Journal of Experimental Psychology:
Learning, Memory, and Cognition
2002, Vol. 28, No. 4, 791–799

0278-7393/02/$5.00 DOI: 10.1037//0278-7393.28.4.791

Dynamics and Constraints in Insight Problem Solving

Thomas C. Ormerod
Lancaster University

James N. MacGregor
University of Victoria

Edward P. Chronicle
Lancaster University

This article reports 2 experiments that investigated performance on a novel insight problem, the 8-coin problem. The authors hypothesized that participants would make certain initial moves (*strategic moves*) that seemed to make progress according to the problem instructions but that nonetheless would guarantee failure to solve the problem. Experiment 1 manipulated the starting state of the problem and showed that overall solution rates were lower when such strategic moves were available. Experiment 2 showed that failure to capitalize on visual hints about the correct first move was also associated with the availability of strategic moves. The results are interpreted in terms of an information-processing framework previously applied to the 9-dot problem. The authors argue that in addition to the operation of inappropriate constraints, a full account of insight problem solving must incorporate a dynamic that steers solution-seeking activity toward the constraints.

The introduction of the study of insight into modern psychology is generally credited to the Gestalt psychologists: most importantly, Kohler's (1925) work on intelligent problem solving in chimpanzees, Wertheimer's (1959) studies of the role of restructuring in productive thinking, and Duncker's (1945) program of experiments on insight problems. This pioneering work was followed by a relatively long hiatus that, judging by the recent resurgence of interest, has now ended (e.g., Sternberg & Davidson, 1995). A common theoretical thread that runs through recent contributions is that insight requires the removal of one or more unnecessary constraints imposed by the solver on the actions that they take in attempting to solve the problem (e.g., Adams, 1974; Davidson, 1995; Gick & Lockhart, 1995; Ohlsson, 1992; Smith & Blankenship, 1991). For example, consider the nine-dot problem, a problem that is simple to state, yet notoriously difficult to solve. The task is to draw four straight lines that, together, intersect each dot of a regular 3×3 grid of dots, without retracing and without lifting the pen off the paper until the end of the final line. The traditional Gestalt explanation for the problem's difficulty is that solvers impose an implicit constraint that lines may not violate the boundary of the square formed by the nine dots (e.g., Scheerer, 1963). More recent explanations (e.g., Lung & Dominowski, 1985; Weisberg & Alba, 1981) point to different constraints but share the view that the locus of problem difficulty is centered on the solver's constrained representation of the problem.

Despite widespread agreement that inappropriate constraints are the main source of problem difficulty and that their removal allows insight to occur, the mechanisms by which they are removed, and the processes that enable attention to be focused on more fruitful aspects of a problem, remain puzzling. This puzzling nature of the insight process has been expressed in the following, almost paradoxical, terms. If a problem is eventually solved, then the solver clearly has the knowledge or the competency to do so. Why, then, does the impasse arise in the first place? On the other hand, given that an impasse has arisen, what, then, makes it go away? (Knoblich, Ohlsson, Haider, & Rhenius, 1999; Ohlsson, 1984). The answer Knoblich et al. (1999) and Ohlsson (1992) proposed was that past experience biases the initial representation of the problem in a manner that hinders finding the solution and that to overcome this, a change in the problem representation is required. Ohlsson's (1992) general insight framework has been developed by Knoblich et al. (1999) into a more precise and testable theory that proposes key roles for both constraint relaxation and *chunk decomposition*, a particular type of reencoding, as sources of insightful moves. Knoblich et al. argued that the probability of any particular problem constraint being relaxed is inversely related to its scope, that is, how much of the current problem representation is affected by the constraint. Similarly, they argued that the probability of reencoding any particular piece of problem information by decomposition is an inverse function of the tightness with which that information is chunked in the current representation. Chunks are loose if they can be decomposed into constituent elements that themselves are recognizable chunks, whereas they

Thomas C. Ormerod and Edward P. Chronicle, Department of Psychology, Lancaster University, Lancaster, United Kingdom; James N. MacGregor, Department of Public Administration, University of Victoria, Victoria, British Columbia, Canada.

This research was partially supported by a research grant to James N. MacGregor from the Natural Science and Engineering Research Council of Canada. We gratefully acknowledge the statistical help of Mike Hunter, Department of Psychology, University of Victoria, and Damon Berridge, Centre for Applied Statistics, University of Lancaster, who advised us on the analysis of binary data. Any errors of understanding are ours alone.

Correspondence concerning this article should be addressed to Edward P. Chronicle, Department of Psychology, Lancaster University, Lancaster LA1 4YF, United Kingdom. E-mail: E.Chronicle@lancaster.ac.uk

are tight if the elements cannot be meaningfully encoded as chunks.

Modern impasse-based theories in general, and the theory of Knoblich et al. (1999) in particular, resemble earlier Gestalt approaches to insight in a number of respects: (a) The proposal that past experience biases problem representation is similar to the idea of "set" or "Einstellung" (Luchins, 1942); (b) the idea that resolution requires a change in representation, which in particular circumstances may be achieved by chunk decomposition, recalls the Gestalt concepts of restructuring and unit segregation (Koffka, 1935; Kohler, 1925); and (c) the concept of a "constraint" appears to have much in common with the Gestalt notion of a "barrier," and the concept of "constraint relaxation" appears to have much in common with the reduction or cessation of restraining forces (Lewin, 1936). However, a central feature of the Gestalt view of insight that appears to be absent from modern impasse-based theories is a dynamic, or driving force (Lewin, 1936), that leads solvers to select some problem moves in preference to others. This missing element is discussed below.

One of the earliest paradigms of insight problem solving studied by Gestalt psychologists was the detour problem (Kohler, 1925; Lewin, 1935). In the introduction to *Mentality of Apes*, frequently cited as the formative work in the study of insight, Kohler writes ". . . the experiments described in the following pages are of one and the same kind: the experimenter sets up a situation in which the direct path to the objective is blocked, but a roundabout way left open. The animal is introduced into this situation which can, potentially, be wholly surveyed. So we can see . . . whether it can solve the problem in the possible 'roundabout' way." (Kohler, 1925, p. 6.) Apparent in this description are the seeds of the questions more recently raised by Ohlsson (1984). Because the situation can be "wholly surveyed," what is it that initially prevents the animal from seeing the route to goal? What is it that eventually allows the route to be discovered? The Gestalt answer to the second question—restructuring—is the same as Ohlsson's (1984); however, the answer to the first question is different. What prevents the organism from seeing the "roundabout" route is not the bias of past experience per se but the presence of driving forces that focus attention and steer behavior toward the most direct approach to the goal (Lewin, 1936). When the driving force encounters the restraining force of the barrier, the resulting "quasi-stationary equilibrium" (Lewin, 1952, p. 470) forms the impasse. The organism is essentially trapped until a shift in the strength or direction of the driving forces permits some form of restructuring or escape (Barker, Dembo, & Lewin, 1943).

The language of forces (along with other dynamic concepts introduced by the Gestaltists, such as "vectors" and "valences") has little currency in contemporary cognitive psychology. Nevertheless, the dynamic perspective of Gestalt psychology seems to capture a characteristic of the insight process that is not apparent in modern approaches. Its description of a system in a state of tension contrasts sharply with the modern view that the impasse phase of insight problem solving is essentially quiescent or inert. The phenomenology of the "aha" experience, with its sometimes dramatic release of tension, suggests that the older position may have been closer to the truth in this respect and that an impasse consists of not one but two essential characteristics: (a) a constraint and (b) a dynamic component that drives and directs problem-solving activity against the constraint.

Recently, drawing not only on the notions discussed above but also on modern theoretical refinements (e.g., Kaplan & Simon, 1990), we developed a model to explain behavior in the nine-dot problem that included a dynamic component (Chronicle, Ormerod, & MacGregor, 2001; MacGregor, Ormerod, & Chronicle, 2001). The model assumes that people adopt a locally rational strategy of drawing lines that intersect the maximum possible number of adjacent dots. This enables them to approach the goal by the most seemingly direct route. Provided that the resultant move satisfies a *criterion of progress*, it will be selected. When it does not, *criterion failure* occurs, resulting in an impulse to seek alternative moves, which may lead to the relaxation of constraints. In the standard nine-dot problem, adoption of the locally rational strategy drives people down a blind alley because, although seemingly fruitful in terms of the initial criterion, the moves will always end in failure.

MacGregor et al. (2001) reported five experiments supporting the theory, in two of which (Experiments 4 and 5) participants were shown the nine-dot problem with a correct first line drawn. In some cases, this was a horizontal line intersecting the top row of dots and extending beyond the square to the nondot location at the first inflection point of a correct solution. This provided a visual hint that violated the constraint that supposedly needed to be relaxed to achieve a solution. In the contrasting condition, the diagonal line was drawn between the top left and bottom right dots. This did not extend beyond the outline of the figure. However, the model predicted that because criterion failure occurred more frequently in the latter case, the impulse to restructure would be stronger, making solutions potentially more likely. This surprising prediction was supported by both experiments.

We believe these results to be unexpected and potentially important. However, because they have so far been limited to the nine-dot problem their implications for general theories about insight remain in question. The main purpose of the work reported here is to attempt to extend the findings concerning locally rational strategies and criterion failure to another, unrelated, insight problem. A major difficulty in achieving this goal was to identify an appropriate problem, because we required one that involved a series of overt steps or moves, to allow us to observe whether they conformed to locally rational strategies over the course of a solution attempt. This requirement ruled out the ingenious and otherwise flexible matchstick algebra tasks of Knoblich et al. (1999), as well as most other standard insight problems that require only one move or in which the "moves" occur internally. As a result, we devised a set of new insight problems. The remainder of this article presents one of the new problems and reports two experiments designed both to investigate properties of the new problem and to test whether we could observe effects of (a) criterion failure and (b) a visual hint, similar to those observed with the nine-dot problem (MacGregor et al., 2001).

The Eight-Coin Problem

The generic task, of which the eight-coin problem is one example, is to alter an array of *x* coins by moving *y* coins only, to create a final array in which each coin touches exactly *z* others. In the eight-coin problems used in this article, the starting array must be altered by moving two coins so that each coin touches exactly three others. Figure 1 shows two possible variants of the initial

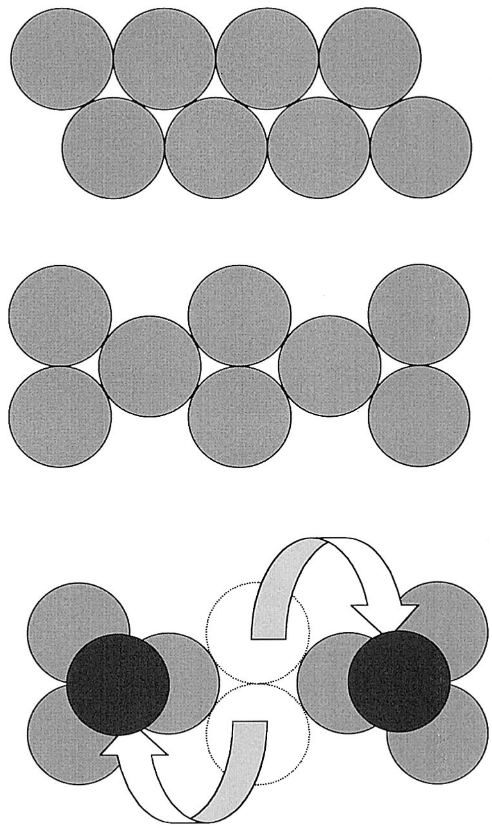

Figure 1. The two starting configurations used in the pilot study for no-move-available (upper panel) and move-available (middle panel) conditions, and a solution for the move-available condition (lower panel). The darker shading represents a coin being stacked on top of other coins.

problem configuration. In each case, the correct solution (also shown in Figure 1) requires taking two coins from the center of the array and stacking one on top of each of the two resulting coin triads.

The primary insight necessary for solving the problem is to switch from moving coins in two dimensions to considering three-dimensional moves. In this respect, the problem is similar to the six-matchstick problem (Scheerer, 1963) and to the four-tree problem (Metcalfe, 1986). In the former, participants must arrange an array of six matchsticks to form four equilateral triangles, a problem that requires the formation of a three-dimensional pyramid as its solution. In the latter, the problem is to specify how a landscape gardener could place four trees so that each one is the same distance from each of the others. In the six-match, four-tree, and eight-coin problems, the source of problem difficulty appears, at least anecdotally, to be that participants constrain their early solution attempts to a two-dimensional horizontal plane. We return to the question of the source of this constraint in the General Discussion.

The general approach of our model of nine-dot problem solving offers a framework for understanding the effects of this two-dimensional constraint. In essence, the extent to which participants limit themselves to two dimensions was proposed to be a function of the availability of two-dimensional moves that can be made in accordance with a locally rational strategy. We proposed that if such moves are available, they will automatically be explored, excluding any immediate consideration of moves in three dimensions. The basic notion that initial moves in multistep problems may be governed by local strategies is not new. Simon and Reed (1976) have, for example, argued persuasively that in the case of well-defined problems, such as Missionaries and Cannibals, the task environment may mean that ". . . subjects must rely on simple local strategies—selecting moves that balance the numbers of missionaries and cannibals, or those that take the greatest number across the river—which permit only a short look ahead" (p. 87).

Before considering local strategies for the eight-coin problem, a note on terminology is necessary. In our previous article on the nine-dot problem (MacGregor et al., 2001), a distinction was made between a strategy for selecting moves and a criterion for evaluating them. The strategy was to select the line that canceled the maximum number of dots; this selection was evaluated against a criterion based on the minimum number of dots that needed to be canceled at each move.

For the eight-coin problem, there are a number of potential strategies and related criteria that might govern initial moves. Figure 1 shows two possible starting configurations of the problem that differ in the number of moves available under several possible strategies. The strategy most analogous to the one we proposed for the nine-dot problem would be to seek and select the move that maximized the number of coins that touch exactly three others. Moves selected under this strategy would then be evaluated under a criterion based on the minimum of coins that needed to touch exactly three others at each move. However, the cognitive operations that would be required to apply this strategy seem unfeasibly demanding.

Slightly less complex would be an improvement strategy based on increasing (rather than maximizing) the total number of coins touching three at the end of a move. This is marginally less demanding than a maximizing strategy, because it is self-terminating rather than exhaustive, but nevertheless it would impose a high cognitive load. More likely, we believe, would be a strategy in which moves are selected so that the coin being moved simply comes to rest touching three others. It is important to note that in the case of this simplest strategy, the criterion is necessarily satisfied at the point of move selection, and conversely, criterion failure necessarily occurs if the strategy fails to identify any moves. This strategy entails the lowest cognitive load and still contains some element of progress.

To determine whether this simplest strategy influences move selection, we conducted a pilot study using the starting configurations shown in Figure 1. The two arrangements differed in the number of first moves available in which a coin could be moved to touch exactly three others (in two dimensions). In the upper arrangement (the no-move-available [NMA] condition) there were no such moves, whereas in the middle arrangement (the move-available [MA] condition) there were 20. If people selected moves on this basis, then there would have been a relatively large number to exhaust in the latter condition before criterion failure occurred. Therefore, the conditions necessary for seeking alternative moves (e.g., those in three dimensions) would not arise, at least in early problem-solving attempts. In contrast, the fact that no such two-dimensional moves were available in the former condition would, we hypothesized, lead participants to entertain other moves at an earlier stage in the problem-solving process. This might lead them to discover a solution sooner.

The pilot study used 24 participants who were randomly assigned one of the two starting configurations shown in Figure 1 (with equal numbers in each group). They were instructed to rearrange the coins into a configuration in which each coin touched exactly three other coins. Two moves were allowed, where a move was defined as moving one coin at a time. An initial period of 4 min was allowed, in which participants made as many attempts as they wished.

Because the baseline success rates were unknown, the procedure also included two verbal hints to participants so that potential floor effects could be avoided. After 4 min, participants were read the first verbal hint, that "the coins can end up in two separate groups." After one additional minute, the second verbal hint was read, that "a coin can come to rest on top of other coins." This was followed by another minute to attempt solution.

Eleven (92%) of the 12 participants in the NMA condition solved the problem, 2 following the first hint and 9 more following the second. Eight participants (67%) in the MA condition solved the problem, 4 prior to the first hint and 4 more following the second hint.

Three main points of interest were apparent in these pilot data. First, by the end of the procedure, solution rates in the NMA condition were slightly higher than in the MA condition. Although not quite statistically significant, the direction of this difference in solution rates was consistent with our general expectations. Second, solution rates prior to any hint were relatively low, at 17%. Third, contrary to expectations, solution rates prior to any hint favored the MA condition.

With regard to this third point, visual inspection of the configurations shown in Figure 1 suggests that they differ, not only with respect to move availability but in their figural integrity, particularly the extent to which they facilitate the separation of the configuration into two groups. It may be that the MA condition inadvertently allowed the easier identification and separation of the coin triads that form bases for the solution components.

The results of this pilot study were encouraging in that they suggested a role for move availability. They also made it clear that figural factors should be controlled and that the verbal hints would be required in order to achieve adequate solution rates. Experiment 1 was therefore designed to take these points into account.

Experiment 1

The results of the pilot study suggested two factors that might have influenced participants' solution attempts. One was the availability of moves in two dimensions that met, or were consistent with, the progress criterion (that the coin being moved comes to rest touching three others). The other was the degree of figural integrity. The direction of the results was consistent with the hypothesis that the former factor affects overall solution rates (after hints) but also suggested that the latter factor may have a general influence on the likelihood of solution. To further address the potential roles of these factors, we systematically varied both in Experiment 1.

Move availability was varied in terms of the number of first moves that allowed a coin to come to rest touching exactly three other coins (in two dimensions). Figural integrity is less amenable to a simple definition. However, in generating figures that appear to have greater or less figural integrity, we took into account two main factors: (a) symmetry, because a symmetrical figure should have relatively more figural integrity than an asymmetrical one, and (b) decomposability, on the grounds that a figure that can be split apart by moving few elements should have relatively less figural integrity than one that requires moving many elements. In this sense, figural integrity is related to the notion of chunk decomposition (Knoblich et al., 1999), in which integrated figures are equivalent to tight chunks and less integrated figures are equivalent to loose chunks. We recognize that other factors may be relevant to figural integrity, but for convenience we adopted the *tight* and *loose* terms used by Knoblich et al. as labels for the two levels of figural integrity.

The stimulus configurations are shown in Figure 2. The upper and lower figures represent higher and lower levels of the figural integrity factor, showing tight and loose figures, respectively. If figural integrity is the key factor mediating performance, then solution rates should be higher for loose figures than for tight figures. The factor of move availability is represented laterally, with the two figures on the left having no moves available in which a coin can be moved to touch exactly three others, whereas the figures on the right have some (5 moves for the arrangement in the upper right and 10 moves for the lower right). If move availability affects performance, then solution rates should be higher for NMA than for MA problems.

Method

Participants. The 60 participants were graduate and undergraduate students at Lancaster University, Lancaster, United Kingdom.

Materials. Eight hexagonal metal tokens were used instead of coins, with a length of side of 15 mm and a thickness of 3 mm. Hexagonal tokens were used because we felt that they might make it easier for participants to evaluate the number of mutual contacts.

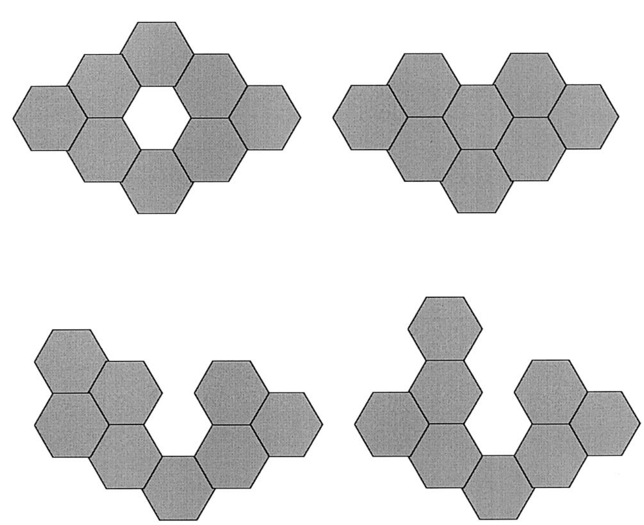

Figure 2. The four starting configurations in Experiment 1. The upper and lower figures represent tight and loose stimuli, respectively. The left and right figures represent no-move-available and move-available stimuli, respectively.

Design. The experiment used two between-subjects factors, each with two levels (represented by the four different starting arrangements shown in Figure 2). The main dependent measure was whether the problem was solved. Also recorded were the first moves made.

Procedure. Participants were randomly assigned to one of the four experimental conditions and tested individually. They were instructed to rearrange the coins displayed in front of them, moving two coins only, so that all eight coins were touching exactly three others. If after moving two coins, participants had not solved the problem, the coins were returned to their original positions and another attempt was made. Participants were given 6 min to work on the problem and were allowed to make as many solution attempts as they wished. At the end of 2 min, participants who had not found the correct solution were told that "the solution requires that the coins should be arranged in two separate groups" (henceforth referred to as the *grouping hint*) and after an additional 2 min, they were told that "the solution requires the use of three dimensions" (henceforth referred to as the *verbal 3-D hint*). Participants' attempts were scored as successful or unsuccessful during the no-hint period, after the grouping hint, and after the verbal 3-D hint. First moves were also recorded.

Results

Four participants were not naive to the problem (having discussed the experiment with previous participants) and were excluded from analysis, leaving 14 participants per condition for the four conditions. Table 1 shows the number of participants who solved the problem in each condition and over the course of the experiment (during the no-hint period, after the grouping hint, and after the verbal 3-D hint). Final numbers of participants who solved the problem at the end of the experiment were analyzed with a procedure for factorial designs with binary data described by Cox and Snell (1989). This procedure was used to compute logistic factorial standardized contrasts, distributed as *Z* scores, for the two main effects and for the interaction of move availability and figural integrity. The main effect of move availability was significant ($Z = 2.16$, $p < .05$): Participants solved the problem more often in the NMA condition. Neither the main effect of figural integrity nor its interaction with move availability were significant ($Z = 1.14$ and $Z = 0.20$, respectively). In addition, we reanalyzed the results using the more familiar analysis of variance (ANOVA) procedure, the use of which with binary data has been widely defended (Gabrielsson & Seeger, 1971; Greer & Dunlap, 1997; Lunney, 1970). The pattern of results was almost identical to the above, with a significant main effect of move availability, $F(1, 52) = 5.20$, $MSE = 0.22$, $p < .05$, but no significant effect of figural integrity, $F(1, 52) = 1.30$, or of their interaction, $F(1, 52) = 0.00$.

An analysis of first moves showed that 7 of the 28 participants (25%) in the MA conditions moved a coin to touch three others in the plane as their first move. It is not completely clear how to calculate chance expectancy here because the instructions placed no constraints on moves. However, the majority of first moves involved a full edge contact with at least one other coin (88%), and the total number of such moves for the two conditions combined was 232, of which 15 (6%) resulted in touching three other coins. The observed value of 25% is significantly different from this chance expectation. (The estimate is conservative, in that if other legal and observed moves, such as stacking coins or removing them from contact with others are included, the population of possible moves increases.) An additional 3 participants made a first move that involved stacking a coin to touch three others.

Table 1
Cumulative Number of Problem Solvers in Each Condition Before and After Verbal Hints in Experiment 1

Condition	No hint	After grouping hint	After verbal 3-D hint
No move available			
Tight	1	3	10
Loose	0	1	12
Total	1 (4)	4 (14)	22 (79)
Move available			
Tight	0	0	6
Loose	0	3	8
Total	0 (0)	3 (11)	14 (50)
Overall total	1 (2)	7 (13)	36 (64)

Note. Numbers in parentheses are percentages. $n = 14$ in each condition.

Discussion

Experiment 1 compared the relative influence of move-availability and figural-integrity factors on performance. The results suggest, for these problem configurations at any rate, that the key factor in determining final solution rates is move availability. Significantly more participants solved the problem in the NMA conditions (79%) than in the MA conditions (50%). In comparison, the difference in solution rates between tight (57%) and loose (71%) conditions was not significant, although it was in the direction expected from the theory of Knoblich et al. (1999). As in the pilot study, the number of participants solving the problem prior to the verbal 3-D hint was low.

Analysis of participants' first moves in the MA conditions showed a higher proportion of first moves that moved a coin to touch three other coins than would be expected by chance alone. A more general tendency to maximize the number of coins that are touched by the moved coin is also apparent in the data. For example, 4 participants in the tight–NMA condition produced a first move in which a coin was moved from the extreme right or left into the center of the coin array.

Experiment 2

Participants in Experiment 1 were provided with two explicit hints, the grouping hint and the verbal 3-D hint. Despite the critical nature of the verbal 3-D hint, not all participants were able to capitalize on it. The data of Experiment 1 seem to indicate that the success of this hint in facilitating the discovery of solutions was a function of the availability of moves under a simple local strategy. However, it may be that the verbal hints used were, for some participants at any rate, ambiguous or unsatisfactory. For example, it may be that preconscious problem constraints cannot be entirely overridden by consciously processed strategic instructions. Alternatively, it may be that attempts to solve the problem prior to the hints established a set to restrict the moves in some way that could not be broken by the later presentation of a hint. Furthermore, it may be that some participants simply needed to work on the hints for longer than other participants before they could capitalize on them in finding a solution.

Thus, Experiment 2 was conducted to investigate the effects of providing an additional visual hint about the relevance of three dimensions, as part of the initial coin configuration. The visual hint consisted of one of the coins being placed directly on top of another, in the initial configuration (see Figure 3). We hypothesized that if the failure to capitalize on solution-relevant information can be explained solely by the quality of the hints, then participants in the visual-hint (VH) conditions should solve the problem more often than participants in the no-visual-hint (NVH) conditions. If, on the other hand, the simple local strategy of attempting to move a coin to touch three other coins is dominant, then participants in the NMA conditions should solve the problem more often than participants in the MA conditions, regardless of the presence or absence of the visual hint. A specific prediction was available stemming from the results of MacGregor et al. (2001; Experiments 4 and 5), who found that criterion failure was more effective in facilitating solutions than a visual hint to extend lines beyond the boundary of the nine dots. The analogous prediction for the present experiment was that performance in the VH–MA cell of the design should be worse than in the NVH–NMA cell. In other words, we hypothesized that the absence of a criterion-satisfying move would have more effect than the presence of a constraint-relaxing hint.

Method

Participants. The 52 participants were graduate and undergraduate students at Lancaster University, excluding students of psychology.

Materials. As in the previous experiment, the stimulus for each condition consisted of an array of eight coins, as illustrated in Figure 3. In this experiment, however, the VH conditions had one coin stacked vertically over another coin as a visual hint to participants to work in three dimensions.

Design. The experiment used two between-subjects factors, each with two levels (represented by the four different starting arrangements shown in Figure 3). The main dependent measure was whether the problem was solved. Also recorded were the first moves made.

Procedure. The procedure was the same as in the previous experiment, with the exception that participants in all groups worked on the problem for an initial period of 6 min. If they had not solved the problem after 6 min, they were presented with the grouping hint. If they still had not solved after an additional minute, they were given the verbal 3-D hint and were then allowed a further minute to find a solution.

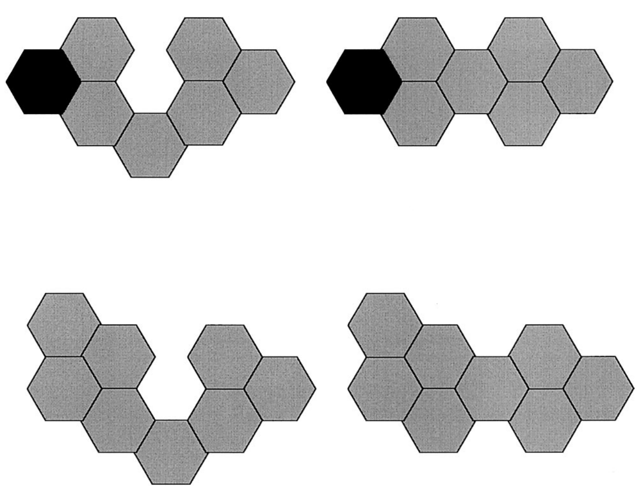

Figure 3. The four starting configurations of Experiment 2. The upper and lower figures represent visual-hint and no-visual-hint stimuli, respectively, with the darker shading representing a stack of two coins. The left and right figures represent no-move-available and move-available stimuli, respectively.

Results

The numbers of participants solving the problem in each condition, over the course of the experiment, are shown in Table 2. As in Experiment 1, the final numbers of participants solving the problem at the end of the experiment were analyzed using the Cox and Snell (1989) procedure. There was a significant main effect of move availability ($Z = 2.51$, $p < .05$): Participants solved the problem more often in the NMA condition. There was no significant effect of visual hint or of the interaction of both factors ($Z = 0.24$ in each case). Reanalysis using ANOVA produced a similar pattern of significant and nonsignificant results: move availability, $F(1, 48) = 7.59$, $MSE = 0.21$, $p < .01$; figural integrity, $F(1, 48) = 0.09$, *ns*; interaction, $F(1, 48) = 0.09$, *ns*. Because a specific prediction had been made, it was also appropriate to make a planned comparison between the VH–MA and NVH–NMA conditions. This comparison was statistically significant (Fisher's exact test, $p = .048$): Participants solved the problem more often in the NVH–NMA condition (11 of 13 vs. 6 of 13).

An analysis of first moves showed that 12 of the 26 participants (46%) in the MA conditions moved a coin to touch three others in the plane as their first move. This is significantly greater than the chance expectation of 6% (derived as in Experiment 1). An additional 6 participants made a first move that involved stacking a coin to touch three others.

Discussion

Experiment 2 was conducted to investigate the effects of providing a visual hint to work in three dimensions, in combination with the effect of varying move availability. As in Experiment 1, the results show clearly that by the end of the experimental procedure, participants in the NMA conditions solved the problem more frequently than did participants in the MA conditions. There was no overall effect of visual hint; in fact, the final solution rate for the VH conditions was slightly lower than that for the NVH conditions (65% vs. 69%). It is important to note, however, that this was not simply a result of using an inadequately clear hint. Some participants were able to see the relevance of the visual hint early in the procedure, with 10 participants (38%) in the VH conditions solving the problem prior to any verbal hint compared with 2 (8%) in the NVH conditions (see Table 2).

The finding of the planned comparison, that there were significantly more solutions in the NVH–NMA condition than in the VH–MA condition, is particularly important because it shows how the availability of moves that meet the criterion influences the effectiveness of hints. It suggests that some participants in the VH–MA condition were unable to override the use of the simple local strategy for move selection, despite the seemingly high salience of the visual hint (especially after the verbal 3-D hint had been provided). In contrast, although few participants in the NVH–

Table 2
Cumulative Number of Problem Solvers in Each Condition of Experiment 2 Before and After Verbal Hints

Condition	Initial 6-min period	After grouping hint	After verbal 3-D hint
No move available			
Visual hint	6	10	11
No visual hint	2	2	11
Total	8 (31)	12 (46)	22 (85)
Move available			
Visual hint	4	5	6
No visual hint	0	1	7
Total	4 (15)	6 (23)	13 (50)
Overall total	12 (23)	18 (35)	35 (67)

Note. Numbers in parentheses are percentages. $n = 13$ in each condition.

NMA group realized spontaneously (that is, prior to any verbal hint) that a move in three dimensions might contribute to problem solution, they reliably found solutions as soon as the verbal 3-D hint had been presented.

We suggest that the experience of criterion failure in the NMA conditions created a state of preparedness necessary for capitalizing on novel solution-relevant information. In the case of the VH–NMA condition, most participants were able to use the visual hint, especially after receiving the grouping hint, and did not require the verbal 3-D hint. They were able to do so, we argue, because the absence of moves meeting the criterion in the initial configuration created the conditions under which the visual hint could be used. In the case of the NVH–NMA condition, once participants were presented with the verbal 3-D hint, they were able to capitalize on it. Fewer participants were able to do so in the NVH–MA condition because they did not experience criterion failure (they continued to make moves that met the criterion). These results are highly consistent with the effects found by MacGregor et al. (2001; Experiments 4 and 5) with the nine-dot problem, suggesting that criterion failure plays a general role in creating the conditions necessary to achieve insight.

One difficulty with both Experiments 1 and 2 is that factor levels of move availability and figural integrity are instantiated by unique problem arrangements. These problem arrangements, it may be argued, might also differ along some uncontrolled dimension and thus influence problem solving. Although the problem arrangements in the two experiments were carefully selected to minimize gross differences of appearance, it would technically be preferable to select randomly from all possible problem arrangements the specific arrangements for each condition of the experiments. We are currently investigating the feasibility of such an approach in the generic coin task.

General Discussion

The present article presented the novel eight-coin problem for exploring insight problem solving, along with two experiments that tested predictions about the role of criterion failure. In particular, these experiments tested the hypothesis that the value of a constraint-relaxing hint depends critically on the experience of criterion failure. Both experiments supported this general prediction.

The work presented here has three main outcomes. First, it demonstrates the utility of the eight-coin problem in allowing a test of specific hypotheses through manipulation of problem features such as starting configuration and format of solution-relevant hints. Participants' performance with the problem (after at least one hint) appeared to be neither at floor nor at ceiling.

Second, this work provides further support for an information-processing approach to insight problem solving, developed with respect to the nine-dot problem (MacGregor et al., 2001) and generalized here to the eight-coin problem. Our approach was to propose that solvers tackle insight problems by adopting a simple and locally rational strategy and by monitoring the performance of that strategy against a criterion. Moves that satisfy the criterion are more likely to be selected and retained than other moves but will result in failure to solve if they do not correspond to the necessary moves. Insight can be achieved when criterion failure occurs, thus signaling the need to abandon the current operator and to search for an alternative operator. In the eight-coin problem, the availability of moves that follow a simple local strategy of moving the first coin to touch three others (in two dimensions) seemed to be the primary determinant of a successful solution. When no such move was available, criterion failure occurred early in the experimental procedure, so that when appropriate hints were provided, an alternative operator was quickly discovered. However, when such moves were available, the alternative successful move remained elusive.

Third, this work provides evidence that the successful achievement of insight does not depend solely on release from unwarranted problem constraints. Instead, insight may also require that a state of preparedness be reached in which the solver is disposed to attend to solution-relevant information. The results challenge accounts in which insight is seen to occur as a natural result of the release from unwarranted constraints (e.g., Isaak & Just, 1995). Although the concepts of a self-imposed constraint and of constraint relaxation are important in interpreting aspects of our findings, they do not explain the observed differences across conditions. This is because the presumed constraint against stacking coins in the third dimension should have been equivalent across conditions (except, arguably, in the conditions in which the visual hint was provided from the outset). The results indicate that although the release from some form of constraint may be a necessary condition for insight to occur, it is not always sufficient.

The concept of preparedness is also central to the opportunistic assimilation hypothesis of Seifert, Meyer, Davidson, Patalano, and Yaniv (1995). In this account, experiences of impasse are recorded in long-term memory as failure indices. During a subsequent period of incubation, solvers may serendipitously encounter an external object or event that triggers the failure indices. If a solver has not reached an impasse, then there are no failure indices to be triggered. In common with Seifert et al., we maintain that some experience of failure is necessary for insight to occur. The opportunistic-assimilation hypothesis cites actual failure, or impasse, as being critical to creating the conditions necessary for insight. In our theory, criterion failure alone may be sufficient for insight to occur. Criterion failure may or may not be accompanied by the experience of impasse, depending on whether participants are able to identify alternative operators at the time criterion failure occurs.

A remaining question is why participants were constrained initially, and so forcefully, to moves that were available in two dimensions even though there were moves available that satisfy the criterion in three dimensions. In the case of coin problems, three-dimensional arrangements clearly lie within the experience of participants, and face-to-face stacks of coins (e.g., rolls of quarters) are the standard way of obtaining change in U.S. and U.K. banks. The hexagonal tokens used in the experiments may be insufficiently similar to coins for this experience to generalize, but if so, the visual hint provided in Experiment 2 would seem to be a sufficient prompt. We suggest that constraints need not arise solely through the retrieval of prior experience. Rather, two-dimensional moves may be considered from the outset simply because the problems are initially presented in two dimensions. Even in the case of the VH conditions of Experiment 2, in which one coin was stacked on top of another, the most readily available slots for coins to touch three others were available in two dimensions. The search for moves that meet a criterion will, therefore, be constrained initially by the presented display. In this sense, insight problem solving has elements in common with accounts of display-based problem solving (Larkin, 1989; Zhang, 1997; Zhang & Norman, 1994). Nonetheless, it remains possible that there were other, hitherto unidentified, constraints at work in the initial problem arrays that influence move selection and that might arise through prior experience of some kind.

Criterion failure should, according to our theory, signal the search for alternative operators. Yet, in the NMA conditions of the current experiments in which criterion failure was immediate, few participants solved the problem prior to the presentation of verbal hints. There are, in principle, an infinite number of alternative solution possibilities (e.g., cutting, melting, or otherwise changing the physical structure of the coins; putting the coins in motion; discovering tricks in the problem wording). In an infinite space of possible operators, there is no guarantee that participants will spontaneously hit on an alternative problem representation that contains the correct operators. Thus, facilitation in our experiments required both criterion failure to provide an incentive to search for new moves and hints that usefully constrain the space of moves considered.

At the same time, a few participants in each of the conditions did discover solutions without the need for hints. Clearly, an important requirement of solution discovery in insight problem solving is to test new moves. We have previously identified, in the case of the nine-dot problem, move attempts that were unsuccessful but nonetheless predictive of subsequent solution. For example, such a move might extend lines beyond the dot array in only one of the two corners where such line extensions are necessary for solution (MacGregor et al., 2001). In the current experiments, an incorrect move that often occurred immediately prior to solution was to place each of two coins on top of three others while failing to separate the array into two groups of coins. However, identifying a previously untested move was not always sufficient to guarantee solution, even when such a move captures the conceptual insight necessary to solve the problem. Many participants in our nine-dot studies made attempts in which lines were extended beyond the dot array but subsequently returned to making moves that remained within the array. Similarly, in the current experiments we witnessed a number of participants make a three-dimensional move attempt in which a coin was balanced on its edge, before returning to try moves again in two-dimensions.

The reason why some moves are retained and others are rejected when an attempt fails is a function not simply of the conceptual insight inherent in the move but also of the extent to which the move increases criterion satisfaction. Move types are likely to be repeated when they make progress in meeting the current criterion relative to previous attempts. Thus, the key to discovering and retaining an insightful move lies, we suggest, in the application of the same simple goal-directed and locally rational strategy that constrained performance in the first place. Our account of how insightful moves emerge differs from that offered by Seifert et al. (1995) in this respect. According to their opportunistic-assimilation hypothesis, participants serendipitously encounter an external object or event that triggers failure indices recorded in previous attempts. In contrast, we propose that the discovery and retention of insightful moves is not serendipitous but is guided by the search for novel moves that maximize criterion satisfaction.

In conclusion, the present results support the view that impasses arise from the conjunction of two factors, the first being a constraint, the second, some form of drive or dynamic that steers activity into the constraint. The former factor has been emphasized by modern theoretical accounts of insight. The latter has not, although it is compatible with both the Gestalt approach to insight and the information-processing approach to general problem solving, in which strategies based on means–ends analysis may be viewed as providing the guiding component that steers activity along the most direct route to the goal.

References

Adams, J. L. (1974). *Conceptual blockbusting*. New York: Freeman.

Barker, R. G., Dembo, T., & Lewin, K. (1943). Frustration and regression. In R. G. Barker, J. B. Kounin, & H. F. Wright (Eds.), *Child behavior and development* (pp. 441–458). New York: McGraw-Hill.

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: Human Experimental Psychology*, *54*(A), 903–919.

Cox, D. R., & Snell, E. J. (1989). *Analysis of binary data* (2nd ed.). London: Chapman & Hall.

Davidson, J. E. (1995). The suddenness of insight. In R. J. Sternberg & J. E. Davidson (Eds.), *The nature of insight* (pp. 125–155). Cambridge, MA: Bradford Books/MIT Press.

Duncker, K. (1945). On problem solving. *Psychological Monographs*, *58*, 1–113.

Gabrielsson, A., & Seeger, P. (1971). Tests of significance in two-way designs (mixed model) with dichotomous data. *British Journal of Statistical Psychology*, *24*, 111–116.

Gick, M. L., & Lockhart, R. S. (1995). Cognitive and affective components of insight. In R. J. Sternberg & J. E. Davidson (Eds.), *The nature of insight* (pp. 197–228). Cambridge, MA: Bradford Books/MIT Press.

Greer, T., & Dunlap, W. P. (1997). Analysis of variance with ipsative measures. *Psychological Methods*, *2*, 200–207.

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.

Kaplan, C. A., & Simon, H. A. (1990). In search of insight. *Cognitive Psychology*, *22*, 373–419.

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*, 1534–1556.

Koffka, K. (1935). *Principles of Gestalt psychology.* New York: Harcourt, Brace & World.

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

Larkin, J. (1989). Display-based problem solving. In D. Klahr & J. Kotovsky (Eds.), *Complex information processing* (pp. 319–341). Cambridge, MA: MIT Press.

Lewin, K. (1935). *A dynamic theory of personality.* New York: McGraw-Hill.

Lewin, K. (1936). *Principles of topological psychology.* New York: McGraw-Hill.

Lewin, K. (1952). Group decision and social change. In G. E. Swanson, T. M. Newcomb, & E. L. Hartley (Eds.), *Readings in social psychology* (pp. 459–473). New York: Holt.

Luchins, A. S. (1942). Mechanization in problem solving. *Psychological Monographs*, *54*, 1–95.

Lung, C. T., & Dominowski, R. L. (1985). Effects of strategy instructions and practice on nine-dot problem solving. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *11*, 804–811.

Lunney, G. H. (1970). Using analysis of variance with a dichotomous dependent variable: An empirical study. *Journal of Educational Measurement*, *4*, 263–269.

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.

Metcalfe, J. (1986). Feeling of knowing in memory and problem solving. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *12*, 288–294.

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

Ohlsson, S. (1992). Information-processing explanations of insight and related phenomena. In M. T. Keane & K. J. Gilhooly (Eds.), *Advances in the psychology of thinking* (Vol. 1, pp. 1–203). London: Harvester.

Scheerer, M. (1963). Problem solving. *Scientific American*, *208*, 118–128.

Seifert, C. M., Meyer, D. E., Davidson, N., Patalano, A. L., & Yaniv, I. (1995). Demystification of cognitive insight: Opportunistic assimilation and the prepared-mind perspective. In R. J. Sternberg & J. E. Davidson (Eds.), *The nature of insight* (pp. 65–124). Cambridge, MA: Bradford Books/MIT Press.

Simon, H. A., & Reed, S. K. (1976). Modeling strategy shifts in a problem-solving task. *Cognitive Psychology*, *8*, 86–97.

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

Sternberg, R. J., & Davidson, J. E. (Eds.). (1995). *The nature of insight.* Cambridge, MA: Bradford Books/MIT Press.

Weisberg, R. W., & Alba, J. W. (1981). An examination of the alleged role of "fixation" in the solution of several "insight" problems. *Journal of Experimental Psychology: General*, *110*, 169–192.

Wertheimer, M. (1959). *Productive thinking.* Chicago: University of Chicago Press.

Zhang, J. (1997). The nature of external representations in problem solving. *Cognitive Science*, *21*, 179–218.

Zhang, J., & Norman, D. A. (1994). Representations in distributed cognitive tasks. *Cognitive Science*, *18*, 87–122.

Received December 6, 2000
Revision received November 17, 2001
Accepted November 17, 2001 ■