, spatial perception, insights from arithmetic, and semantic knowledge to far exceed random trial-and-error processes that computational algorithms would have to use without specific domain knowledge.

The main missing element in any theory that derives from GPS is how to impose a similarity metric on pairs of problem states. This is essential because the key to means-ends analysis is making moves that reduce the difference between one's current state and

a goal state. For many move problems, selecting a similarity metric seems post hoc, and in practice it is often developed for a particular problem after seeing the general nature of human move sequences. In fact, problems like Problem 3.1 and Problem 3.2 challenge the adequacy of the GPS to handle more complex move problems involving insight in the selection of a similarity metric. To our best account, Problem 3.1 is due originally to Gamow and Stern (1958). It is posed as a move problem; however, there is a very large number of ways that one might try and tile the 62 squares. Apart from exhausting them all by brute force, it would not be possible to arrive at the fact that the remaining 62 squares cannot successfully be tiled. In fact Kaplan and Simon (1990) wrote a program to solve the problem by trial and error, and it took 758,145 domino placements to exhaust all the possibilities. Discovering the idea of restructuring the problem by coloring the squares like a checkerboard (the problem is sometimes called the mutilated checkerboard problem, incorporating the coloring hint) as presented in the Appendix is not something that the GPS is capable of doing. This coloring scheme imposes a parity property on the states in the sense that adjacent squares have opposite colors, and once imposed, it can inform the similarity metric between pairs of states.

Problem 3.2 concerning the cheese-eating worm is also similar to Problem 3.1 because a solution is aided by imposing parity on the problem representation. In particular, the addition of an integer-numbering scheme on the 27 little cubes in the problem representation can lead to an immediate solution (see Appendix). The key insight is that the walls next to any visited little cube will have an even (odd) number if the visited cube has an odd (even) number. If GPS were able to solve this problem without enumeration of all possible cheese-eating routes, it would have to impose something like the parity scheme in its effort to develop a similarity metric between problem states. Unfortunately, unlike the precise representation of a move problem itself, the search space for similarity metrics on a problem's state space is huge, and to be manageable one would seem to have to develop general search heuristics. In that spirit, the tiling problem (3.1) was posed to the developers of the GPS in the late 1960s as a challenge to see if GPS could be upgraded to incorporate restructuring heuristics into the theory. It is precisely in this aspect of GPS that one might seek ways to extend the GPS into a computational theory of insight problem solving; however, to our knowledge, there have not been any computational upgrades to the theory that work generally across a variety of insight problems.

MacGregor, Ormerod, and Chronicle's (2001) theory for the nine-dot problem (3.6) discussed in Section 4.5 does add a productive concept to the ideas incorporated in GPS. They provide a general rationale for why certain unproductive moves are attractive to problem solvers during the early stages of the problem solving efforts. Their assumption is that if a move seems to be making good progress toward the goal, then it, and equivalent variations, may be repeatedly tried, eliminating, or at least postponing discovery of the solution. This idea was developed into a more general theory called the Criterion of

Satisfactory Progress (CSP) theory. Chu, Dewald, & Chronicle (2007) developed the CSP theory and applied it to understand solution behavior in the so-called cheap necklace problem. In this problem, subjects are presented with four, three link pieces of chain, and their task is to make a single necklace with twelve links. This problem is a well-defined move problem in terms of the formal definition discussed earlier and in Section 4.6. The initial and goal states are clearly presented, and the allowable operations are cutting a link, for 2 cents, or closing an open link, for three cents. The goal is achieved by a series of allowable moves that holds the expense to 15 cents or less. The authors of the study argue that subjects feel they are making satisfactory progress by early moves that partly construct the desired necklace, and this leads them to fail to see that the best early moves are to sever all the links in one of the four pieces of chain so that they can be used to splice the other three chains together into the necklace.

As with the theoretical analysis of the nine-dot problem, the CSP theory is couched in terms that are analogous to those used in Newell and Simon's formal theory of move problem solving discussed earlier (also see Mayer, 1992) such as hill climbing to minimize the difference between the current state and the goal state. While it may require insight to break out of the pattern of repeating unproductive moves sequences, it is not surprising that the CPS theory works well for these problems. However, the move problems discussed above fail to satisfy many of the attributes of insight problems as defined in the introductory section of this paper. In particular, the representation of the problems is already provided and the solution is a series of moves rather than a single step. While the CSP is a plausible theory for such problems, it is not at all clear how to apply the notions of the CPS theory to most of the brainteasers from Section 3 discussed in this section.

Despite the successful implementation of a computational model for move problems, no one to our knowledge has been successful in constructing a convincing general computational model for insight problem solving. There have been attempts, for example a recent one is Hélie and Sun (2010), who provide a theory called EII (explicit-implicit interaction theory) based on a connectionist architecture called CLARION (Sun, 2002). Roughly speaking, the theory is based on the simultaneous interaction of rule-based, explicit processes and associative, implicit processes. The computational theory is motivated by trying to incorporate the four stages of problem solving proposed by Wallas (1926). These stages are preparation, incubation, illumination (i.e. insight), and verification. Wallas' theory closely mirrors Köhler's (1925) theory of problem solving developed from observations of apes, and while plausible, it was not developed from an extensive set of empirical findings. The EII theory is mostly concentrated on the stages of incubation and illumination, and Helie and Sun (2010) describe six extant theories of incubation and four for illumination. The theories are quite different from each other; however, somewhat surprisingly, EII is able to provide a reinterpretation of all of them.

Incubation refers to solution facilitation that occurs in some way as a consequence of laying a problem aside for a while, and no one can doubt that this is sometimes beneficial, although it has proven to be a difficult process to study experimentally. In a meta-analysis of studies involving incubation, Sio and Ormerod (2009) found that there are problem specific effects of incubation, and that incubation tends to be beneficial when a problem requires a lot of knowledge activation. But they found relatively little convincing evidence that incubation plays an important role in problem restructuring or when a problem has one as opposed to many possible solutions.

The EII theory did capture some aspects of insight problem solving in a few specially designed verbal problems. For example, subjects were asked how to accomplish the following task described in Schooler, Ohlsson, and Brooks (1993, p. 183),

> A giant inverted steel pyramid is perfectly balanced on its point. Any movement of the pyramid will cause it to topple over. Underneath the pyramid is a $100 bill. How would you remove the bill without disturbing the pyramid?

This problem has been posted many times on the Internet, and several possible solutions have been suggested, such as burning the bill. In order for EII to find solutions, the possible solutions have to be loaded into a layer of CLARION. Then when EII is run, solutions bubble up when their "internal confidence level" reaches a threshold. While EII does not come close, at least yet, to solving insight problems like the ones in Section 3, it may be a start. In fairness, it seems a safe conclusion that successful effort to develop computational theory for insight problem solving lags considerably behind such work in other areas of higher process cognition.

5. Some Possibilities for Progress

The discussion in Section 4 lays out our reasons for why there has been little empirical or theoretical progress in understanding human insight problem solving, especially when compared to the level of progress that has been made in other higher process areas of cognition. We detailed several principled reasons why it is very difficult if not impossible to have a deep, satisfactory theory of human insight problem solving. Despite our pessimism, we do see several possibilities that might advance the field, and hopefully future work will prove our pessimistic assessment incorrect. We cover these possibilities in the next subsections.

5.1. *Some Useful Classes of Insight Problems*

As discussed in Section 4.1, one of the greatest barriers to the development of a solid empirical base of studies on insight problem solving is the lack of large collections of well-calibrated problems. Chu and MacGregor (2011) also discuss this problem, and they

suggest three classes of insight problems that do not suffer from this limitation. These classes are matchstick arithmetic problems (Knöblich et al., 1999), compound remote association problems (Bowden & Jung-Beeman, 2003), and rebus puzzles (MacGregor & Cunningham, 2008). An example of a matchstick arithmetic problem is given in Section 4.5, and a compound remote association problem is given in Section 4.4. Rebus puzzles involve determining a global meaning of a visual display of letters and numbers, for example a solution to the rebus, '1t345', is 'tea for two'. Solutions to most problems in these three classes come with an "aha!" experience, and it often occurs when a solver is able to access a productive search space. Especially the last two of these classes have many problem instances, and there have been efforts to calibrate them for similarity and difficulty. Another advantage of working with these large classes of brain teasers is that it is possible to set up training studies to understand what aspects of solving one of these problems transfers to another similar problem.

We propose further work with three other classes of insight problems that share some of the advantages of the three classes suggested by Chu and MacGregor (2011). These classes are Bongard problems, tricky series completion problems, and self-referential logic problems. Bongard problems are discussed extensively in Hofstadter (1979). These problems present solvers with a visual pattern recognition task in which they have to discern the difference between two sets of relatively simple diagrams. The instances in one of the sets share an attribute not shared by any of the diagrams in the other set, such as polygons with an odd number of sides, or a particular feature in a fixed orientation. Some of the pattern rules are pretty simple, but others are quite difficult even for experienced human solvers. The solution to a Bongard problem requires the solver to explore visual features at various levels of abstraction, and when a problem is solved it is usually accompanied with an "aha!" experience. Bongard (1970) invented and published 100 of these problems, and at the time of this writing, Harry Foundalis (2006), maintains a website with a growing index of 280 such problems. Also there have been some efforts to grade them for level of difficulty.

Another largely undeveloped class of insight problems that we believe has the potential to have some of the advantages as the ones discussed above is tricky series completion problems. An example is Problem 3.8, in which a solver is asked to continue an alphabetic series. In general, series continuation problems are standard for tests of intelligence; however, for those series problems the proper search space is usually clear. Problem 3.8 is set up to mislead the solver into searching for a solution of the wrong type. When this problem is posed to students in the problem solving class described in Section 2, there are almost no immediate solvers. Then a series of hints are given. First, it is stated that humanities majors solve this problem faster than math majors, and then, after a delay, it is mentioned that elementary students are faster than college students on the problem. By this point several hands go up signaling a perceived solution, but still the bulk

of the class remains clueless. The final hint is to look at the shape of the letters, and that is usually followed by an instantaneous solution along with a big "aha!" experience. What is interesting about this problem is that its structure inevitably leads solvers to initially look at patterns in the sequence of clump sizes of up and down letters. As with Problem 3.6, successful solvers must look for solutions that are outside of the box, so to speak.

After presenting this problem, students are treated to several other series completion problems involving different productive search spaces, for example

$$ℲＭ ℼＳ ℰℨ Ⱥ ⁊...$$

$$OTTFFSS...,$$

and

$$335443....$$

All three of these new series involve, in one way or another, sequences of the first few positive integers: the first involves visual forms of the numbers reflected in a mirror, the second is the first letter of each of the numbers, and the third involves a sequence of the number of letters in each of the number words. After this exposure, students are given a homework problem to make up their own tricky series completion problem. Lots of creative ideas are generated, and from the ease with which students take to the task, we would conjecture that people would exhibit a lot of improvement from training on these problems. It would be a potentially useful task to create a number of these tricky series completion problems and calibrate them for similarity and difficulty level.

There is a large class of classical insight problems that deal with self-referential statements that might be amenable to computational models. Though some of these paradoxes arise from deep issues in mathematical logic, one does not need any training in formal logic to understand some of the classical ones. The earliest examples of self-referential paradoxes appear to have arisen in Greek philosophy, and they take various forms of the so-called liar's paradox. For example, in 4 BC, Eubulides of Miletus is said to have made the statement, "I am lying." An analysis of this statement reveals the paradox, if Eubulides is lying, then the statement is false, and so he isn't lying (at least in the usual sense of two-valued logic), and alternatively, if the statement is true, then it must be a lie, so it is false, and he isn't lying.

A large number of problems dealing with self-referential statements involves a setting where all the natives on an island belong to one of two groups, where in one members always tell the truth (T), and in the other where members always tell lies (L). Many college students have encountered brainteasers that derive from this setting. There are several academic writers of popular books of brainteasers who tie self-referential paradoxes to major advances in mathematical logic, such as Smullyan (1978, 1988) and Hofstadter (1979).

Problem 3.17 is from Smullyan (1988), and it describes a wizard who offers a deal to make money by uttering a single true or false statement. The author claims that it captures an essential aspect of Gödel's famous incompleteness theorem in logic. Students in the problem solving class described in Section 2 are given a couple of days to try and solve the problem, after they have been exposed to various forms of the liar's paradox in class. Typically only a few students succeed in finding a solution like the one given in the Appendix, but when the problem is explained, most students seem to experience insight. Occasionally a student will discover a neat trick that one might play on the wizard, namely, say, "You will give me $10." This paralyzes the wizard in a double bind, using logic similar to the liar's paradox, but it does not result in any monetary gain.

Another example of a self-referential paradox is Problem 3.18, where one examines the possibility of making a perfect map of England. There are of course a number of practical reasons why one could not draw a prefect map of England no matter where they lived; but the problem asks the solver to find an interesting sense that would make it impossible in London but not in New York. In a group problem solving setting, very few students in the problem solving class come up with plausible reasons. Nevertheless in pre-solution discussions, it is quickly recognized that the key must be that London is in England and New York City is not. The next phase of the group discussion turns on the concept of what exactly is a 'perfect map,' and only if it is seen that a perfect map must depict everything, will the group have a chance to solve the problem. Sometimes insight is achieved when a hint about the infinite regress of reflections in two mutually reflecting mirrors is provided, but many students continue to express puzzlement and never seem to get the point. This is surprising because unlike Problem 3.17 and other classical self-referential problems, the statement of the map problem is designed to focus on the self-referential component. It seems that many solvers are unwilling to stretch their minds beyond the fact that it is impossible, from a practical point of view to make a perfect map.

It might be a productive strategy to extract a large number of these problems from the many available sources. Then one could grade them for difficulty for naive solvers. In addition one could conduct experiments with individual subjects over a long time span in an effort to develop expert problem solvers in this area. A study of their problem solving strategies might shed light on some of the latent cognitive processes that naïve solvers bring to bear on such problems.

5.2 Suggestions for Experimental Work

It seems clear to us that some new experimental approaches are needed if insight problem solving is to reach a mature status among the higher process areas in cognition. We suggest three. The first is to accept that insight problem solving is quite a bit more difficult to understand scientifically than other higher process areas of cognition. The usual approaches that have helped shed light on other higher process areas such as simple

behavioral experiments with naive subjects, creative computational models, and brain imaging studies are unlikely to help for a variety of reasons as detailed in Section 4. In particular, very little of the work so far on insight problem solving is likely to be helpful in establishing a formal computational theory of insight, and new ideas and new varieties of experiments will be necessary.

The second step is to continue work on establishing a large body of insight problems in the categories discussed in Section 5.1. Along with an increasing stock of problems in each class, more normative studies establishing their similarity and level of difficulty would be helpful. A third step is to focus on training expert problem solvers in each of these problem domains. It is a reasonable conjecture that many humans can considerably improve their ability to solve these types of problems with extensive experience in generating and solving them. Such experts might have the capacity to rapidly shift their search spaces until the type of space that contains the solution occurs to them. Experimental study of insight in problem solvers who are experienced experts in a particular problem domain seems to us to provide a more productive basis for developing theory than the usual studies with inexperienced subjects such as college sophomores. While we cannot assure that our proposed experimental strategies for studying insight problem solving would advance the field, we are almost certain that if insight problem solving is approached in the experimental ways that it is currently studied, very little sustainable progress will be forthcoming.

5.3 Possibilities for a Computational Theory

It was pointed out in Section 4.6 that so far there is not a generally recognized successful computational model for insight problem solving. In this section we discuss the possibility for a computational theory in less pessimistic terms. There is an ever-present tension between two strategies in developing computational models cognitive activity. One strategy is to develop computational models that are designed to embody some theoretical ideas for how humans perform a cognitive task, and the other strategy is to write computer programs that are designed to perform a cognitive task that humans do as well as possible. This is the familiar difference between the computer simulation and artificial intelligence approaches, and Newell and Simon (1972) explicitly pursue the first of these strategies with their GPS. We discussed some of the impasses in extending the GPS to handle problems that require representational changes in Section 4.6. In this section we discuss both the computer simulation and the artificial intelligence strategies for developing computational models for special classes of insight problems.

In Section 4.6, we argued that it is not realistic to expect that a general computational model of insight problem solving will be forthcoming. An alternative strategy would be to try and develop computational models within particular large insight problem domains to represent the processes of expert problem solvers. Some of the insight problem domains

discussed in Section 5.1 would seem natural to handle with a computational model. For example, a computational model for the class of compound word association problems might be developed using current computational models of semantic organization such as the topics model, as in Griffiths, Steyvers, & Tenenbaum (2007), or word association networks, as in De Deyne and Storms (2008). Further, the extensive word association norms of Nelson, McEvoy, and Schreiber (1998) are already in great use to predict the many aspects of word association experiments.

The domain of matchstick arithmetic problems is much smaller than the domain of compound word association problems, although there are those matchstick move problems that do not deal with arithmetic formula that could increase the size of this domain. It might be possible to develop a computational theory for the entire domain using principles from the computational theories of move problems. The RCT of Knoblich et al. (1999) discussed in Section 4.5 could plausibly be a start because its notions would seem to lend themselves to formal rules.

The other four classes of insight problems discussed in Section 5.1 are rebuses, Bongard problems, tricky series completion problems, and self-reference problems. These classes of problems might pose more challenges for constructing a computational model for an expert problem solver. For example, in series completion problems, one could probably find many logically justifiable ways to continue a given sequence; yet in most cases humans exhibit a clear preference for only one of them. This is similar to the law of prägnanz in Gestalt psychology, which is designed to explain why humans are biased to see a pattern of sensation in one particular way even though logically it could be seen in many different ways. Because of this, computational models for series completion problems would have to represent human principles of good series completion.

The case of rebus puzzles and Bongard problems would probably play out similarly, because humans usually agree on a single solution to most such problems once the solution is revealed. On the other hand, machine intelligence that does not represent human biases could probably generate multiple candidates for the meaning of a rebus or the solution to a Bongard problem. In the case of Bongard problems efforts have been made to construct an artificial intelligence system that solves them, as in Foundalis (2006), and despite progress in this area it is clear at this point that humans are better than the available computer programs for such problems. Finally, in the case of self-reference problems, we know of no efforts to develop a model for how humans represent and solve them. Because of the underlying tight logical structure of such problems, this might prove to be a viable task given data from expert problem solvers in this area.

In addition to developing computational models of expert problem solvers in a particular problem domain, it would be informative to pursue an artificial intelligence approach as well. There are several examples of creative artificial intelligence in several areas of human problem solving such as chess and Go (Newborn, 1997). A more recent

success was the development of a computer program that could beat expert problem solvers in the popular television game of Jeopardy (Baker, 2011). A team of IBM researchers headed by David Ferrucci in 2005 accomplished this success by creating 'Watson.' Watson is an artificial intelligence program that can take encyclopedia type questions in natural language and answer them by a rapid search of its vast store of knowledge. Indeed, it is likely that computer scientists might be attracted to pursue an artificial intelligence route, if cognitive scientists could convince the public that achieving insight in machines was at the same level of challenge as defeating human experts in chess or competing with humans in the Jeopardy game.

There are clear difficulties that must be overcome for an artificial intelligence insight problem solver to succeed. Problem 3.7 illustrates some of these difficulties. To solve this problem, a solver must know pragmatic information about light bulbs in sockets as well as things about the consequences of touching a light bulb shortly after it was in use. Since computers do not have sensory experience, this knowledge would have to come from a vast storehouse of semantic information. That would not be an impossible barrier, for example such an encyclopedia is essentially what Watson has.

In addition to having a knowledge base similar to that of Watson, an artificial intelligence insight problem solver would have to extract productive search spaces from very short problem statements. This challenge does not seem impossible to overcome. In fact, at a joint computer science and psychology conference at Dagstuhl castle in 2011, the senior author read a paper that discussed many of the insight problems in Section 3 and described some of the difficulties in the current state of work on insight problem solving (Batchelder, 2011). In discussion, the conference participants coined the term the "Dagstuhl-Batchelder challenge" (van Rooij, Haxhimusa, Pizlo, & Gottlob, 2011), which is a challenge to computer scientists to create a program that would solve insight problems using as input the problem statements verbatim in natural language exactly as the brain teasers in Section 3 were presented. Even if the challenge were successfully met, it would remain to see if any of the knowledge learned by creating such a program could be transferred to the task of constructing a computational model for human insight problem solving. In this, one is tempted to be pessimistic, but we would hold out some hope for advancement in a human theory of insight problem solving by pursuing this route.

The reason for pessimism is that insights into human theory from artificial intelligence successes do not always occur. For example, for many years artificial intelligence programs have been constructed to play games like chess and Go. Throughout many of these years, computer scientists working with chess masters had expected that aspects of human reasoning during game play would be valuable in aiding the construction of computer chess playing programs. Instead what has been learned by all these efforts is that computer chess programs did not succeed in any way by following aspects of human

reasoning, and instead vast, powerful deep search engines were the key to beating the world chess champion (Hsu, 2004).

Another example of disconnect between human participants and computer programs in game playing occurs in the game of Go. Go is combinatorially much more complicated than chess, and so far there are no computer programs capable of defeating professional Go masters that play without handicap on the usual 19 by 19 board. There are some programs that play expertly or even perfectly on smaller boards, but the received view in computer science is that the day when the world champion at Go will be a computer program is a long way off, if ever. The first computer program that played a complete game on a 19 by 19 board was due to Albert Zobrist (1969) in his prize winning doctoral dissertation. His approach was to build in pattern recognition processes and downplay the role of look-ahead. At that time, Newell and Simon (1958) were supporting the idea that human chess skill is mostly pattern recognition rather than look ahead, so this was a natural start. In 1970, the senior author was invited to join Zobrist's PhD Defense Committee and to play a game against the program with a handicap. The requirement was a 15 stone handicap, a handicap about equivalent to a queen and a rook in chess. After the first 50 moves of the game, the program appeared to have improved its position; however, when it was engaged in close hand-to-hand combat, in general not a good strategy in Go, it proceeded to lose one safe group of stones after another, and most of its territory was eventually wiped out. The lesson here, as in the case of chess, is that successes in computer game playing programs have come from abandoning efforts to incorporate human reasoning strategies into the program.

This divergence between artificial intelligence approaches and human approaches in problem solving is echoed in a report on the November 2008 Workshop on Human Problem Solving held at Purdue University (Goldstone & Pizlo, 2009). These authors came to the same conclusion that we do about the fact that the way that successful computer chess programs operate bears little or no relationship to the mental processes of human chess players. They go on to remark that in the area of solving Bongard problems, discussed in Section 5, humans can outperform extant computer programs. This is interesting because computer programs have replaced humans in many areas of human pattern recognition such as recognizing speech and writing. We conjecture that in the future computer programs will be able to solve Bongard problems; however, it remains to be seen if any of the knowledge needed to accomplish this goal will transfer into ideas for a computational theory of human pattern recognition while solving Bongard problems.

In conclusion, human insight problem solving has proven to be a very complex subject to study scientifically. We do not think that this is because of any inadequacies on the part of the researchers who have worked in this area. Instead, we believe that there are intrinsic reasons why insight problem solving is so refractive to study using the standard approaches that have been successful in studying other areas of higher process cognition.

We follow on the suggestions of others who have worked on insight problem solving to suggest some new approaches to insight problem solving that may prove to be productive.

Acknowledgements

References

Adams, J. L. (1974). *Conceptual blockbusting*. San Francisco: W. H. Freeman.

Anderson, J. R. (1990). *The adaptive character of thought*. Hillsdale, NJ: Erlbaum.

Applegate, D. L., Bixby, R. E., Chvátal, V., & Cook, W. J. (2006). *The traveling salesman problem: A computational study*. Princeton: Princeton University Press.

Ash, I. K., Cushen, P. J., & Wiley, J. (2009). Obstacles in investigating the role of restructuring in insightful problem solving. *The Journal of Problem Solving*, *2*(2), 6–41.

Ash, I. K., Jee, B. D., & Wiley, J. (2012). Investigating insight as sudden learning. *The Journal of Problem Solving*, *4*(2), 1–27.

Ashby, F. G. (Ed.). (1992). Multidimensional models of perception and cognition. Hillsdale, NJ: Erlbaum.

Aziz-Zadeh, L., Kaplan, J. T., & Iacoboni, M. (2009). "Aha!": The neural correlates of verbal insight solutions. *Human Brain Mapping*, *30*(3), 908–916. http://dx.doi.org/10.1002/hbm.20554

Baddeley, A., Eysenck, M. W., & Anderson, M. C. (2008). *Memory*. London: Psychology Press.

Barrett, H. C., & Kurzban, R. (2006). Modularity in cognition: Framing the debate. *Psychological Review*, *113*(3), 628–647. http://dx.doi.org/10.1037/0033-295X.113.3.628

Baker, S. (2011). *Final Jeopardy: Man vs. machine and the quest to know everything*. Boston: Houghton Mifflin Harcourt.

Batchelder, W. H. (2011, August). *Some issues in developing a theory of human problem representation*. Paper presented at Dagstuhl Seminar 11351: Computer Science and Problem Solving: New Foundations, Dagstuhl, Germany.

Batchelder, W. H., & Riefer, D. M. (1999). Theoretical and empirical review of multinomial processing tree modeling. *Psychonomic Bulletin & Review*, *6*(1), 57-86. http://dx.doi.org/10.3758/BF03210812

Birch, H. G., & Rabinowitz, H. S. (1951). The negative effect of previous experience on productive thinking. *Journal of Experimental Psychology*, *41*(2), 121–125. http://dx.doi.org/10.1037/h0062635

Born, A., Hurken, C. A. J., & Woeginger, G. J. (2008). The Freudenthal problem and its ramifications (Part III). *Bulletin of the European Association for Theoretical Computer Science*, *95*, 201–219.

Bowden, E. M., & Jung-Beeman, M. (2003). Normative data for 144 compound remote associate problems. *Behavior Research Methods*, *35*(4), 634–639. http://dx.doi.org/10.3758/BF03195543

Bowden, E. M., Jung-Beeman, M., Fleck, J., & Kounios, J. (2005). New approaches to demystifying insight. *Trends in Cognitive Sciences*, *9*(7), 322–328. http://dx.doi.org/10.1016/j.tics.2005.05.012

Bradley, M. M., & Lang, P. J. (1999). *Affective norms for English words (ANEW): Instruction manual and affective ratings* (Technical Report C-1). Gainesville: University of Florida, Center for Research in Psychophysiology.

Chronicle, E. P., MacGregor, J. N., & Ormerod, T. C. (2004). What makes an insight problem? The roles of heuristics, goal conception, and solution recoding in knowledge-lean problems. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *30*(1), 14–27. http://dx.doi.org/10.1037/0278-7393.30.1.14

Chu, Y. (2009). *Human insight problem solving: Performance, processing, and phenomenology*. Saarbrücken, Germany: VDM Verlag Dr. Müller.

Chu, Y., Dewald, A. D., & Chronicle, E. P. (2007). Theory-driven hints in the cheap necklace problem: A preliminary investigation. *The Journal of Problem Solving*, *1*(2), 18–32.

Chu, Y., & MacGregor, J. N. (2011). Human performance on insight problem solving: A review. *The Journal of Problem Solving*, *3*(2), 119-150.

Cosmides, L., & Tooby, J. (1994). Origins of domain specificity: The evolution of functional organization. In L. A. Hirschfeld and S. A. Gelman, Eds. *Mapping the Mind: Domain Specificity in Cognition and Culture*, pp. 523–543. Cambridge, England: Cambridge University Press. http://dx.doi.org/10.1017/CBO9780511752902.005

De Deyne, S., & Storms, G. (2008). Word associations: Network and semantic properties. *Behavior Research Methods*, *40*(1), 213–231. http://dx.doi.org/10.3758/BRM.40.1.213

Dow, G. T., & Mayer, R. E. (2004). Teaching students to solve insight problems; Evidence for domain specificity in creativity training. *Creativity Research Journal*, *16*(4), 389–398. http://dx.doi.org/10.1080/10400410409534550

Duncker, K. (1945). On problem-solving. *Psychological Monographs*, 58, Whole No. 270. American Psychological Association.

Dunlosky, J., & Bjork, R. A. (Eds.) (2008). *Handbook of metamemory and memory*. New York: Psychology Press.

Erdfelder, E., Auer, T-S, Hilbig, B. E., Aßfalg, A., Moshagen, M., & Nadarevic, L. (2009). Multinomial processing tree models: A review of the literature. *Zeitschrift für Psychologie–Journal of Psychology*, *217*(3), 108–124. http://dx.doi.org/10.1027/0044-3409.217.3.108

Ericsson, K. A. (2003). Valid and non-reactive verbalization of thoughts during performance of tasks: Towards a solution to the central problems of introspection as a source of scientific data. *Journal of Consciousness Studies*, *10*(9–10), 1–18.

Estes, W. K. (1996). *Classification and cognition*. New York: Oxford University press.

Fodor, J. A. (1983). *The modularity of mind: An essay on faculty psychology*. Cambridge, MA: MIT Press.

Fodor, J. (2000). *The mind doesn't work that way: The scope and limits of computational psychology*. Cambridge, MA: MIT Press.

Foundalis, Harry E. (2006). Phaeaco: A cognitive architecture inspired by Bongard's problems. *Dissertation Abstracts International, 67*(04). 461B. (UMI No. 3215222)

Freudenthal, H. (1968). Problem No. 223. *Nieuw Archief voor Wiskunde*, *17*, 152.

Friendly, M., Franklin, P. E., Hoffman, D., & Rubin, D. C. (1982). The Toronto Word Pool: Norms for imagery, concreteness, orthographic variables, and grammatical usage for 1080 words. *Behavior Research Methods*, *14*(4), 375–399. http://dx.doi.org/10.3758/BF03203275

Gamow, G., & Stern, M. (1958). *Puzzle-Math*. New York: Viking Press.

Goldstone, R., & Pizlo, Z. (2009). New perspectives on human problem solving. *Journal of Problem Solving*, *2*(2), 1–5.

Gardner, M. (1978). *Aha! Insight*. New York and San Francisco: Scientific American and W. H. Freeman.

Gardner, M. (1982). *Aha! Gotcha: Paradoxes to puzzle and delight*. New York: W. H. Freeman.

Gintis, H. (2009). *Game theory evolving: A problem-centered approach to modeling strategic interaction (2nd ed.)*. Princeton, NJ: Princeton University Press.

Grasby, P. M., Frith, C. D., Friston, K. J., Bench, C., Frackowiak, R. S. J., & Dolan, R. J. (1993). Functional mapping of brain areas implicated in auditory—verbal memory function. *Brain*, *116*(1), 1–20. http://dx.doi.org/10.1093/brain/116.1.1

Greene, J. D., Sommerville, R. B., Nystrom, L. E., Darley, J. M., & Cohen, J. D. (2001). An fMRI investigation of emotional engagement in moral judgment. *Science, 293*(5537), 2105–2108. http://dx.doi.org/10.1126/science.1062872

Greeno, J. G. (1974). Hobbits and orcs: Acquisition of a sequential concept. *Cognitive Psychology*, *6*(2), 270–292. http://dx.doi.org/10.1016/0010-0285(74)90014-0

Griffiths, T. L., Steyvers, M., & Tenenbaum, J. B. (2007). Topics in semantic representation. *Psychological Review*, *114*(2), 211–244. http://dx.doi.org/10.1037/0033-295X.114.2.211

Grunwald, T., Lehnertz, K., Heinze, H. J., Helmstaedter, C., & Elger, C. E. (1998). Verbal novelty detection within the human hippocampus proper. *Proceedings of the National Academy of Sciences of the United States of America*, *95*(6), 3193–3197. http://dx.doi.org/10.1073/pnas.95.6.3193

Hélie, H., & Sun, R. (2010). Incubation, insight, and creative problem solving: A unified theory and a connectionist model. *Psychological Review*, *117*(3), 994–1024. http://dx.doi.org/10.1037/a0019532

Hofstadter, D. (1979). *Gödel, Escher, Bach: An eternal golden braid*. New York: Basic Books.

Hsu, F-H. (2004). *Behind Deep Blue: Building the computer that defeated the world chess champion*. Princeton, NJ: Princeton University Press.

Jones, G. (2003). Testing two cognitive theories of insight. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *29*(5), 1017–1027. http://dx.doi.org/10.1037/0278-7393.29.5.1017

Jung-Beeman, M., Bowden, E. M., Haberman, J., Frymiare, J. L., Arambel-Liu, S., Greenblatt, R., Reber, P. J., & Kounios, J. (2004). Neural activity when people solve verbal problems with insight. *PLoS Biology, 2*(4), 500–510. http://dx.doi.org/10.1371/journal.pbio.0020097

Kahneman, D., & Tversky, A. (1979). Prospect theory: An analysis of decision under risk. *Econometrica, 47*(2), 263–292. http://dx.doi.org/10.2307/1914185

Kaplan, C. A., & Simon, H. A. (1990). In search of insight. *Cognitive Psychology*, *22*(3), 374–419. http://dx.doi.org/10.1016/0010-0285(90)90008-R

Knight, R. T. (1996). Contribution of human hippocampal region to novelty detection. *Nature*, *383*, 256–259. http://dx.doi.org/10.1038/383256a0

Knoblich, G., Ohlsson, S., Raney, G. E. (2001). An eye movement study of insight problem solving. *Memory & Cognition*, *29*(7), 1000–1009. http://dx.doi.org/10.3758/BF03195762

Knoblich, G., Ohlsson, S., Haider, H., & Rhenius, D. (1999). Constraint relaxation and chunk decomposition in insight problem solving. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *25*(6), 1534–1555. http://dx.doi.org/10.1037/0278-7393.25.6.1534

Köhler, W. (1925). *The mentality of apes*. New York: Harcourt Brace.

Kounios, J., & Jung-Beeman, M. (2009). The aha! moment: The cognitive neuroscience of insight. *Current Directions in Psychological Science*, *18*, 4210–216. http://dx.doi.org/10.1111/j.1467-8721.2009.01638.x

Langman, H. (1934). Problem E: (rope around the earth). *Mathematical Monthly*, *41*, 103.

Loyd, S. (1914). *Cyclopedia of puzzles*. New York: Lamb Publishing Co.

Luce, R. D., and Raiffa, H. (1957). *Games and decisions: Introduction and critical survey*. New York: Wiley.

Luo, J., & Niki, K. (2003). Function of hippocampus in "insight" of problem solving. *Hippocampus*, *13*(3), 316–323. http://dx.doi.org/10.1002/hipo.10069

Luchins, A. (1942). Mechanization in problem solving: The effect of *einstellung*. *Psychological Monographs*, 54, whole no. 248.

MacGregor, J. N., & Cunningham, J. B. (2008). Rebus puzzles as insight problems. *Behavior Research Methods*, *40*, 263–268. http://dx.doi.org/10.3758/BRM.40.1.263

MacGregor, J. N., & Chu, Y. (2011). Human performance on the traveling salesman and related problems: A review. *The Journal of Problem Solving*, *3*(2), 1–29.

MacGregor, J. N., Ormerod, T. C., & Chronicle, E. P. (2001). Information-processing and insight: A process model of performance on the nine-dot and related problems. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *27*(1), 176–201. http://dx.doi.org/10.1037/0278-7393.27.1.176

MacRae, N. (1992). *John von Neumann: The scientific genius who pioneered the modern computer, game theory, nuclear deterrence, and much more.* Providence, RI: American Mathematical Society.

Mayer, R. E. (1992). *Thinking, problem solving, cognition* (2nd ed.). New York: W. H. Freeman.

Metcalfe, J. (1986). Premonitions of insight predict impending error. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *12*(4), 623–634. http://dx.doi.org/10.1037/0278-7393.12.4.623

Nelson, D. L., McEvoy, C. L., & Schreiber, T. A. (1998). The University of South Florida word association, rhyme, and word fragment norms. http://w3.usf.edu/FreeAssociation/Intro.html

Newborn, M. (1997). *Kasparov versus Deep Blue: Computer chess becomes of age.* New York: Springer.

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

Norman, D. A. (Ed.) (1970). *Models of human memory.* New York: Academic Press.

Ormerod, T. C., MacGregor, J. N., & Chronicle, E. P. (2002). Dynamics and constraints in insight problem solving. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *28*(4), 791–799. http://dx.doi.org/10.1037/0278-7393.28.4.791

Petković, M. S. (2009). *Famous puzzles of great mathematicans.* Providence, RI: American Mathematical Society.

Qiu, J., Li, H., Jou, J., Liu, Y., Luo, Y., Feng, T., Wu, Z., & Zhang, Q. (2010). Neural correlates of the "aha" experiences: Evidence from an fMRI study of insight problem solving. *Cortex*, *46*(3), 397–403. http://dx.doi.org/10.1016/j.cortex.2009.06.006

Pizlo, Z., Stefanov, E., Saalweachter, J., Li, Z., Haxhimusa, Y., & Kropatsch, W. G. (2006). Traveling salesman problem: A foveating pyramid model. *The Journal of Problem Solving*, *1*(1), 83–101.

Rumelhart, D. E., McClelland, J. L., & PDP Research Group. (1987). *Parallel distributed processing, Vol.1: Foundations.* Cambridge, MA: MIT Press.

Samuels R. (1998). Evolutionary psychology and the massive *modularity* hypothesis. *British Journal for the Philosophy of Science*, *49*(4), 575–602. http://dx.doi.org/10.1093/bjps/49.4.575

Schooler, J. W., Ohlsson, S., & Brooks, K. (1993). Thoughts beyond words: When language overshadows insight. *Journal of Experimental Psychology: General*, *122*(2), 166–183. http://dx.doi.org/10.1037/0096-3445.122.2.166

Shallice, T. (1982). Specific impairments of planning. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *298*(1089), 199–209. http://dx.doi.org/10.1098/rstb.1982.0082

Sio, U. N., & Ormerod, T. C. (2009). Does incubation enhance problem solving? A meta-analytic review. *Psychological Bulletin*, *135*(1)*,* 94–120. http://dx.doi.org/10.1037/a0014212

Smullyan, R. (1978). *What is the name of this book: The riddle of Dracula and other logical puzzles.* Englewood Cliffs, N.J.: Prentice Hall.

Smullyan, R. (1988). *Forever undecided: A puzzle guide to Gödel.* New York: Oxford University Press.

Sperber, D. (1994). Understanding verbal understanding. In J. Khalfa (Ed.), *What is intelligence*? (pp. 179–198). New York: Cambridge University Press.

Sternberg, R. J., & Davidson, J. E. (Eds.) (1996). *The nature of insight*. Cambridge, MA: MIT Press.

Sun, R. (2002). *Duality of the mind: A bottom up approach toward cognition*. Mahwah, NJ: Lawrence Erlbaum.

Taatgen, N., & Anderson, J. A. (2010). The past, present, and future of cognitive architectures. *Topics in Cognitive Sciences, 2*(4), 693–704. http://dx.doi.org/10.1111/j.1756-8765.2009.01063.x

Tulving, E., Kapur, S., Markowitsch, H. J., Craik, F. I., Habib, R., & Houle, S. (1994). Neuroanatomical correlates of retrieval in episodic memory: Auditory sentence recognition. *Proceedings of the National Academy of Sciences of the United States of America*; *91*, 2012–2015. http://dx.doi.org/10.1073/pnas.91.6.2012

Tulving, E., Markowitsch, H. J., Craik, F. I. M., Habib, R., & Houle, S. (1996). Novelty and familiarity activations in PET studies of memory encoding and retrieval. *Cerebral Cortex*, *6*(1), 71–79.

Van Overschelde, J. P., Rawson, K. A., & Dunlosky, J. (2004). Category norms: An updated and expanded version of the Battig and Montague (1969) norms. *Journal of Memory and Language*, *50*(3), 289–335. http://dx.doi.org/10.1016/j.jml.2003.10.003

van Rooij, I., Haxhimusa, Y., Pizlo, Z., & Gottlob, G. (Editors). Dagstuhl Seminar 11351: Computer Science and Problem Solving: New Foundations. *Dagstuhl Reports*, *1*(8), 2011.

von Neumann, J., & Morgenstern, O. (1958). *The theory of games and economic behavior*. Princeton, NJ: Princeton University Press.

Wallas, G. (1926). *The art of thought.* New York: Harcourt Brace.

Weisberg, R. W. (1992). Metacognition and insight during problem solving: Comment on Metcalfe. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *18*(2), 426–431.

Wertheimer, M. (1954). *Productive thinking*. New York: Harper & Brothers.

Wickelgren, W. A. (1974). *How to solve problems: Elements of a theory of problems and problem solving*. San Francisco: W. H. Freeman.

Wittgenstein, L. (1958). *Philosophical investigations* (2nd ed.; G. E. M. Anscombe, Trans.). Oxford, England: Blackwell.

Yassa, M. A., & Stark C. E. (2008). Multiple signals of recognition memory in the medial temporal lobe. *Hippocampus*, *18*(9), 945–954. http://dx.doi.org/10.1002/hipo.20452

Zobrist, A. (1969). A model for visual organization for the game of GO. *Proceedings of the AFIPS Spring Joint Computer Science Conference*, *34*, 103–112.

Appendix

In this section, we give the solutions and some notes concerning each of the 18 problems in Section 3. The solutions will be very cryptic; however, in the body of the paper we will discuss some of the issues surrounding the problems including the original source when available.

Solution Problem 3.1. Imagine that the original 8″ by 8″ was a chessboard with green and white squares. Observe that the upper right and lower left squares are the same color. However, each domino placement must cover one white and one green colored square; therefore, it is impossible to tile the 62 square array after the upper right and lower left squares are cut off.

Solution Problem 3.2. Suppose alternate cubes are colored green and white, where cubes sharing a face have different colors. Color the upper, front left cube in Figure 1 green. By the coloring scheme, it is clear that the middlemost cube is colored white. Starting with the Green upper, front left cube, it is clear that the worm eats green colored cubes on odd numbers and white ones on even numbers. But the middlemost cube is eaten as the 27th cube and it is white. Thus there can be no sequence that starts at the upper, front left cube and leads to eating the middlemost cube last. It's no help to start at another cube, because by this coloring scheme 13 cubes are colored white and 14 green, so if the worm starts eating a white cube, he runs out of white cubes to eat before the 27th cube.

Solution Problem 3.3. The correct answer is 402 pieces of paper. In books, pieces of paper are numbered on both sides. Furthermore, examining the way the books are stacked in Figure 2, the worm only eats one piece of paper in the first and tenth volumes; where as it eats 50 pieces of paper in the other eight volumes.

Solution Problem 3.4. Let the circumference and radius of the earth be, respectively, C and R. We know that $C = 2\pi R$, so the circumference is directly proportional to the radius; thus $(1/2\pi)'' \approx 0.16''$ is added to the radius, regardless of the original size of C. The thickness of a dime is less that 0.16″, so you could fit a dime under the belt.

Solution Problem 3.5. It is not difficult to see that there are exactly 16 little cubes with both colors.

Solution Problem 3.6. The usual stipulation of the nine-dot problem requires four lines, and the statement of the problem reduced it to three. Figure 6 shows solutions for these two stipulations.

Figure 6. Solutions to the Nine- Dot Problem
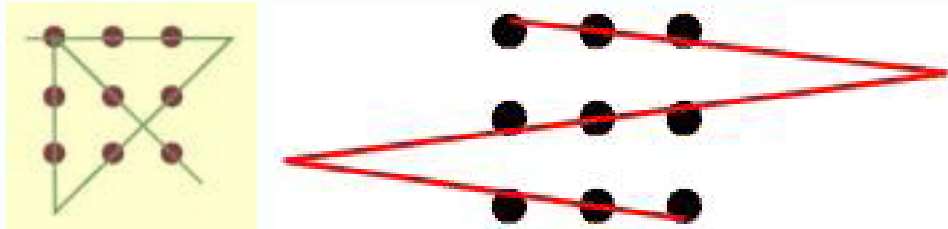
Solution Problem 3.7. Light bulbs in sockets get warm when they are on. Turn one switch on and leave it on. Turn another switch on for a few minutes and then turn it off. Enter the closet without touching the third light switch. The first switch controls the light that is on, the second controls the light that is off but warm, and the untouched switch controls the light that is off and not warm.

Solution Problem 3.8. A reasonable completion is obtained by noting that the capitol letters with straight lines are on the top, and the letters with a curvy feature are on the bottom.

Solution Problem 3.9. The solution follows from the logic of the solution to Problem 3.10 next, namely that any volume of cream that ends up in the coffee is exactly matched by a displacement equal to the amount of coffee in the cream mixture. Thus both beakers are equally diluted.

Solution Problem 3.10. Clearly, every blue bead that was transferred into the red bead jar matches with a red bead that remains in the blue bead jar. So in all cases, there are as many red beads in the blue jar as there are blue beads in the red jar.

Solution Problem 3.11. The trains collide in one hour, and the SUPERFLY has been going 100 mph, therefore he has traveled 100 miles in his zig zag journey.

Solution Problem 3.12. There will be exactly one place on the trail where George was on both days at the same time of day. To see this, imagine that a 'shadow' of George proceeds up the trail on Sunday morning exactly as George did the previous day. Clearly the shadow will pass by George on his descent somewhere on the trail, and their watches will read the same time.

Solution Problem 3.13. Trip average speed is distance divided by travel time. The distance is 40 miles, and the up travel of 20 miles at 30 mph has taken 40 minutes (two thirds of an hour). So to average 60 MPH one would have to go infinitely fast downhill, and of course that is impossible.

Solution Problem 3.14. The product of the ages is 36 implies that the ages are three positive integers that multiply to 36. Table 1 lists them along with their sum. It is clear that

knowing the sum determines which row one is in unless the sum is 13. Thus the house number is 13, else the census taker would already know the ages. Since the oldest taught the younger ones to play chess, the ages must be 2, 2, and 9 because if it were 1, 6, and 6 there would not be two youngest.

Table 1. Possible Ages of the Three Daughters in Problem 3.14.

D11	D2	D3	SUM
1	1	36	38
1	2	18	21
1	3	12	16
1	4	9	14
1	6	6	13
2	2	9	13
2	3	6	11
3	3	4	10

Solution Problem 3.15. The first man is not seeing two red hats because if he did he would know his hat was white. The second man knows this, so if he saw a red hat on the third man, he would know his hat was white. Therefore, because neither guessed the color of their hat, the third man knows his hat is white.

Solution Problem 3.16. The estate rules were ill formed, because ½ + 1/3 + 1/9 = 17/18 rather than one.

Solution Problem 3.17. The following statement will force the wizard to give you a trillion dollars. "You will either give me $10 or you will give me a trillion dollars." If the statement is false, he must give you $10 but that would make the statement true. If he gives you any amount other than $10 or a trillion dollars, the statement is false and he would have to give you $10. Ergo, he must give you a trillion dollars.

Solution Problem 3.18. Of course there are a bunch of 'uninteresting' reasons why it cannot be done in either place. The key to an interesting one is to characterize a perfect map as one that depicts everything. Then the map of England drawn in London would have to depict the mapmaker on the map drawing a map that has the mapmaker on it drawing . . . ad infinitum. Drawing one on NYC would not have that problem.