1 The shape of problems

Philip N. Johnson-Laird

Tutta la teoria del problem-solving, da Simon ai giorni nostri, può venire considerata come uno sviluppo diretto o indiretto dei lavori dei gestaltisti. [All theories of problem solving from Simon to the present day can be considered as a development, direct or indirect, of the work of the Gestalt psychologists.]
(Paolo Legrenzi, 1978, p. 173)

Introduction

Psychology in Italy during the twentieth century was unique. Vittorio Benussi (1878–1927) initiated an Italian variety of Gestalt psychology, deriving more from Meinong than the great German school; Cesare Musatti (1899–1979) continued the tradition; and Gaetano Kanizsa (1913–1993) brought it to fruition in his startling and original demonstrations of perceptual phenomena, especially illusory contours. Paolo Legrenzi is a crucial figure in this history, because he built the bridge from Gestalt theory in his collaborations with Kanizsa to information-processing psychology (see, e.g., Kanizsa & Legrenzi, 1978; Kanizsa, Legrenzi, & Sonino, 1983). As the epigraph to the present chapter shows, he has never lost sight of the Gestalt origins in the study of thinking. For over 30 years, he and I have carried out research together, usually in collaboration with his wife Maria Sonino, and mostly on deductive reasoning. He is thus both a colleague and a dear friend. In this chapter, my aim is to honour him with an analysis of human problem solving, a topic that goes back to some of his own studies (e.g., Legrenzi, 1994) and to those of his Gestalt forebears (e.g., Duncker, 1945; Katona, 1940; Köhler, 1925; Luchins, 1942; Maier, 1931; Wertheimer, 1945/1959).

A common phenomenon when you are struggling to solve a problem is that you have a burst of inspiration – an *insight* – and the solution suddenly emerges into your consciousness, seemingly from nowhere. "Eureka," you say, like Archimedes: "I have it". Though you do not normally leap from your bath and run stark naked through the streets, shouting the solution to the world. The experience seems very different from the normal process of thought in which you advance step by step towards the solution of a problem, as, say, when you work out your income tax returns. Indeed, the Gestalt psychologists assumed that insight was special. It is a creative process, they argued, that depends on a sudden restructuring (*Umstruktuierung*) or recentring (*Umzentrierung*) of the perceptual field (see Kohler, 1929/1947; Wertheimer, 1945/1959). But critics both early (Bulbrook, 1932) and late (e.g., Perkins, 1981; Weisberg, 1986), have argued that there is nothing special about insight. Normal thinking is analogous to a search for the number that opens a combination lock. When at length you stumble on to the solution, it is bound to be sudden, but it does not depend on any processes that differ from those that hitherto failed.

An empirical discovery about problem solving contrasts with such claims. Janet Metcalfe and her colleagues asked participants in an experiment to rate how "warm" they were in their efforts to solve problems (see, e.g., Metcalfe & Weibe, 1987). With some problems, the participants made progress towards the solution, as shown in the increase in their ratings of warmth. But, with other problems – those typically used in studies of insight, there was no increase in ratings of warmth: Individuals felt remote from the solution right up to the moment that it popped into their heads. When the ratings of warmth did grow incrementally for insight problems, they tended to presage a failure to solve the problem. Hence, some problems are solved in a sudden and unexpected way. These solutions, we shall say, depend on *insight*. The question is what are its underlying mental processes?

The short answer is: no one knows. There are, however, three main schools of thought. The first view is the Gestalt account of insight as dependent on restructuring or recentring. The trouble is that these notions are obscure (Ohlsson, 1984a, 1984b): no computer model implementing them exists.

The second school of thought is that insight depends on overcoming an erroneous approach to a problem (see, e.g., Isaak & Just, 1995; Knoblich, Ohlsson, Haider, & Rhenius, 1999; Weisberg, 1993). Individuals become fixated on inappropriate methods based on their prior experience, but a period of "incubation" in which they think about other matters allows the misleading cues to become less accessible with a consequent greater chance of recovering the correct cues (Smith & Blankenship, 1991). Such perseverations certainly occur, e.g., designers tend to conform to any initial example that they are given, incorporating its features into their finished designs even when they are asked not to (Jansson & Smith, 1991). The trouble with this approach, however, is its assumption that individuals know the right methods, though cannot access them. The claim may be correct for some problems (see Keane, 1989), but it is false for others. Arguments to the contrary resemble Socrates' leading questions to a slave boy to show that he knew the solution to a geometric problem (see Plato's *Meno*). In fact, a major theme of the present chapter is that attempts to solve a problem can lead to the creative discovery of knowledge.

The third school of thought is based on Newell and Simon's (1972) concept of a "problem space", which is defined as all possible sequences of the mental operations pertinent to the solution of a problem. The problem solver needs to search this space for a sequence of operations leading from the initial state of the problem to its goal. Kaplan and Simon (1990) extended this idea to insight. They argued that it depends on switching from a search in the problem space to a search in the meta-level space of *possible* problem spaces for a new representation of the problem (see also Korf, 1980). In a case of insight, this new representation yields the solution to the problem. For example, most people are at first defeated by the problem of the "mutilated" chessboard. The goal is to cover a chessboard with dominoes, or else to prove that the task is impossible. Each domino is in the shape of a rectangle that covers two adjacent squares on the chessboard. However, the board has been mutilated by the removal of a square at one corner and another square at the diagonally opposite corner. Individuals who imagine laying out dominoes on the board get nowhere. They have the wrong representation of the problem. The key insight depends on constructing a new representation, which concerns parity: Each domino covers one white square and one black square, but the diagonally opposite corners are of the same color. Hence, after they are removed, the remainder do not include an equal number of black and white squares, and so the task is impossible.

Tom Ormerod and his colleagues have formulated an alternative account of insight (see, e.g., Chronicle, Ormerod, & MacGregor, 2001; MacGregor, Ormerod, & Chronicle, 2001; Ormerod, MacGregor, & Chronicle, 2002). They argue that when individuals tackle a problem, they select moves to maximise progress towards a hypothesised goal but to minimise the expansion of the problem space. They relax this constraint only if they have to, but in this way they may discover a new sort of move. This theory postulates that problem solvers can assess whether a putative move does make progress towards a goal, and so they can rely on a "hill-climbing" strategy. The assumption is plausible for certain problems, but perhaps less plausible for others, such as those problems in daily life that call for insight (or those discussed towards the end of this chapter).

The problem with "problem spaces" is that they are irrefutable. Any computable process can be described as a search through a space of possible sequences of operations. At the limit, of course, there is no search at all, because the process follows a deterministic sequence of operations. This limit, however, is merely a special case. To characterise insight as a search through *possible* problem spaces is highly abstract. As Kaplan and Simon remark (1990, p. 381): "The space of possible problem spaces is exceedingly ill-defined, in fact, infinite." What we need is an account of how the mind carries out the search.

In sum, all three schools of thought contain elements of the truth, but none is wholly satisfactory. The present chapter accordingly outlines a new theory of creative problem solving. It begins with an analysis of creativity and an exhaustive set of possible creative strategies. It describes a test bed for the new theory – a domain of "shape" problems – and a process of creative discovery depending on explorations of possible operations in tackling these problems. This process enables individuals to develop strategies for problem solving. It can also lead to insights that overcome the failure of an initial strategy. Finally, the chapter draws some conclusions about the nature of problem solving.

The nature of creativity

Insight is creative, and so to understand it we need to understand how the mind creates. But what is creativity? Science does not advance by a priori definitions, but elsewhere I have offered a working definition of creativity (see, e.g., Johnson-Laird, 2002). This definition depends on five main assumptions:

(1) *Novelty*: the result of a creative process is novel for the individual who carries out the process. Creativity is not mere imitation or regurgitation.
(2) *Optional* novelty for society: the result may also be novel for society as a whole, but this requirement is optional. The mental processes underlying creativity are the same even if unbeknownst to the relevant individual someone else has already had the same idea.
(3) *Nondeterminism*: creativity depends on more than mere calculation or the execution of a deterministic process, such as long multiplication. When you create something, such as the solution to a problem, alternative possibilities can occur at various points in the process. If you could relive your experience with no knowledge of your first effort, then you might make different choices the second time around. Hence, creation does not unwind like clockwork with only one option for you at each point in the process. In computational theory, a machine with this property is known as *nondeterministic*. It can yield different outcomes when it is in the same internal state, and has the same input if any (see, e.g., Hopcroft & Ullman, 1979). No one knows whether human creativity is truly nondeterministic, but at present we have to make this assumption. It allows for our ignorance and, computationally speaking, it costs us nothing. In principle, anything that can be computed nondeterministically can also be computed deterministically.
(4) *Constraints*: creativity satisfies pre-existing constraints or criteria. For a problem in science, art, or daily life, a crucial constraint is that the solution works, i.e., it is viable. There are usually many other constraints on the operations, physical or mental, that are pertinent to the problem. The aesthetic values of a genre constrain the creation of works of art; a knowledge of robust results constrains the creation of scientific theories; and an awareness of practical realities constrains the creation of solutions to everyday problems. The individual creator is not a closed system,

but is influenced by mentors, collaborators, and leaders (Simonton, 1984). In this way, the values of a culture influence individuals' creative processes, which, in turn, may contribute to the values that are passed on to the next generation.

(5) *Existing* elements: you cannot create out of nothing. There must be existing elements to provide the raw materials for even highly original works of art or science. These existing elements include, of course, the system of mental processes that allow creativity to occur.

Novelty, Optional novelty for society, Nondeterminism, Constraints, and Existing elements are the five components of the NONCE definition of creativity. The definition has an unexpected consequence. On the assumption that creativity is a computable process, there can be only three sorts of creative strategy. A *strategy* is a systematic sequence of elementary steps that an individual follows in solving a problem, and is therefore similar to a computer algorithm for solving a problem.

The first sort of strategy is *neo-Darwinian*. It is analogous to the evolution of species according to the neo-Darwinian synthesis of genetics and natural selection. The strategy has two stages: a generative stage in which ideas are formed by an entirely arbitrary or nondeterministic process working on existing elements; an evaluative stage that uses constraints to pass just those results that are viable (see Figure 1.1, p. 8). Whatever survives evaluation, which may be little or nothing, can serve as the input to the generative stage again. The process can thus be repeated ad libitum with the output from one iteration serving as the input to the next. Neo-Darwinist theories of creativity include accounts based on trial and error, and they have often been proposed by psychologists (e.g., Bateson, 1979; Campbell, 1960; Simonton, 1995; Skinner, 1953). An individual produces a variety of responses, and the contingencies of reinforcement, or some other constraints, select those that are viable and eliminate the remainder. It is crucial to distinguish between a single operation of a neo-Darwinian strategy and its iterative or repeated use, which is much more powerful (Dawkins, 1976). Evolution is thus an archetypal *recursive* process: It applies to its own successful results. It is mimicked in the "genetic algorithms" developed by Holland and his colleagues for finding optimal solutions to problems (e.g., Holland, Holyoak, Nisbett, & Thagard, 1986).

A neo-Darwinian strategy is grossly inefficient, but the only feasible strategy if the generative process cannot be guided by constraints, as in the case of the evolution of species (see Mayr, 1982, p. 537). Yet, if constraints are used in the evaluation of ideas, then they could be used instead to constrain their generation in the first place. Unlike species, ideas could evolve using this process. The second sort of strategy is *neo-Lamarckian* in just this way (see Figure 1.1, p. 8). All the constraints acquired from experience govern the generative stage – by analogy with Larmarck's theory of evolution (see Mayr, 1982, p. 354). If an individual has acquired a comprehensive set of

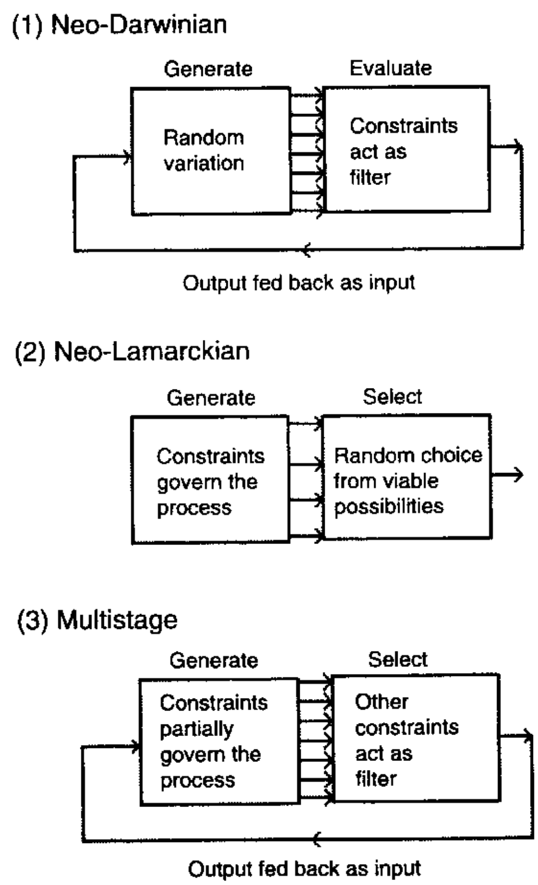


Figure 1.1 The three strategies for creativity

constraints guaranteeing the viability of the results, then the generative stage will yield a small number of possibilities, all of which meet the constraints. But because all the constraints are used to generate a result, by definition none are left over for its evaluation. The constraints will sometimes allow more than one possibility, and so the only way to choose amongst them must be nondeterministic, e.g., by making an arbitrary decision. The strategy has just two stages with no need for recursion: (1) the generation of possibilities according to constraints; (2) an arbitrary selection, where necessary, from amongst them. The strategy is highly efficient, because it never yields hopeless results and never calls for recursion.

The third sort of strategy is a compromise. It is *multistage*: The generative stage uses some constraints like a neo-Lamarckian strategy, and the evaluative stage uses some constraints like a neo-Darwinian strategy (see Figure 1.1). The initial generation of possibilities under the guidance of some constraints may leave something to be desired, and so the individual applies further constraints to evaluate the results. They may need further work, and so the process can be recursive.

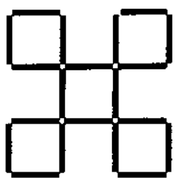
Figure 1.1 summarises the three strategies. Many creative individuals do indeed repeatedly revise the results of their earlier efforts. Since they are applying constraints at each stage, why do they not apply all of these constraints immediately in the first generative stage? Why the need for a time-consuming division of labour over several recursions? From a computational standpoint, it would be more efficient to apply all the constraints in the generative stage in a neo-Lamarckian way. It seems paradoxical for individuals to waste time making an inadequate attempt if they have the ability to perceive its inadequacy and to set matters right. The resolution of the paradox may lie in the way the mind works. Knowledge for generating ideas is largely unconscious, whereas knowledge for evaluating ideas is largely conscious and embodied in beliefs. This dissociation resolves another puzzle, which Perkins (1981, p. 128) refers to as the fundamental paradox of creativity: People are better critics than creators. Criticism can be based on conscious knowledge that can be acquired easily; whereas the generation of ideas is based on unconscious knowledge acquired only by laborious practice in creating. Hence there are two stages in many sorts of creation: a generative stage and an evaluative stage. Hence the greater ease of criticism over imagination. In creation, no substitute exists for a period of apprenticeship. You learn by emulating successful creators, and by trying to create for yourself in a particular domain. Only in this way can you acquire the tacit constraints for creativity. If this account is correct, then it is a blow to the purveyors of universal nostrums for creativity. There can be no effective recipes for enhancing your creativity across all domains. Creation is local to a particular domain of expertise.

Problems of shape

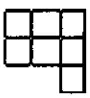
An important distinction in the problems that arise in daily life and in the psychological laboratory is between what I shall refer to as *one-off* problems and *series* problems. One-off problems are those that occur just once. If you have solved a one-off problem, then provided you have not forgotten the solution you can apply it at once to any other instance of the problem. A classic example of a one-off problem is Duncker's (1945) candle puzzle. You are given a candle and a box of thumb tacks (drawing pins), and your task is to fix the candle to the wall. Most people are stumped, unless and until they have the insight that the box containing the thumb tacks can itself be pinned to the wall and thereby support the candle. Once they have had this insight, they retain it, and so the solution to any other instance of the problem is trivial. The solution to other insight problems may not be so easy to retain. Tom Ormerod et al. (2002) argue that retention depends on how easy it is to recode the solution as a single step or a single idea. One corollary of this analysis is that there is no such thing as an "insight" problem per se, i.e., the quest for a decisive categorization is chimerical (pace Weisberg, 1996). All that can exist are problems that call for most *naive* individuals to have an insight if they are to solve them.

In contrast to one-off problems, water jug problems come in a series of different forms, e.g., using just three jugs, measuring respectively 127, 21 and 3 quarts, your task is to measure a quantity of 100 quarts. Infinitely many water jug problems can be posed; they do not all have the same sort of solution and the acquisition of one sort of solution can inhibit the use of another (see, e.g., Luchins, 1942; Luchins & Luchins, 1950). The distinction between one-off and series problems can also be drawn for problems in mathematics and science. The theoretical problem of relating mass, force, and acceleration was solved by Newton, and later revised by Einstein. Their solutions can be exploited to solve a series of practical problems in mechanics.

The study of one-off problems in the psychological laboratory is intriguing, but it has methodological problems. As Tom Ormerod and I discovered to our cost (in an unpublished study), it is difficult to distinguish between the knowledge that individuals bring to a problem and the knowledge that they discover in trying to solve it. Likewise, it is difficult to observe the development of their strategies for coping with the problem. The main experimental procedure is to give hints of one sort or another in order to trigger the knowledge requisite for the solution. This manipulation, however, tells one little about an individual's strategy. We need to understand how people develop both knowledge and strategies, and the task calls for a series of problems. This section of the chapter accordingly examines a series of *shape* problems in which individuals have to make or to modify a shape by the addition, removal, or rearrangement of its component pieces (see, e.g., Katona, 1940). Here is an example:

> Given the following arrangement of five squares:
>
> 
>
> in how many ways can you move three pieces to make exactly seven squares of the same size as the originals and with no loose ends, i.e., additional pieces left over?[1]

1 One solution is:



There are seven other ways to solve the problem, i.e., a total of eight.

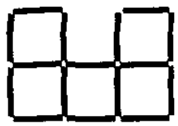
Infinitely many shape problems can be posed, but there are three main sorts:

- problems calling for the removal of a certain number of pieces
- problems calling for the addition of a certain number of pieces
- problems calling for the rearrangement of a certain number of pieces.

Most shape problems are simple, yet they call for sufficient thought to be a staple of puzzle books (see, e.g., Orleans & Orleans, 1983). Some shape problems, however, are sufficiently difficult that they defeat most naive individuals, where "naive" here means an individual who has not had any relevant experience in solving these problems.

Shape problems illustrate most of the features of human problem solving, and in what follows, the chapter uses them to illustrate a theory of human problem solving. We begin with the simplest possible shape problem:

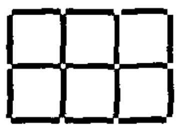
(1) Given the following five squares, add one piece to make six squares:



The problem is trivial. But, why? There are at least two reasons. First, the symmetry of the figure ensures that the missing piece "pops out" of the display. This missing piece is highly salient. Human perceivers can subitize these squares, i.e., they can apprehend them in a single glance and see that a piece is missing. This ability depends on parallel processing. In contrast, a serial computer program can merely convolve a function for identifying a square (typically a template) with each cell in the visual array, and thereby reveal the presence of five squares and the three sides of a square with a missing piece. The process does not yield an outcome akin to the "pop out" phenomenology of human perception. Second, the addition of only a single piece to a shape makes a neo-Darwinian strategy feasible. You add a piece arbitrarily here or there, and sooner rather than later you will get the right answer.

For a naive participant, who has not seen the previous problem, the following problem is slightly harder:

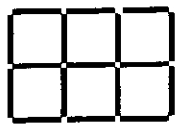
(2) Given the following six squares, remove one piece to leave exactly five squares with no loose ends, i.e., pieces that are not connected to any other piece at one or other end:



In this case, the solution does not pop out, and naive individuals are likely to try one or two experimental moves at least in their mind's eye using a neo-Darwinian strategy.

The following problem is more interesting:

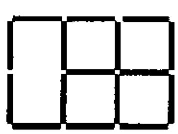
> (3) Given the following six squares, remove five pieces to leave exactly three squares with no loose ends:



You are invited to tackle this problem for yourself before you read on. You are welcome to use matches, tooth picks, or pen and pencil as an aid. A neo-Darwinian strategy can solve this problem, but it is laborious: There are over 6000 different ways to remove five pieces, and only ten of them yield a solution to the problem. A computer program that I have written to model the neo-Darwinian strategy churns away until by chance it derives a solution. The generative stage removes a piece at random, and the evaluative stage checks the result according to a single constraint: Does it solve the problem? Human solvers almost certainly lack the patience and memory to tackle the problem in such a stupid way. They either know what to do in order to solve it (they may be wrong) or they are lost. In the latter case, depending on their motivation, they flail around for some time. In unpublished studies, Louis Lee and I have indeed observed this process of flailing around when individuals first tackle shape problems (see Lee & Johnson-Laird, 2003). They try out seemingly arbitrary choices, which they undo when these moves fail to yield the solution.

Suppose that after you have removed a single piece, you have some way to assess whether your move is a good one moving you towards the solution or a bad one that is not a step towards the solution. With this ability, you could use a neo-Darwinian strategy. You remove a piece at random, and evaluate whether or not your move is a good one. If it is, you retain it, and then return to the generative stage, which removes a second piece at random, and so on. But, if your first move is not good, you abandon it, and return to the generative stage to choose a different first move. You proceed in this way until you have solved the problem. The strategy is neo-Darwinian, but it is tolerably efficient, because you make progress towards the solution move by move, i.e., you are hill climbing (cf. Ormerod et al., 2002). The crux is accordingly whether you can evaluate moves as good or bad without having to solve the problem as a whole.

Let us suppose that you remove a piece at random using a neo-Darwinian strategy:

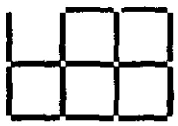


This move transforms problem (3) into a new one: You now have to remove four pieces from the preceding array to leave exactly three squares with no loose ends. So, is your move a good one or not? You are unlikely to have any way to answer this question without carrying out some further moves and assessing their consequences. Hence, you cannot use a neo-Darwinian strategy based on an evaluative stage using only success as one constraint and goodness of move as another. You must develop more powerful constraints or a more powerful strategy. But how? The next section of the chapter aims to answer this question.

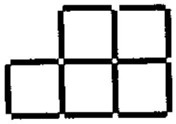
The discovery system

The way people discover how to solve shape problems depends on three mental components that they bring to the task. First, they understand the language in which these problems are posed, and their knowledge includes a grasp of the concepts of *squares* and *pieces*, of *numbers* both cardinal and ordinal, and of the pertinent arithmetical operations of counting, adding, and subtracting. Intelligent adults in some other cultures do not have these concepts, and so they cannot understand, let alone solve, shape problems. Second, people are equipped with the mechanisms of perception and action. They can see squares, and they can envisage the consequences of removing a piece. These mechanisms are unconscious: Introspection does not tell you how you see at once that a piece is missing in problem (1). Third, people are equipped with a *system for discovery*. It enables them both to acquire knowledge about problems from their attempts to solve them, and to develop a strategy for coping with them. Such a strategy exploits the knowledge that they acquire and that can constrain the process of solution. As we will see, the discovery system stores the effects of different sorts of move. It can also shift such a constraint from the evaluative stage of solving a problem to the generative stage.

The process of discovery occurs as individuals explore the effects of different moves in trying to solve a problem. Consider, again, problem (3) in which the task is to remove five pieces from an array to eliminate three squares. You are likely to tackle this problem by removing a piece and assessing the consequences. Suppose you remove a piece from a corner of the shape, then the result is:

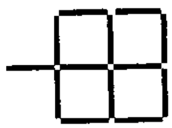
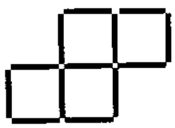
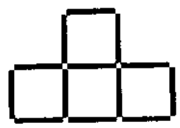


It is now necessary to remove the remaining "loose end" to yield the following shape. You have discovered something: One way to eliminate a square is to

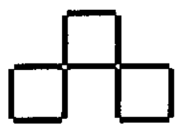


remove the two pieces making up a *corner*. A corner is indeed one of seven sorts of configuration in a shape problem (see Figure 1.2, below).

A further result is that you now have a new problem: to remove three pieces from the preceding shape to eliminate two squares. If you remove another corner, then discounting rotations there are three ways in which you could do so. In each case, your remaining problem is to remove one piece to eliminate one square:

(a)  (b)  (c) 

In case (a), you are obliged to remove the loose end, and so you fail to solve the problem. In case (b), you also fail, because here the only way to eliminate a square as a result of removing a single piece is to remove a piece that is common to two squares, thereby eliminating both of them. You have discovered something new: Removing one *middle* piece, i.e., a piece in the middle of two squares, eliminates both the squares. In case (c), you can succeed. You can remove one piece to eliminate one square, as shown here:



Once more, you have made a discovery: You can eliminate a square by removing an *outer* piece, i.e., a piece in the square that has middle pieces from the same square at both ends, but that is not itself a middle. There are, as we will see presently, quite different solutions to problem (3).

What this analysis has illustrated is how the exploration of possible moves leads to a tactical knowledge of the different sorts of pieces in shape problems, and of the different effects of removing them. As Figure 1.2 shows, there are seven sorts of piece in shape problems. These tactics, in turn, provide constraints on the creative process. When they are first acquired, they are constraints that are discovered in the evaluative stage of trying to create a

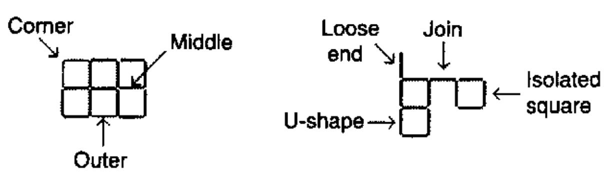


Figure 1.2 Illustrations of the seven sorts of piece in shape problems

solution. A major change occurs, however, when such knowledge is shifted to the generative stage of the process. You no longer have to try out moves in an exploratory way, but, like the participants whom Louis Lee and I have tested experimentally, you can go directly to the solution of at least certain problems with no false moves (Lee & Johnson-Laird, 2003).

The discovery system operates by storing the effects of operations on shapes. The order in which individuals develop a knowledge of the different tactical steps, and the completeness of their knowledge, is likely to differ depending on each individual's experience. Table 1.1 summarises the complete set of tactical steps for shape problems of the present sort.

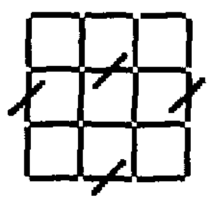
Table 1.1 The seven tactical steps for a set of shape problems (see Figure 1.2 for illustrations of the different sorts of piece)

1. To remove 1 piece and 0 squares, remove a loose end (obligatory).
2. To remove 1 piece and 0 squares, remove a join.
3. To remove 1 piece and 1 square, remove an outer.
4. To remove 1 piece and 2 squares, remove a middle.
5. To remove 2 pieces and 1 square, remove a corner.
6. To remove 3 pieces and 1 square, remove a U-shape.
7. To remove 4 pieces and 1 square, remove an isolated square.

Does the use of visual imagery help you with shape problems? Perhaps. But, it has its limitations. Consider the following problem:

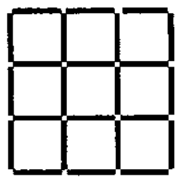
> I take a bunch of thousands of very thin needles, and throw them up into the air, imparting different arbitrary accelerations to each of them. They fall to the ground. As they fall, however, I freeze them in mid-air. Are there likely to be more needles nearly horizontal than nearly vertical, more needles nearly vertical than nearly horizontal, or are the two proportions likely to be roughly equal?

Many people say that the two proportions should be roughly equal. They are wrong. Visual images tend to be two-dimensional. An accurate three-dimensional representation shows that there are many more ways in which a needle can be nearly horizontal than it can be nearly vertical. A horizontal needle can be pointing in any of the 360° compass directions, but a vertical needle must point, in effect, only due North (or due South). Imagery is also difficult in the case of the following figure. You have to imagine that each of the lines that is crossed out has been removed from the figure:

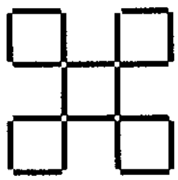


One source of the difficulty is that looking at the figure interferes with your ability to imagine its appearance with the crossed out lines missing. The resulting figure is shown at the end of this section of the chapter.

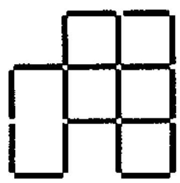
Does symmetry influence the sequence of tactics that problem solvers adopt? Consider the problem of removing four pieces from the following shape in order to eliminate four squares:



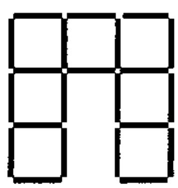
An obvious solution is to remove the symmetrical outer pieces:



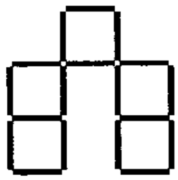
A much less obvious solution is an asymmetrical one:



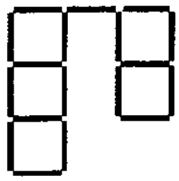
A sequence of the same tactical moves preserves symmetry, and symmetry suggests the use of such a sequence. Indeed, individuals can carry out symmetrical operations in parallel rather than one after the other (pace Newell, 1990). Suppose you are given the following initial shape, and your task is to remove four pieces to eliminate two squares:



You can envisage removing the two top corners simultaneously:



You are quite unlikely to reach the following solution:



Symmetry is important in solving shape problems. As our unpublished studies show, symmetry makes certain moves salient. If they are indeed part of the solution to a shape problem, the problem is simpler to solve than if they are not part of the solution (Lee & Johnson-Laird, 2003).

The development of strategies

Strategies are the molar units of problem solving, *tactics* are the molecular units, and their underlying perceptual and action *mechanisms* are the atomic units. Tactics alone cannot enable you to solve a complex problem. You have to discover how to combine them into useful sequences that make up a strategy. Van der Henst, Yang, & Johnson-Laird (2002) have investigated how individuals develop strategies in order to make deductive inferences. Our results corroborated the following principle of *strategic assembly*: Naive individuals assemble strategies bottom up as they explore problems using their existing tactics. Once developed, a strategy can control thinking in a meta-cognitive way, i.e., individuals can consciously apply it top down (Van der Henst et al., 2002).

The same principle appears to apply to the development of strategies to solve problems. Hence, in some domains, different individuals are likely to develop different strategies. Likewise, different sorts of problem should inculcate different sorts of strategy. Shape problems, however, are sufficiently simple that the intelligent adults usually converge on a fairly comprehensive knowledge of tactics, which they can use to solve shape problems with few false moves (Lee & Johnson-Laird, 2003).

Instruction has an effect on strategies too. Katona (1940) demonstrated this phenomenon. He tested two groups of participants with shape problems. One group was shown the solution to a few problems, and another group was given a lesson about the relations between numbers of pieces and numbers of squares – a lesson, in effect, about tactics. Only this second group showed a positive transfer to new problems. Yet, for shape problems, individuals are unlikely to develop beyond a multistage strategy in which they use some tactical constraints to generate ideas and other constraints, such as the goal of the problem, to evaluate the outcomes. A deep one-off problem is whether there is a neo-Lamarckian strategy for shape problems, that is, a strategy that solves any shape problem whatsoever without ever making a false move.

Readers familiar with the seminal work of Newell and Simon (1972) may wonder whether a means–ends strategy, in which individuals work backwards from the desired goal, is feasible for shape problems. The answer is: no. The goal state in shape problems is deliberately imprecise. It specifies how many squares should remain, but not how they are arranged. A precise goal with such an arrangement would amount to a solution of the problem. Granted an imprecise goal, it is impossible to use a means–ends strategy. Feasible strategies for shape problems call for a certain amount of working forwards from the initial shape.

An initial strategy that intelligent adults seem to adopt is to work forwards in the way that I sketched in the previous section. As they discover tactical constraints, they adopt them to constrain the possibilities that they envisage. For example, a person whose immediate problem is to remove one piece and thereby eliminate two squares will look at once to see whether the shape contains a *middle* piece. Likewise, a person who has to remove four pieces to eliminate two squares will look to see whether the shape contains two *corners*. A more systematic strategy uses the following sorts of constraint:

- If the number of pieces to be removed is greater than the number of squares to be eliminated, then remove an isolated square, a U-shape, or a corner.
- If the number of pieces to be removed is greater than the number of squares to be eliminated and no further squares are to be eliminated, then remove a loose end or a join.
- If the number of pieces to be removed is smaller than the number of squares to be eliminated, then remove a middle.
- If the number of pieces to be removed is equal to the number of squares to be eliminated, then remove an outer.

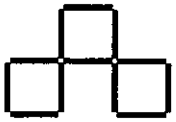
Naive reasoners often begin by focusing either on the number of matches to be removed or else the number of squares to be eliminated, and only later realise the need to think about both these numbers at each step.

A more advanced strategy aims to decompose complex problems into tactical moves. It starts by making an inventory of the different sorts of piece in the shape. The shape may contain, for example, four corners (two pieces each), seven middles, and two outers. If the goal is to remove five pieces in order to eliminate three squares, then the solution must satisfy two equations:

5 pieces to be removed = $n_1$2.corners + $n_2$middle + $n_3$outer

3 squares to be eliminated = $n_1 + 2n_2 + n_3$

Two simultaneous equations containing three unknowns are ill-formed, i.e., they can have more than one integral solution. Yet, they do constrain the set of possible solutions. One solution, for instance, is $n_1 = 2$, $n_2 = 0$, and $n_3 = 1$. The final stage is to try to construct such a solution. In fact, for problem 3, it yields the familiar answer:

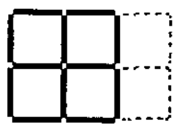


Two difficulties exist for the successful use of this strategy. The first is that naive individuals have to discover how to solve simultaneous ill-defined equations. An insight that yields a solution to this one-off problem is to treat it as an "occupancy" problem, i.e., to compute the different ways in which P pieces can be put into N cells, where P is the number of pieces to be removed, and N is the number of squares to be removed. This step is likely to be beyond many individuals. The second difficulty is that the inventory of a problem can change, not merely quantitatively but also qualitatively, as each piece is removed from a shape. Hence, at each removal step, it is prudent to categorise the remaining pieces in the resulting shape all over again.

One feasible strategy, which I owe to my colleague Uri Hasson, is that individuals think at the level of squares, and so they treat problem (3) as calling for the removal of three squares. There are six squares in the initial shape, i.e., four squares at the corners and a vertical row of two squares in the middle. Hence, individuals think about the different ways of removing individual squares: If they remove three corner squares, they remove six pieces, which fails to solve the problem; if they remove two corner squares and one square in the middle, then they can solve the problem; and if they remove both squares in the middle and one corner square, they can also solve the problem. The details of the solutions, however, call for some careful thought – working forwards from the initial shape – about which particular pieces to remove. We have observed a similar phenomenon: Individuals soon realise that there are higher order configurations than those shown in Figure 1.2. They learn, for example, to remove an E-shape (five matches making up a corner and a U-shape) in a single move (Lee & Johnson-Laird, 2003).

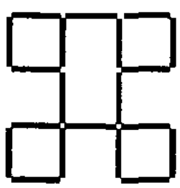
Another sort of strategy depends on reformulating problems. Given problem (3) to remove five pieces and thereby eliminate three squares, a solver can reformulate the problem taking into account the initial shape: Use 12 pieces to make three squares. The tactical knowledge that each isolated square requires four pieces immediately yields the goal of constructing three isolated squares from the initial configuration. This goal in turn yields the familiar solution shown in the preceding diagram.

Still another strategy can be based on a different reformulation of the problem: Remove any five arbitrary pieces, and then move any of the remaining pieces into positions that were previously occupied but that make up three squares. The result of removing the five pieces might be as follows, where the dotted lines show the pieces that were previously in the shape:



It is obvious that moving one or two pieces cannot solve the problem. A move of three pieces, however, can yield the familiar solution illustrated above. A corollary of these observations about strategies is that the verbal framing of a problem may bias individuals to adopt one strategy rather than another (see Van der Henst et al., 2002).

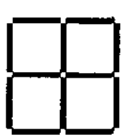
A final note on imagery. The result of removing the lines crossed out in the example earlier in this section is the following shape:



Insight

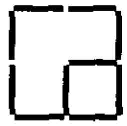
The previous sections of this chapter have explained how an exploration of problems yields a knowledge of tactics and of ways of sequencing them into problem-solving strategies. The human discovery system appears to lay down a record of the outcome of various operations on problems, and the knowledge cumulates as a result of experience. But what about sudden insights? Do they have no role to play in solving shape problems? Let us consider three further problems:

> (4) Given the following shape, remove two pieces to leave two squares with no loose ends (see Levine, 1994, p. 91; Perkins, 2000, p. 111, for versions of this problem):



At first sight, the problem may seem to be impossible, but of course it isn't.

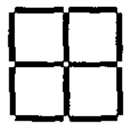
> (5) Given the following shape, add two pieces to make five squares, where a square cannot have any pieces inside it:



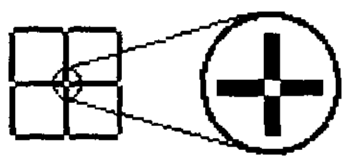
> (6) There are exactly four squares of the same size in a shape. Remove one piece to eliminate all of them with no loose ends.

Before you read on, you should make every effort to solve each of these problems. To solve each of the three problems, most readers are likely to need an insight. The insight for problem (4) is that squares can be of different sizes. This point is normally obvious, but assiduous readers who have tackled all the problems in the chapter may by now have made the tacit assumption that all squares in the solution to a shape problem must be the same size. The solution to problem (4) is indeed the same shape as the starting shape for problem (5): one small square is contained within another larger square. The insight required to solve this problem accordingly calls for you to relax the constraint that all squares are the same size.

Problem (5) is much harder. It depends on a deeper insight. The solution is the following shape:



You may object that there are only four squares here, or you may think instead that the fifth square is the large one containing the four smaller squares. This large square, however, is disqualified because the problem specifies that a square cannot contain any pieces inside it. But look closely at the centre of the shape, which is enlarged here:



The separate ends of the four pieces meeting at the centre form a tiny square, which is the required fifth one. When I have revealed this solution to people who have given up on the problem, they usually give me the sort of old-fashioned look reserved for perpetrators of a bad pun. Yet, the solution is a genuine one, and it depends on an insight. You have to relax the constraint that only the sides of pieces can be parts of squares: Ends of pieces can make squares too.

At this point, if you have not solved problem (6) then I urge you to try it one last time. Its solution also depends on an insight, which contravenes the nature of all the shape problems that you have so far encountered in the chapter. This solution depends on relaxing the constraint that shapes are two-dimensional. They can also be three-dimensional, and Figure 1.3 shows the required initial shape. If you remove the central vertical piece, then you eliminate all four squares, because they each become a rectangle.

When you have an insight, at one moment you are unable to solve a problem, and then at the next moment you are aware of the solution. The three preceding problems illustrate the present theory's account of insight. It depends on changing the constraints currently governing the problem-solving strategy. The current constraints fail to yield a solution. Indeed, the problem may be provably insoluble given the current constraints. The process of insight accordingly depends on the following steps:

(1) The current strategy fails to yield a solution.
(2) There is a tacit consideration of the constraints in the strategy.
(3) The constraints are relaxed in a new way.
(4) Many changes in constraints lead nowhere, but, with perseverance, a change may be made that leads at once to the solution of the problem.

For example, the constraint that squares are the same size is relaxed to the constraint that squares can be of different sizes. In some cases, as Kaplan and Simon (1990) argued, the change in constraints can lead to a new representation of the problem. For instance, if squares can be in a three-dimensional arrangement, then four squares (or more) can have a piece in common. Once such an insight has occurred, it may be used henceforth to constrain the generative stage itself: Individuals look at once for a three-dimensional solution when two-dimensional attempts all fail.

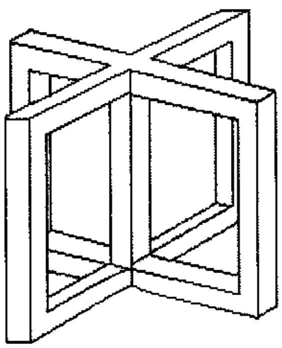


Figure 1.3 A three-dimensional shape that solves problem (6)

Conclusions

This chapter has presented a theory of how people solve problems. According to this theory, as they explore different "moves" in trying to solve a problem, they discover a variety of tactical possibilities. These discoveries can occur even if individuals fail to solve the problem in hand. But once they have acquired a repertoire of tactics, exploration can also lead to the development of a strategy for coping with the problems. The knowledge derived from tactical exploration yields constraints on the process of problem solving. The constraints are used at first in the evaluative stage of thinking. But, this neo-Darwinian strategy is inefficient. Hence, individuals are likely to shift some constraints to the actual generation of ideas to create a multistage strategy. This shift is part of the process of discovering how to solve problems. The problem-solver's dream, however, is to acquire sufficient constraints on the generative process that it never yields any erroneous results. Such a neo-Lamarckian strategy has to be acquired by those who master the improvisation of ideas in real time, e.g., professional jazz musicians (Johnson-Laird, 2002). Whether anyone develops a neo-Lamarckian strategy for shape problems is an open question.

The present theory was illustrated by a series of shape problems, and the reader may wonder whether it also applies to one-off problems, such as Duncker's candle puzzle. For a complex one-off problem, the same process of discovery – yielding a knowledge of tactics – is feasible. But, most one-off laboratory problems call for insight. That is, they are designed so that the problem-solver's initial strategy is almost certain to fail. Insight, on the present account, depends on the discovery of new constraints. They do not always lead to a solution, but when they do yield one suddenly, all the hallmarks of the classic Gestalt phenomenon occur.

The Gestalt psychologists pioneered the study of problem solving. Paolo Legrenzi, to whom this chapter is dedicated, is important because he brought Gestalt ideas into modern cognitive psychology, and showed how they related to mental models (see, e.g., Legrenzi, 1994). The present theory has focused not on this aspect of the representation of problems, but on how individuals acquire tactical knowledge and use it to develop strategic thinking.

Acknowledgements

The research reported in this chapter was supported in part by a grant from the National Science Foundation (Grant BCS 0076287) to study strategies in reasoning. The author thanks Tom Ormerod for many conversations about problem solving, Louis Lee for collaborating in experimental studies of shape problems, Uri Hasson for scholarly advice and thoughts about strategies, and the following colleagues for their helpful criticisms: Vittorio Girotto, Sam Glucksberg, Geoff Goodwin, Cathy Haught, and Sanna Reynolds. He also thanks the participants of the workshop for pointing out some of his errors. And, most of all, thanks to Paolo Legrenzi and Maria Sonino Legrenzi: best friends, outstanding colleagues.

References

Bateson, G. (1979). *On mind and nature*. London: Wildwood House.

Bulbrook, M. E. (1932). An experimental study into the existence and nature of "insight". *American Journal of Psychology*, *44*, 409–453.

Campbell, D. (1960). Blind variation and selective retention in creative thought as in other knowledge processes. *Psychological Review*, *67*, 380–400.

Chronicle, E. P., Ormerod, T. C., & MacGregor, J. N. (2001). When insight just won't come: The failure of visual cues in the nine-dot problem. *Quarterly Journal of Experimental Psychology*, *54A*, 903–919.

Dawkins, R. (1976). *The selfish gene*. Oxford: Oxford University Press.

Duncker, K. (1945). On problem solving. *Psychological Monographs*, *58*(5), whole number 270.

Holland, J., Holyoak, K. J., Nisbett, R. E., & Thagard, P. (1986). *Induction: Processes of inference, learning, and discovery*. Cambridge, MA: Bradford Books, MIT Press.

Hopcroft, J. E., & Ullman, J. D. (1979). *Formal languages and their relation to automata*. Reading, MA: Addison-Wesley.

Isaak, M. I., & Just, M. A. (1995). Constraints on thinking in insight and invention. In R. J. Sternberg, & J. E. Davidson (Eds.), *The nature of insight* (pp. 281–325). Cambridge, MA: Bradford Books, MIT Press.

Jansson, D. G., & Smith, S. M. (1991). Design fixation. *Design Studies*, *12*, 3–11.

Johnson-Laird, P. N. (2002). How jazz musicians improvise. *Music Perception*, *19*, 415–442.

Kanizsa, G., & Legrenzi, P. (1978). *Psicologia della gestalt e psicologia cognitivista*. Bologna: Il Mulino.

Kanizsa, G., Legrenzi, P., & Sonino, M. (1983). *Percezione, linguaggio, pensiero*. Bologna: Il Mulino.

Kaplan, C. A., & Simon, H. A. (1990). In search of insight. *Cognitive Psychology*, *22*, 374–419.

Katona, G. (1940). *Organizing and memorizing: Studies in the psychology of learning and teaching*. New York: Columbia University Press.

Keane, M. (1989). Modelling problem solving in Gestalt "insight" problems. *Irish Journal of Psychology*, *10* (2), 201–215.

Knoblich, G., Ohlsson, S., Haider, H., & Rhenius, D. (1999). Constraint relaxation and chunk decomposition in insight problem-solving. *Journal of Experimental Psychology: Learning, Memory & Cognition*, *25*, 1534–1556.

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

Köhler, W. (1929/1947). *Gestalt psychology*. New York: Liveright.

Korf, R. E. (1980). Toward a model of representational changes. *Artificial Intelligence*, *14*, 41–78.

Lee, L., & Johnson-Laird, P. N. (2003). How individuals learn to solve problems. In preparation.

Legrenzi, P. (1978). Rilettura dei temi della Gestalt in termini di psicologia cognitivista. In G. Kanisza & P. Legrenzi (Eds.), Psicologia della Gestalt e psicologia cognitivista (pp. 165–174). Bologna: Il Mulino.

Legrenzi, P. (1994). Kanizsa's analysis of "prägnanz" as an obstacle to problem solving and the theory of mental models. *Japanese Journal of Psychology*, *36*, 121–125.

Levine, M., (1994). *Effective problem solving* (2nd ed.). Englewood Cliffs, NJ: Prentice Hall.

Luchins, A. S. (1942). Mechanization in problem solving: The effect of einstellung. *Psychological Monographs*, *54*(6), whole number 248.

Luchins, A. S., & Luchins, E. S. (1950). New experimental attempts at preventing mechanization in problem solving. *Journal of General Psychology*, *42*, 279–297.

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*, 176–201.

Maier, N. R. F. (1931). Reasoning in humans: II. The solution of a problem and its appearance in consciousness. *Journal of Comparative Psychology*, *12*, 181–194.

Mayr, E. (1982). *The growth of biological thought: Diversity, evolution, and inheritance*. Cambridge, MA: Belknap, Harvard University Press.

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

Newell, A. (1990). *Unified Theories of Cognition*. Cambridge, MA: Harvard University Press.

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

Ohlsson, S. (1984a). Restructuring revisited I: Summary and critique of Gestalt theory of problem solving. *Scandinavian Journal of Psychology*, *25*, 65–76.

Ohlsson, S. (1984b). Restructuring revisited II: An information processing theory of restructuring and insight. *Scandinavian Journal of Psychology*, *25*, 117–129.

Orleans, S., & Orleans, J. (1983). *The great big book of pencil puzzles*. New York: Pedigree.

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*, 791–799.

Perkins, D. N. (1981). *The mind's best work*. Cambridge, MA: Harvard University Press.

Perkins, D. (2000). *The eureka effect: The art and logic of breakthrough thinking*. New York: Norton.

Simonton, D. K. (1984). *Genius, creativity, and leadership: Historiometric inquiries*. Cambridge, MA: Harvard University Press.

Simonton, D. K. (1995). Foresight in insight? A Darwinian answer. In R. J. Sternberg & J. E. Davidson (Eds.), *The nature of insight* (pp. 465–494). Cambridge, MA: MIT Press.

Skinner, B. F. (1953). *Science and human behavior*. New York: Macmillan.

Smith, S. M., & Blankenship, S. E. (1991). Incubation and the persistence of fixation in problem solving. *American Journal of Psychology*, *104*, 61–87.

Van der Henst, J.-B., Yang, Y. R., & Johnson-Laird, P. N. (2002). Strategies in reasoning. *Cognitive Science*, *26*, 425–468.

Weisberg, R. W. (1986). *Creativity: Genius and other myths*. New York: Freeman.
Weisberg, R. W. (1993). *Creativity: Beyond the myth of genius*. New York: Freeman.
Weisberg, R. W. (1996). Prolegomena to theories of insight in problem-solving: A taxonomy of problems. In R. J. Sternberg & J. E. Davidson (Eds.), *The nature of insight* (pp. 157–196). Cambridge, MA: MIT Press.
Wertheimer, M. (1945/1959). *Productive thinking*. Chicago: University of Chicago Press.