What makes an insight problem? The roles of heuristics, goal conception and solution recoding in knowledge-lean problems.

Edward P. Chronicle

University of Hawaii at Manoa

Honolulu, HI, USA

James N. MacGregor

University of Victoria

Victoria, British Columbia, Canada

Thomas C. Ormerod

Lancaster University

Lancaster, UK

To appear in Journal of Experimental Psychology: Learning, Memory & Cognition.

Abstract

Four experiments investigated transformation problems with insight characteristics. Experiment 1 examined two versions of the six-coin problem with different solution properties. Performance on one version with a concrete and visualizable solution followed predictions derived from assuming a hill-climbing heuristic. The determinant of performance on the second version, in which the solution potentially required insight, was unclear. Experiment 2 concluded that the difficulty of this second version stems from the same hill-climbing heuristic, which creates an implicit conceptual block. Experiment 3 investigated four problem variants and confirmed that the difficulty of the potential insight solution is conceptual, not procedural. Experiment 4 compared the six-coin problem with the ten-coin (triangle) problem, and observed the same principles of move selection on both insight and transformation problems. We argue that hill-climbing heuristics provide a common framework for understanding problem difficulty and solution discovery in both transformation and insight problems. We suggest that at least part of the phenomenology of insight may be accounted for by processes of post-solution recoding.

What makes an insight problem? The roles of problem-solving heuristic, goal conception and solution recoding.

In the cognitive psychological literature there has been a recurrent debate as to whether insight represents a distinct class of problem-solving activity. The roots of this debate lie in the Gestalt tradition, with its emphasis on conceptual restructuring as the mechanism of insight problem-solving. The Gestalt explanation has bequeathed modern cognitive science a view of insight as a step function, rather than as a steady, incremental approach towards a goal. In its more recent incarnation, the debate comes down to one between a 'business as usual' view (e.g., Simon, 1986), and a 'special process' view (e.g., Wertheimer, 1985; Schooler, Ohlsson & Brooks, 1993).

A particular source of difficulty for this debate, as Metcalfe and Wiebe (1987) recognise, is determining what exactly is an insight problem. Often, the only selection criterion for problems used in the study of insight is to have been used as an insight problem in a previous study (Weisberg, 1996). Three a-priori approaches to defining insight problems may be identified in the literature. The first approach defines insight problems in terms of their phenomenology. For example, Metcalfe and Wiebe (1987) characterize insight problems as those showing an absence of incremental "feeling of warmth" ratings prior to solution. The second approach emphasizes changes in conceptual knowledge necessary for insightful solutions to be found (Seifert, Meyer, Davidson, Patalano and Yaniv, 1996; Knoblich, Ohlsson and Raney, 2001). The third approach identifies processes underlying insight problem-solving (Kaplan & Simon, 1990; MacGregor, Ormerod & Chronicle, 2001). Each of these definitional approaches has its merits, but differing theoretical stances are still apparent. The phenomenological and conceptual change approaches emphasize the special nature of insight problem-solving, and the process approach emphasizes the similarities between insight and non-insight problem-solving.

Despite different emphases, the majority of approaches recognize that insight problem-solving involves some kind of restructuring of the initial problem representation. What constitutes restructuring, however, and whether the processes underlying restructuring are special or not, are open questions. In an attempt to unpack the Gestalt notion of restructuring to make it more amenable to empirical test, Weisberg (1996) distinguishes between discontinuity and restructuring in problem-solving. A discontinuity in thinking, according to Weisberg, involves a change in the moves that are sampled, while a restructuring involves a change in the underlying representation of a problem, that is, a re-conceptualization of the initial or goal states of a problem, the operators that are available for assembling moves, or the constraints under which moves are sampled. He proposes the following decisions to diagnose whether a problem involves insight or not: first, if the solution process shows a discontinuity (change in approach), then it may be an insight problem; second, if the discontinuity requires restructuring (change in problem representation) then it may be an insight problem (if not, the solution requires discontinuity but not insight); finally, if restructuring is the only way a solution can occur, then it is a "pure" insight problem (if it can be solved by other means, for example trial and error, it is a "hybrid"). Applying these diagnostic criteria to a set of 24 problems previously described as "insight" problems, Weisberg (1996) concluded that four were discontinuity but not insight problems (e.g., anagrams), five were "hybrid" types (e.g., the nine-dot problem), and 15 were "pure" insight problems (e.g., the Matchsticks problem). Later in the paper we discuss data from empirical studies involving problems characteristic of each of these types.

One argument in favor of insight as a special process is the failure of an information-processing approach to make significant inroads into the explanation of insight (Wertheimer, 1985), despite its success in explaining many other kinds of problem-solving (e.g. Newell & Simon, 1972; Anderson, 2000). One exception to this is the work of Kaplan and Simon (1990), who applied an information-processing framework to explain performance on the mutilated checkerboard problem. They argued that solvers apply heuristics to narrow the space of possible moves, and specifically identified a heuristic for detecting invariant features of the problem across attempts. Their account may be limited by a lack of generality (Knoblich et al, 2001), since it is not clear what invariants might enable solutions to be found for other insight problems. What are lacking from current theories of insight problem-solving are general problem-solving heuristics that might apply across a wider range of insight problems.

General heuristics have been widely cited as providing the basis for solving many transformation problems (e.g., Newell & Simon, 1972; Lovett & Anderson, 1996), defined by Greeno (1978, pp.241) as problems in which the solver must apply a finite set of operators to find a sequence of moves that transform an initial situation into a goal state. Heuristics such as hill-climbing and means-ends analysis operate to select moves that appear to make progress towards the goal state. One reason why general heuristics such as these might not appear immediately applicable to insight problem-solving is because the goal state of many insight problems is ill-defined, rendering the evaluation of progress made from a current state towards an unknown goal state seemingly impossible

(VanLehn, 1989). However, as well as evaluating moves against a concrete and visualizable goal state, individuals may also evaluate moves against “locally rational” criteria that indicate whether progress is being made toward partial or intermediate goal properties. For example, Simon and Reed (1976) propose that individuals switch between three locally evaluated heuristics in solving the Missionaries and Cannibals problem: early moves balance the numbers of missionaries and cannibals on each side of the river, intermediate moves maximize progress from one side to the other, and later moves avoid re-visiting previous states. It seems plausible that, in the absence of complete goal information, individuals might also attempt insight problems by selecting locally rational moves that make progress towards partial or intermediate goals (inferred from the problem description or current problem state).

We have recently proposed that a hill-climbing heuristic underlies the selection of moves across a range of variants of the classic nine-dot problem (MacGregor, Ormerod & Chronicle, 2001), and a novel insight problem, the eight-coin problem, where the goal is to transform a given arrangement of 8 coins into one where each coin touches exactly three others, in a specified number of moves (Ormerod, MacGregor & Chronicle, 2002). According to our account, individuals evaluate potential moves against a criterion of satisfactory progress. In the nine-dot problem, the criterion is that each line must cancel a number of dots given by the ratio of dots remaining to lines available. In the eight-coin problem, a range of criteria may be selected, the simplest one being that moves should end with the coin being moved touching exactly three others. What these criteria have in common is that they specify progress in terms of goal properties inferred from the problem statement (e.g., dots must be cancelled, coins must touch three others), rather than in terms of movement towards a known goal state. Individuals fail to solve, we argue, because selecting criterion-meeting moves drives them away from moves that lie on the correct solution path. So, solution attempts on the nine-dot problem stay within the square shape of the dot array because of the many criterion-meeting moves available within that square shape. Solution attempts on the eight-coin problem are restricted to two dimensions because of the ready availability in two dimensions, the form in which the problem is presented, of moves that end in the moved coin touching three other coins. Individuals fail to make the necessary 'insights' to search for moves outside the representation in which the problem is first presented because there is no apparent need to do so and because there is no information presented in the problem statement regarding the value of moves in different dimensions.

When a search fails to yield moves that meet the criterion for satisfactory progress (e.g., when all criterion-meeting moves and their offspring have been exhausted), then according to our account individuals will relax the requirement to maximize progress. If a non-maximal move allows a subsequent move to make more progress than previous attempts, then it is retained as a 'promising state' for future trials. For example, in experiments on the nine-dot problem, we found that participants often drew solution attempts that went outside the dot array (MacGregor et al, 2001). Where an attempt cancelled more dots than previous attempts, participants were likely to repeat lines drawn beyond the array, but if no progress was made then they generally returned to lines drawn within the boundary of the dot array. Thus, in insight problems such as these, hill-climbing provides both the restriction on move sampling that underlies failure, and an incentive to retain promising moves that might permit eventual success. We recognize that our account has – so far – only been demonstrated to generalize across knowledge-lean problems, that is, problems that do not require any expertise in a particular domain. Nonetheless, we feel that our articulation of insight *processes*, contained both in previous studies (Chronicle, Ormerod & MacGregor, 2001; MacGregor et al, 2001; Ormerod et al, 2002) and in the following four experiments, has the potential to generalize further. We return to this point in the General Discussion.

This paper explores the kinds of information that individuals use to derive and confirm their inferences about goal state properties, across a range of problem types. It does so in three ways. First, in Experiment 1, we introduce a problem (the six-coin problem: see Gardner, 1977) that has not previously been investigated in the literature, and that can be configured to reflect characteristics attributed both to transformation and to insight problems. Second, in Experiments 2 and 3, we explore the ways in which individuals identify and confirm hypotheses about the goal properties necessary for a hill-climbing heuristic from problem statements that have under-specified goals. Third, in Experiment 4, we test predictions from a hill-climbing approach in a direct comparison between transformation and insight problems, pitting the six-coin problem against the ten-coin 'triangle' insight problem (see Metcalfe, 1986; Schooler, Ohlsson & Brooks, 1993; Weisberg, 1996). In addition, we investigate the reproducibility of correct sequences of moves as evidence of recoding into a single solution concept. In doing so, we demonstrate process commonalities between transformation and insight problem-solving, and raise further issues about the nature of insight.

Experiment 1

Experiment 1 investigated performance on two versions of the six-coin problem, as illustrated in Figure 1. In the first version (left column, Figure 1), the starting state is two offset rows of three coins, and the goal is shown as a ring of coins. The task is to transform the starting state into the goal state in three moves. A move consists of sliding a single coin, with the constraints (a) that other coins may not be disturbed during the move, and (b) that the coin being moved must come to rest touching exactly two other coins. This version appears to be a transformation task with many of the properties that made the Towers of Hanoi an important vehicle for problem-solving research. Like the Towers of Hanoi, its initial and goal states, single operator, and constraints are explicitly defined from the outset. In addition, and unlike so-called insight problems, finding the solution does not demand any obvious conceptual insight into previously inaccessible moves, but simply the discovery of a sequence of moves that apply a known operator. It might be objected that the standard Towers of Hanoi task does not impose a set number of moves; rather, the instruction is to complete the task in the minimum number of moves possible. In the six-coin problem, three is indeed the minimum number of moves required, given constraints (a) and (b), above. It therefore seems reasonable to regard this version of the six-coin problem as a transformation task in which the participant has one additional piece of information, that is, the minimum number of moves required. On the other hand, anecdotal reports suggest that the problem is highly resistant to solution: individuals often reach an apparent impasse (as defined by Knoblich *et al*, 2001) in generating solution attempts, yet the solution appears deceptively simple when demonstrated. Occasional successes are met with something akin to an ‘aha’ experience.

The second version of the problem used in this experiment (and discussed further below) is shown in the right column of Figure 1. The starting state, operator and constraints are defined exactly as before, but the goal is given abstractly, as “each coin must touch exactly two others”. It should be noted that this abstract version is open to both the ring and two-group solutions shown in Figure 1.

We have determined the entire state space for the six-coin problem. This was done by an exhaustive computation that produced every three-move sequence that was legal according to the problem statement. The computation permitted move sequences that back-tracked and repeated (e.g., in Figure 1, move coin 4 to touch coins 1 and 2, move it back to its original position, then repeat the move to touch coins 1 and 2). The state space is large: there are 7426 legal move sequences, of which 2 reach the ring solution and 176 reach a two-group solution (in a variety of configurations relative to the original array). Interestingly, the 7426 move sequences are not equiprobable under the assumption that moves are selected randomly. This is because the number of available legal second moves varies depending on which of the 24 legal first moves was selected, and likewise, the number of available legal third moves then varies depending on which of the second moves was selected. The overall probability of finding a correct solution by random move selection is not, therefore, given simply by the number of correct move sequences divided by the total number of sequences. Rather, the overall probability of finding a certain type of solution is the sum of the products of the conditional probabilities of each sequence of first, second and third moves that leads to that type of solution. For the ring solution, the overall probability is .00015, and for the two-group solution it is .01866.

We assume, for the moment, that participants attempting the problem employ a hill-climbing heuristic if the ring goal is provided. An available property of states prior to the goal state is the number of adjacent coins that are on the goal configuration, that is, a ring. We propose that this is used to monitor progress through the problem space, by comparing progress against a criterion, in a similar way to what we proposed for the nine-dot problem (MacGregor et al, 2001). In the latter case, we hypothesized that people monitored progress against a criterion defined in terms of the number of dots remaining after a move relative to the number of moves remaining. Thus, for example, at the outset, there are nine dots to be cancelled in four moves, yielding a criterion of 9/4=2.25 dots to be cancelled by the first move. The most commonly-chosen first moves, intersecting three dots, meet this criterion (MacGregor et al, 2001). In the present case we propose an analogous criterion, that with each move the number of coins on the ring should be increased by the difference between the goal state (6 coins on the ring) and the current state, divided by the number of moves remaining. At the outset, there are 4 coins on the ring, which is 2 less than the goal state, and three moves available in which to eliminate this difference. The average increase required per move is therefore 2/3, which yields a criterion for the first move of an increase of 0.67 coins on the ring, or 1 in integer terms. This means that the first move must result in 5 coins on the ring, to meet the criterion of satisfactory progress. (This is assuming that the solver is looking only one move ahead.) If an individual finds a move that meets or surpasses this criterion, then we predict that they will select it. If they cannot find such a move, then "criterion-failure" occurs, initiating relaxation of the requirement to maximize and allowing a search for alternative moves that may lie on the solution path.

Combining the above assumptions regarding a criterion of satisfactory progress with the characteristics of moves available allows predictions about the relative frequency of various moves. Considering first moves, there are 24 legal moves available, and 12 if we ignore symmetry. Of these 12, two result in 5 coins on the ring (and meet the criterion), nine result in 4 on the ring, and one in 3. The mean number of coins on the ring if move selection is random is therefore 4.08 (s.d. 0.49). With reference to Figure 1, the moves resulting in 5 on the ring are to move Coin 1 to touch Coins 4 and 5, or to touch Coins 5 and 6. (The symmetrical moves are to move Coin 4 to touch 1 and 3 or to touch 2 and 3.) Since these moves increase the number of coins on the ring by a margin that surpasses the current criterion, then such moves should be selected by individuals considering only one move ahead. Selecting one of these moves would result in an immediate failure to solve, since they entrap a coin (5 or 2) so that it cannot be moved without violating the non-nudging constraint. Someone looking further ahead might reject a maximizing first move after mentally considering what second moves would then be available. Nevertheless, so long as there are sufficient participants operating at one lookahead, we predict a higher than chance level of first moves that result in five coins in the goal state. This prediction was tested in this experiment.

Correct first moves are possible either by individuals operating at more than one lookahead or, alternatively, by individuals operating at one lookahead on later attempts once they have seen the maximizing moves lead to a dead-end. In either of these cases we anticipate that a strict requirement for maximization would be relaxed, allowing first moves to be sampled that result in 4 coins on the goal configuration, the next highest number possible. Since there are nine such moves, one of which is correct, we predict that correct first moves may occur with a frequency of up to 1 in 9, or 11%, over multiple attempts.

A correct first move results in four coins on the goal configuration, so that the criterion of progress for the second move becomes (6-4)/2=1. That is, an acceptable second move must increase the number of coins in the goal state from four to five. Following a correct first move, there are 23 legal second moves, three of which meet this criterion. These are: move Coin 1 to touch Coins 2 and 5; move Coin 1 to touch Coins 5 and 6; move Coin 5 to touch Coins 1 and 2 (see Figure 1). The first two of these moves will entrap Coin 5. The last is correct. Thus, someone operating at one lookahead has a one in three chance of choosing a correct second move following a correct first. Finally, if a correct second move is made, the criterion for the third move becomes (6-5)/1=1. Only one move meets this criterion, the correct move of Coin 1 to touch Coins 5 and 6. Experiment 1 also tested the prediction that the first move of the problem will be the most difficult, with a probability of success of 11%, followed by the second move, with a probability of success of 33%, while success on the third move should have a probability of 1 once correct first and second moves have been accomplished. In contrast, the conditional probabilities of correct first, second and third moves based on random selection from all possible moves are 8.3%, 4.3% and 4.2%, respectively.

The second version of the problem presents the goal abstractly, in the absence of the ring display, as “each coin must touch exactly two others”. The aim of comparing performance on the ring version of the problem with this abstract version was to examine whether participants’ move selections on problems that lack a concrete and visualizable goal are influenced by the same kinds of evaluation processes used in assessing progress towards known goal states. As previously mentioned, the abstract version is open to both the ring solution and the two-group solution (Figure 1). Participants who attempt the abstract version have both solutions open to them, as both are absolutely consistent with the abstract goal. If participants envisage the goal as a ring of coins, we anticipate that performance should be determined as for the ring condition. However, participants who do not interpret the goal as requiring a ring shape may avoid the conflict raised by attempting to maximize coins on the ring, and may thus be able to discover the alternative solution. Moreover, there are considerably more routes to the alternative solution with the abstract version, and the state space of the problem predicts that the abstract version should be solved more often than the ring version by chance alone. The abstract version of the problem therefore permits an examination of how performance varies in the face of goal uncertainty.

Method

Participants. Forty student volunteers from Lancaster University were quasi-randomly assigned to one of two experimental conditions, "Ring" and "Abstract". In this and all subsequent experiments, age and gender information were not collected.

Materials and Procedure. Participants were tested individually, and their solution attempts were video-recorded. For both conditions, participants were shown the same starting arrangement of coins, and then they read the following instructions "Your task is to rearrange the coins such that each coin touches exactly 2 others. In attempting to solve the problem, you must abide by the following rules: (a) you have THREE moves, no more and no fewer; (b) in each move, slide one coin only (do not pick it up); (c) when you slide a coin, it must not disturb any other coins; (d) at the end of each move, the coin you are sliding must be touching TWO other coins (no more and no fewer)". In the Abstract condition, no further information was provided; in the Ring condition, a drawing of the ring solution was visible. Instructions were available throughout the session. 20 participants were tested in each condition. Participants were allowed up to ten minutes to make as many solution attempts as necessary. On each attempt, the coin array was reset to the start state by the experimenter after the participant had moved three coins. At the end of the ten-minute period, participants in the Abstract condition were asked whether they had envisioned a goal state during problem-solving and the point at which this occurred.

Results and discussion

A problem with the video recording led to the loss of data for one participant from the Ring condition, while one participant in the Abstract condition had seen the six-coin problem previously, during pilot testing. Both were excluded from further analysis.

Of the remaining 19 in the Ring Condition, 6 (32%) found the solution within 10 minutes. The ring solution was hypothesized to be difficult because selecting moves that maximize coins on the ring is incompatible with the correct solution path. The mean number of coins on the ring after the first move was 4.42, which was significantly higher than the chance mean of 4.08, $t(18)=3.02$, $p<.01$. The results also provide a close fit with the predicted frequencies of successive correct moves, further corroborating our hill-climbing account of move selection with this problem. There were a total of 177 first moves in the Ring condition, of which 20 (11.3%) were correct, very close to the predicted probability of 11%. Of these 20 correct first moves, 6 (30%) were followed by a correct second move. Again, this corresponds very closely to the predicted probability of 33%. Of the 6 correct second moves, all six (100%) followed with a correct third move, exactly as predicted.

In the Ring condition, the predicted probability of success on an attempt was .036 (.11x.33x1), assuming a hill-climbing heuristic. Combining this probability across multiple attempts as a series of Bernoulli trials, the proportion of participants expected to solve within *n* attempts would be $1\text{-}(1\text{-}.036)^n$. Substituting for *n* the observed average number of solution attempts, 9.3, this evaluates to 29% of the 19 participants, or 5.5. If the calculation is based on the average solution attempts of non-solvers, which was 10, the predicted value becomes 5.9. Neither of these predicted values was significantly different from the 6 observed successes. In contrast, substituting the chance probability of success of .00015, the expected number solving within either 9.3 or 10 attempts was virtually zero (.03). The obtained number, 6, was significantly greater than this, by the binomial test ($p<.001$). Clearly, solutions were governed by intentional rather than chance processes.

The inferred goals reported by participants in the Abstract condition were categorized. Seventy-four percent inferred a ring-shaped solution. In spite of this, no ring solutions were found in the Abstract condition. This was significantly fewer than in the Ring condition, $p<.01$ by the Fisher exact test. Perhaps the failure to find the ring solution in the Abstract condition was because their degree of goal certainty was lower, since it was an inference and not a given. In addition, since those in the Abstract condition were not restricted to one solution, they may have been more willing to explore other avenues, discovering the two-group solution in the process, either by chance or design. It should be noted that, in some move sequences such as that shown on the right of Figure 1, a two-group arrangement that satisfies the goal requirements of the Abstract condition appears after two moves. Although this is not counted as a solution (because it uses too few moves), it may be that its emergence in attempts at the problem could contribute to discovery of the legal two-group solution.

The results do not, however, refute the possibility that solutions appeared by chance. On average, participants made 6.5 solution attempts in this condition, with 5 of the 19 participants (26%) discovering the two-group solution. Using a 1-sample binomial test, this was not significantly greater than 2.2, the number expected to solve by chance (calculated by substituting 6.5 for n and .0187 for p in the previous formula: basing the calculation on the number of attempts of non-solvers increases the predicted number only slightly, to 2.6). With the observed proportion of solutions being 5/19, the proportion expected on the basis of chance being 2.2/19, and a conventional criterion for significance of 0.05, the statistical power of the binomial test is modest, at 0.38. The result should therefore be viewed with some caution. Nonetheless, it is consistent with the interpretation that some of the successes in this condition could have occurred randomly. (Of note, two participants in the Ring condition produced the two-group arrangement, presumably by chance, since it was not a valid solution in that condition.)

In order to check that the sample of participants was not becoming contaminated over the course of the experiment by word-of-mouth information about solutions to the problems, the number of solutions of the first 9 participants in sequence was compared with the last 10 participants. In the Ring condition, the numbers were 4 and 2, and in the Abstract condition, 2 and 3. It seems unlikely that there was any such contamination.

The results for the Abstract condition raise the question of why the two-group solution occurs no more often than chance (bearing in mind the earlier caveat concerning statistical power). There was an indication that the two-group solution required not only a physical separation of the pieces but a psychological discontinuity in the types of move made. The relative proportion of states in the population of all possible final states that result in separate groups of coins is quite high, at 16.1%. The observed frequency of final states that separated the coins was significantly lower than this, at 4.8% (that is, of a total of 124 final states across all solution attempts, only 6 exhibited two separate groups of coins), χ^2 =11.68. Similarly, in the Ring condition only 7 out of 177 observed final states (4.0%) separated the coins. What is not clear at this point is whether this apparent resistance to creating two groups of coins is a consequence of inferring the single-figure ring solution, or if it is a separate constraint in its own right. We return to this issue in the second experiment.

If the two-group solution involves a discontinuity then it is a candidate insight problem, in Weisberg's (1996) scheme. To confirm its status, what would be required in addition is that the discontinuity involves restructuring. In the Abstract condition, participants tended to think of the solution as a single ring, whereas the two-group solution is, effectively, *two* rings. The cognitive shift required to move from a single ring to a double ring representation could be considered an instance of restructuring.

The experiment also collected information on the reproducibility of the ring solution. First, the six participants in the Ring condition who found the ring solution were immediately asked to reproduce it, and allowed one minute to do so. Only two succeeded. Second, the ring solution was demonstrated to all the participants in the

Abstract Condition immediately following their procedure. They then performed a filler task for six minutes, following which they were given the six coin problem again and given one minute to produce the ring solution. Only 6 of the 19 participants were able to do so, and required a mean of two attempts. The thirteen who failed were shown the solution a second time, and asked to reproduce it immediately. Six failed to do so within one minute. The solution was demonstrated to these six for a third time, and again they were invited to repeat it. Within the one minute allowed, one succeeded while five failed. Third, the 14 participants who were able to reproduce the solution at some point were shown the problem again, but with the starting state in reverse orientation. Five (36%) were unable to reproduce the solution. The nine who did reproduce it successfully did not do so immediately, but required a mean of 2.2 attempts. These data suggest that participants find it difficult to remember the ring solution to the six-coin problem, even after several demonstrations.

Experiment 2

Participants in both conditions of the first experiment rarely made moves that separated the configuration of coins, which may have hindered those in the Abstract condition from discovering the two-group solution. This apparent reluctance is explained by our theoretical approach as resulting from selecting moves to make progress towards a given or inferred ring goal. Moves that separate the configuration are avoided because they generally result in a decrease rather than increase in the number of coins on the ring. This reasoning applies equally to the Abstract condition, since the majority of participants in that condition conceived of the solution as a ring. In essence, the conception of the goal as a ring creates an implicit conceptual block that precludes exploration of the areas of the problem-space within which the two-group solution lies. The hypothesis bears some resemblance to the concept of "fixation"— the adherence to an inappropriate representation of a problem that blocks insight (Dominowski & Dallob, 1996; Smith, 1996). In the case of the six-coin problem, however, the ring hypothesis is quite appropriate in the sense that it is a correct solution, although it conflicts with the other solution.

However, there are several other possible explanations for the low frequency of moves that separated the coin configurations in Experiment 1. One is that there may be a more general tendency to avoid decomposing chunks into less conceptually or visually coherent groups, as has been clearly shown in a different context (Knoblich, Ohlsson, Haider & Rhenius, 1999). Another is that participants may seek solutions within dimensions that are bounded by the initial problem presentation, in this case single composite coin figures.

Experiment 2 was designed primarily to test among these alternatives. The two initial states used in Experiment 2 are shown in Figure 2, and participants were restricted to finding solutions in two moves only. In both cases, the two-group solution can be achieved in two moves while the ring solution is impossible. Each of the two figures can be decomposed into two parts by moving any of the four interior coins, and so are equally likely to be separated by chance. They appear to form approximately equally "good" figures and should therefore equally resist decomposition. Both are unitary, and should equally constrain moves to other unitary figures. The two states were not equivalent in their chance probabilities of success, which were .055 for the Partial-ring and

.041 for the Straight-line. However, we conjectured that a representation of the goal as a ring would be strengthened in the Partial-ring condition and weakened in the Straight-line condition. Consequently, there should be greater perseverance in pursuing a ring-like solution in the Partial-ring condition, making the two-group solution relatively less available. We therefore predicted a relatively greater number of two-group solutions in the Straight-line condition than in the Partial-ring condition, contrary to the predictions based on chance alone.

Method

Participants. The participants were 54 final-year high school student volunteers visiting Lancaster University. Participants were assigned randomly in equal numbers to either the Partial-ring condition or the Straight-line condition.

Materials and Procedure. Testing was conducted in a group setting. Participants received the same instructions as the Abstract condition of Experiment 1, except that the number of moves allowed was limited to two (thereby excluding ring solutions). Each participant received an envelope containing six UK penny coins, a pen and a sheet of paper showing a template of the initial state. They were instructed to place the coins on the template. They were allowed to make as many solution attempts as they wished within a time period of three minutes. The time period was reduced from that in Experiment 1, as pilot work had suggested that the problems would be easier. At the end of three minutes, participants were asked to draw the shape of their solution.

Results and Discussion

The present experiment manipulated the initial layout of the coins in a manner that created a partial ring in one condition but not in the other. We anticipated that this would reinforce a ring interpretation of the goal state more strongly in the Partial-ring than in the Straight-line condition and thus make the two-group solution less accessible, even though that solution was attainable in each condition in two moves. The results supported this expectation. Four participants (15%) in the Partial-ring condition and 11(41%) in the Straight-line condition produced a correct solution. The difference was significant, $\chi^2(1, 54) = 4.52$, $p<.05$. The procedure did not allow recording of the number of attempts, but if we base an estimate on the rate of attempts from the Abstract condition of Experiment 1, then the chance number of solutions for the Straight-line condition is 2.1 solutions. The observed frequency, 11, was significantly greater than this, by the binomial test ($p<.001$). In contrast, the observed number in the Partial-ring condition, 4, was not significantly greater than the expected number of chance solutions, of 2.8.

The results indicate that participants in the Straight-line condition were sufficiently liberated from the ring hypothesis that they were able to discover the alternative solution at greater than chance levels. In contrast, participants in the Partial-ring condition, like those in the Abstract condition of Experiment 1, did not find the two-group solution significantly more often than chance. These outcomes support the hypothesis that participants in the previous experiment failed to produce moves that separated the coins because of their focus on a single ring solution. This fixation would have prevented a wider exploration of the problem space that could have led to the two-group solution. This leads to the parsimonious conclusion that the difficulty of both the two-group solution and the ring solution arise from the pursuit of a hill-climbing strategy towards the ring goal. The experiment also provides evidence that, in the absence of a concrete and visualizable goal, individuals make hypotheses about properties of the goal state and use these hypothesized properties to derive a test for satisfactory progress against which to evaluate alternative moves.

Experiment 3

In Experiment 3, we manipulated the goal information given to participants. The experiment employed four conditions and, in all four, participants were instructed that the correct solution resulted in each coin touching exactly two others, and that <u>two different solutions</u> were possible. In the first condition (Abstract), no further information was given. In the second (Ring example) the ring solution was shown, in the third (Two-group example) the two-group solution was shown and, in the fourth (Both examples) both solutions were shown.

The experiment addressed several issues. First, we remedied a number of methodological inequalities of Experiment 1, notably holding the space of possible solutions constant across conditions. Second, the experiment explored whether the difficulty of the two-group solution resides wholly in forming a representation of the goal state properties or if some component of difficulty arises in executing the required sequence of moves (as in the ring solution, where executing the moves is the sole source of difficulty). If the difficulty is representational, then two-group solution rates would be close to 100% in the Two-group and Both-example conditions, and significantly higher than in the Abstract and Ring-example conditions. If the difficulty lies more in move execution, then performance should be constant across all four conditions. Third, the experiment investigated the effects on two-group solution rates of providing the ring goal. Does seeing, as opposed to inferring, the ring goal make a difference, when knowledge that alternative solutions are present is held constant? Fourth, the preceding distinction between difficulty of representation and difficulty of execution suggested an additional test. If the move sequence is difficult to execute, we envisaged that it might be difficult to reproduce, once discovered. By contrast, if an adequate unitary representation of the goal state were achieved during the course of a successful attempt at the problem, the solution might be easily reproducible. The experiment therefore collected information on the reproducibility of both types of solution.

Method

Participants. 42 students from Lancaster University, majoring in subjects other than psychology, were randomly assigned to one of four experimental conditions. Participants were paid two pounds sterling.

Materials and procedure. The materials and procedure were similar to those used in Experiment 1. The penny coins used in Experiment 1 were replaced with steel regular hexagons, with length of side of 15mm and thickness 3mm. This change was made because hexagons make it easier for participants to evaluate the number of mutual contacts. For convenience, henceforth we refer to these hexagons as coins. Participants read the instructions of Experiment 1 with the additional line: “There are two general types of solution to this problem, both of which are acceptable”. In the Abstract condition, no further instruction was provided. In the Ring-example condition, a drawing of the ring solution was provided; in the Two-group example condition, a drawing of the two-group solution was provided, and in the Both-examples condition, both drawings were provided. As in Experiment 1, participants were allowed up to ten minutes to make as many solution attempt as they wished. At the end of the procedure, participants in the Abstract condition were asked whether they had any image of the goal state in mind during their solution attempts. Finally, participants in the Abstract and Ring example conditions who produced the two-group solution were asked immediately to reproduce it from a starting state that was reoriented by 180 degrees from the original. In addition, the ring solution was demonstrated to the participants from all conditions who had not produced it, and they were asked immediately to reproduce it from the same starting state.

Results and discussion

The numbers (percentages) finding the two-group solution were 3 (27%), 3 (27%), 8 (89%) and 10 (91%) for the Abstract, Ring-example, Two-Group example and Both-examples conditions, respectively. The difference across conditions was significant, χ^2 (3,42) =16.84; $p<0.015$. In addition, solution rates were close to 100% in conditions where the two-group example was shown, with participants finding the solution on the second attempt on average. (In both conditions, the single participant who failed persisted in attempting to produce a ring solution) The results demonstrate that the difficulty of the two-group solution resides in establishing an appropriate representation of the goal and not in executing the necessary steps to solution. When the goal is provided, the problem becomes relatively trivial.

Because there are relatively many paths that lead to the two-group solution we compared the observed results with chance. The expected numbers of two-group solutions by chance alone (assuming a single attempt) were 1.6, 1.5, 0.5 and 0.7 for the four conditions, respectively (calculated as in Experiment 1). The corresponding observed frequencies were 3, 3, 8 and 10. The former two observed values were not significantly different from chance ($p>.20$), while the latter two were ($p<.001$), by the binomial test. Given that the observed frequencies resulted from multiple attempts, the comparison with chance is a conservative one. The result is consistent with the hypothesis that in the Abstract and Ring-example conditions participants may simply have stumbled upon the two-group solution by chance. This interpretation is supported by the move selections observed in these conditions, as described below.

There was evidence that participants in the Abstract condition entertained a ring shaped goal. Again, the majority of participants (8, or 73%) in the Abstract condition inferred that the solution had a circular form. As in Experiment 1, we expected that when participants were shown the ring example their initial attempts would exhibit a tendency to increase the number of coins on the ring, and this was the case. Analyzing first move data, this tendency was equally evident in the Abstract condition, consistent with participants' reported goal inference. The means (standard deviations) of the number of coins on the ring after the first move were 4.55 (0.52), 4.50 (0.53), 3.89 (0.33) and 4.18 (0.40), for the four conditions respectively. There was an overall difference among the means, $F(3,37) = 4.35$, $MSe = 0.21$, $p =.01$. Post-hoc comparisons using Tamhane's T2 procedure (because of heterogeneity of variance) indicated that the Abstract and Ring-example conditions had significantly higher mean numbers on the ring than the Two-group example condition.

Analysis of the times to reach the two-group solution provided information about the solution processes in the different conditions. The mean solution times (s) for those solving the problem only were 221, 417, 90, and 144, for the Abstract, Ring-, Two-Group- and Both–examples conditions, respectively. A 2x2 ANOVA on solution times, with presence/absence of ring example and presence/absence of the two-group example as independent variables, resulted in significant main effects for both factors. The results showed a significant facilitation in the time required to find the two-group solution with the presence of the two-group example, $F(1,20)=18.08$, $p<.001$, $MSe= 10156$, and a significant inhibition with the presence of the ring example, $F(1,20)=6.92$, $p<.02$, $MSe= 10156$. The interaction effect was not significant. The results suggest that the effects of the ring strategy are stronger where the ring goal is presented (rather than inferred) and that participants will pursue it for longer before considering an alternative solution.

Five participants who produced the two-group solution in the Abstract and Ring-example conditions (those who had not seen the two-group goal) were able to reproduce it, three on the first attempt and two on the second (the data for one participant was missing because of a procedural error). Of the five reproduced solutions, four involved a sequence of moves that differed from the original solution. The result is instructive in two ways. First, although some or even all of these solutions may have occurred by chance, participants were apparently able to quickly recode the solution, since they were able to reproduce it. Second, this recoding was not simply a trace of the previous moves, since the majority of reproductions followed different solution paths and resulted in different orientations of the two group solution. It seems probable that participants recoded the general configuration of the solution and then reconstructed paths that led to it.

The ring solution was demonstrated to the 36 participants who had not produced it. Only seven (19%) were then able to reproduce the solution within a one minute time period. This finding illustrates the difficulty of recoding the required sequence of moves for the ring solution, in sharp contrast with the ease with which moves enabling a two-group solution appear to be recoded.

Experiment 4

Experiments 1-3 established that performance on variants of the six-coin transformation problem is under the control of a hill-climbing heuristic, in which moves are evaluated for selection on the basis of their fit with a criterion for progress towards a known or hypothesized goal. In Experiment 4 we investigated whether the same hill-climbing heuristic determines performance on the ten-coin or 'triangle' problem, widely recognized in the literature as one requiring insight. A reason for selecting the ten-coin problem for comparison with the six-coin problem is that, status as insight or non-insight problem aside, the two problems appear similar superficially, and they can be used with identical instructional constraints.

The initial and goal states of the ten-coin problem are shown in Figure 3. As with the six-coin problem, the task can be stated in terms of transforming a starting state to a goal state by moving three coins, one at a time, under the constraints that a coin being moved (i) must not be lifted, (ii) must come to rest touching exactly two other coins and

(iii) must not displace any other coin. As before, we determined the state space of the ten-coin problem. There are 81147 sequences of legal moves, as compared to 7426 for the six-coin problem. 36 sequences lead to the goal state. The overall chance probabilities of finding a correct solution, calculated in both cases as the sum of the products of the conditional probabilities of each sequence of first, second and third moves that lead to a correct solution, are .00041 for the ten-coin and .00015 for the six-coin problem.

A number of researchers have proposed that the ten-coin problem requires insight for its solution (Metcalfe, 1986; Schooler *et al*, 1993). Metcalfe describes the insight as restructuring the triangle of coins into a central rosette of 7 coins around which the 3 corner coins may be rotated, leading directly to the solution shown in Figure 3. An alternative account of the problem's solution is offered by Weisberg (1996), who proposes that the problem may be solved without insight, using trial and error or other processes. It remains the case that the role of a rotational insight as a precursor to solution, and the conditions under which such an insight might arise, have yet to be tested against trial-and-error and other accounts.

Whether solved through insight or not, the concrete goal and well-defined operators of the ten-coin problem suggests that it may be addressed in a similar manner to the six-coin problem and that, perhaps prior to any rotational insight, move selection may be governed by simple locally-rational progress evaluation criteria. What constitutes reasonable progress will depend on how each participant conceptualizes the goal of the ten-coin problem, specifically the hypotheses that they develop regarding the properties of the path towards the goal state, and there are several possibilities. For example, a participant might translate the goal into a requirement to transform the apex of the triangle into the base. Given such a conception, a reasonable approach would be to move a corner coin from the base to touch the apex coin (with reference to Figure 3, move Coin 7 to touch Coins 1 and 2, or 1 and 3, or move Coin 10 to touch Coins 1 and 3 or 1 and 2). An alternative conception might maintain the current base and translate the rest of the figure across it, shifting the apex from its current location to the bottom of the array by moving Coin 1 to touch Coins 7 and 8, 8 and 9, or 9 and 10. A third conception might direct attention towards rows that appear to have too many adjacent coins (the bottom row) or too few (the top and second rows). This representation could result in a number of locally rational first moves. These include the moves described for the first goal conception, since they reduce the number of coins on the bottom row and increase the number on the top. Alternatively, moving coins 7 or 10 to touch coin 2 and 4 or coins 3 and 6 decreases the bottom row and increases the second row. Of the 10 moves identified for the three goal conceptions described above, 5 are correct and could result in success on a first attempt, while 5 are incorrect and will lead to failure on that attempt.

Given that there appear to be several possible conceptions of goal properties in the ten-coin problem and a relatively large number of moves stemming from them, we will not attempt to elaborate related criteria of progress, though evidence regarding the existence of different goal conceptions is reported below. While many of the predicted first moves are incorrect, all of them involve moving a coin that has to be moved in the correct solution path. This suggests that, while the ten-coin problem may be quite difficult to solve on a first attempt, it may be relatively amenable to solution across multiple attempts. In contrast, the goal conception identified for the six-coin problem (ring goal) predicts first moves that lead to failure. Even if a wrong first move is rejected on subsequent attempts, the probability of selecting a correct first move from the next best alternatives is only 11%. This leads to the prediction that the six-coin problem will be extremely difficult to solve on a first attempt and will continue to be relatively difficult over subsequent attempts. Overall, the ten-coin problem should be relatively easier than the six-coin, both on the first and on subsequent attempts.

Accounts that invoke insight as a precursor to solution (e.g., Metcalfe, 1986), might suggest that the ten-coin problem should be more difficult than the six-coin problem. Restructuring of some kind – perhaps to identify the central "rosette" – is held to be necessary in the ten-coin problem, but no such restructuring seems possible in the six-coin problem, nor is there any evidence for it in the foregoing experiments. In contrast, a trial-and-error account allows the prediction that the six-coin problem should be no more difficult than the ten-coin problem: the chance probabilities of finding a correct solution are diminishingly small for both problems. Thus, the comparison between six- and ten-coin problems in this experiment provides a strong test of our theoretical predictions against a other accounts.

Method

Participants. The participants were 50 student volunteers from Lancaster University, each paid two pounds sterling.

Materials and procedure. The six-coin condition used similar materials and procedures to the Ring condition in Experiment 1. The start and goal states of the ten-coin problem are shown in Figure 3. Participants were instructed as follows: "Show how you can make the triangle of 10 coins…point downward by moving only three of the coins. You must abide by the following rules: You have three moves, no more and no fewer. In each move, slide one coin only (do not pick it up). When you slide a coin, it must not disturb any other coins. At the end of each move, the coin you are sliding must be touching two other coins (no more and no fewer)." In order to keep the total amount of time spent on problem-solving comparable to that in Experiments 1 and 3, and to avoid any confound with fatigue, participants were allowed 5 minutes for each of the two problems (10 minutes total). Participants were tested individually on both the six-coin problem and the ten-coin problem. Half received the six-coin problem first, half the ten-coin first. If a correct solution was given the participant was invited to reproduce it immediately, and was allowed one minute in which to do so. All moves were recorded on videotape.

Results and discussion

The data for two participants, one from each order assignment, were unusable. The results for the remaining 48 were used to test the experimental predictions.

The prediction that the six-coin problem will be more difficult to solve than the ten-coin problem on first and subsequent attempts, was confirmed. The numbers (percentages) solving at the first attempt were zero (0%) and 9 (19%), Wilcoxon $z = 2.65$, $p =.008$, and solving within 5 minutes were 10 (20.8%) and 36 (75%), Wilcoxon $z = 2.67$, $p = .008$, for the six-coin and ten-coin problems, respectively. As in Experiment 1, the numbers of solutions to the six-coin problem in the first and second half of the experiment were examined to ensure that no contamination of the participant sample had taken place. Five solutions were found by the first 24 subjects, and five by the second 24.

The greater difficulty of the six-coin problem supports a hill-climbing rather than trial-and-error account of solution. One possible caveat is that the instructions used for the ten-coin problem imposed additional constraints (the requirement that the coin being moved must come to rest touching exactly two other coins). Ormerod and Gross (2003) tested participants with standard instructions to the ten-coin problem and found that 65% solved within the same time period of five minutes. It does not appear, therefore, that the additional constraints radically altered the problem's difficulty. Rather, it is notable that the clear and significant difference in performance between the six- and ten-coin problems occurred in a situation where care was taken to equate instructions and move constraints.

The first moves made on each participant's first attempt were analyzed, on the grounds that these provide the clearest evidence for the influence of goal conception on move selection, unaffected by outcome. For the six-coin problem, first moves were analyzed for the number of adjacent coins on a ring that they produced and compared with the population distribution of all possible first moves. The expected frequency of moves resulting in 5, 4 or 3 coins on a ring, are 17%, 75% and 8%, respectively, with a mean of 4.08 and a standard deviation of .49. The corresponding obtained frequencies were 50%, 40% and 10%, with a sample mean and standard deviation of 4.40 and 0.68, respectively. The obtained sample mean was significantly higher than the theoretical population mean, $t(47) = 4.53$, $p<.01$, indicating a greater than chance preference for maximizing the number of coins on the ring. These results are consistent with those of Experiment 1, and again suggest that move selection is determined by the operation of a hill-climbing heuristic. Again, the probabilities of correct moves across multiple attempts were close to the predicted values. Of the 352 first moves, 35 (10%) were correct (predicted value 11%). Of these, 25 were followed by a second move, of which 9 (36%)

were correct (predicted value 33%). Of the nine correct second moves, all (100%) were followed by a correct third move (predicted value 100%). Based on these predicted probabilities, the expected number solving within the mean number of observed attempts was 11.3, which was not significantly different from the observed value of 10.

First moves in the ten-coin problem were also examined for consistency with the application of a hill-climbing heuristic to the different goal conceptions identified earlier. First move data from six participants could not be unambiguously transcribed from video. The results for the remaining 42 showed that the number (percentage) selecting first moves consistent with each of the three goal conceptions were 13 (31%), 11(26%) and 5 (12%), respectively. This yielded a total of 29 (69%) first moves consistent with a hill-climbing heuristic, which was significantly higher than the 7.33 (17%) based on a chance selection of first moves from all available moves in the state space, $\chi^2(1, 42) = 77.61$, $p<.01$. Thus, the analysis of first moves yields direct evidence of the kinds of goal conceptions we had hypothesized.

In the case of the 36 participants who solved the ten-coin problem, we examined the successful sequence of three moves for evidence of the “rotational insight” that has been attributed to the problem. We defined this as any sequence of moves where each of the three corner coins was shifted in order one position around the central rosette, as shown in Figure 3, in a clockwise or counter-clockwise direction. Thirty-four of the successful trials were unambiguously classifiable and of these 8 (24%) exhibited this rotational pattern. In contrast, 19 (56%) were consistent with the application of hill-climbing to a goal conception of trying to translate the figure across a horizontal median axis, either by moving the top coin to the center of the bottom row, followed by moving the two lower flanking coins up, or the same moves in reverse order. In his taxonomy of problem types, Weisberg (1996) classified the ten-coin problem as a "hybrid", meaning that it could in principle be solved either through insight or other means. The present results appear to support Weisberg's interpretation, since only 24% of people solving did so with a sequence of moves that were completely consistent with the "rotational" insight. More solvers (56%) used moves consistent with a hill-climbing heuristic.

In addition, the experiment collected information on the reproducibility of solutions to both problems. Of the ten solving the six-coin problem, only 2 (20%) were able to replicate the solution on an immediately subsequent attempt. Of the 36 who solved the ten-coin problem, 35 (97%) succeeded in replicating the solution immediately. We return to these data in the General Discussion, where we propose that they hold the key to understanding a major remaining difference between the ten-coin and six-coin problems, and one that may lie at the heart of a useful definition of insight.

General Discussion

This paper addressed the potential for an information-processing approach to demonstrate and account for commonality in the strategies used by participants solving knowledge-lean problems of both transformation and insight types. It did so in three ways. First, it introduced the six-coin problem, which is configurable to reflect characteristics attributed both to transformation and to insight problems. Second, it explored how individuals identify the information about goal properties necessary to implement a hill-climbing heuristic when problems have under-specified goals. Third, it tested predictions derived from a hill-climbing approach in a direct comparison of transformation and insight problems.

Experiment 1 gathered baseline information on two versions of the six-coin problem, one in which a ring goal was shown to participants, the other in which the goal was described only in abstract terms as "each coin touching exactly two others". While not quite as difficult as the nine-dot problem (e.g., MacGregor et al, 2001), both versions of the problem are challenging, with less than one-third of participants solving either version within 10 minutes. Performances with the ring version of the problem supported the hypothesis of a hill-climbing heuristic, in that (a) first moves maximized the number of coins on the ring significantly more than chance and (b) the observed probabilities of correct first, second and third moves corresponded to those predicted by hill-climbing.

Given the difficulty of the problem, how were participants ever successful in solving the six-coin problem? If participants discovered the correct first move, then they were likely to repeat it and eventually solve. The correct first move appears to be a 'promising state' (cf. MacGregor et al, 2001), in that it allows the discovery of second moves that make progress against the goal in a way that the second moves following all other first moves do not.

While the majority of participants who received the abstract version of the six-coin problem reported holding a ring hypothesis, none found that solution. Some found the alternate two-group solution. The number of such solutions was not significantly greater than chance, though the power of the test was admittedly low. One source of difficulty in finding the two-group solution appeared to be a reluctance to select moves that split the coin array. This difficulty may be explained in a number of ways, such as the possibility that participants are unwilling to decompose the ‘chunk’ that the coins form into less conceptually coherent groups, as has been clearly demonstrated with other stimuli (Knoblich *et al*, 1999). The account deriving from a hill-climbing model is that the relative absence of splitting moves was primarily a by-product of participants inferring and seeking the ring solution. The latter explanation was supported by Experiment 2, which manipulated the perceptual salience of the ring goal while holding the figural coherence of the starting state constant. The results demonstrated that providing participants with a visual cue to reinforce a ring hypothesis inhibited the discovery of the two-group solution, the only solution possible in two moves.

Experiment 3 manipulated the presence of ring and two-group solution examples, while holding key features of the state space constant across conditions. The important findings of this experiment were twofold. First, participants who received the two-group example were generally able to find the two-group solution, in contrast to participants who did not. This finding supports our hypothesis that the key source of problem difficulty with this version of the problem lies in conceiving of the two-group solution, not in executing it. In contrast, few participants who received the ring example were able to find the ring solution, suggesting that the problem with this version lies in executing the ring solution, not conceiving it. The fact that the two-group example gave rise to many more correct solutions than the ring example confirms the hypothesis that a two-group display yielded different and less inhibiting local goal conceptions than the ring display. The application of hill-climbing to the goal conception elicited by the ring example (maximize coins on the ring) initially inhibits selection of the correct move, whereas the same heuristic applied to the goal conception elicited by the two-group example does not appear to inhibit correct moves. The second important outcome of the experiment was that presenting participants with the ring example slowed down the discovery of two-group solutions. This indicated greater perseverance on the ring goal when it was presented, as opposed to inferred, even with the knowledge that an alternative solution was available.

In Experiment 4 we compared performance on the six-coin problem with the ten-coin problem, a well-known problem characterized as necessitating insight by many researchers. As predicted, the six-coin problem was considerably more difficult to solve than the ten-coin problem, despite the apparent absence of a requirement for insight as a precursor to solution, and despite the fact that it has a much smaller state space. There was evidence that first moves in the ten-coin problem were determined by the same hill-climbing heuristic as in the six-coin problem. Nearly 70% of first moves conformed to the application of hill-climbing to the three goal conceptions described earlier. Evidence that a 'rotate around a rosette' insight was a necessary precursor to solution was slight, only 24% of solutions reflecting a move sequence consistent with such an insight. Instead, over 56% of solutions conformed to a non-insightful but nonetheless successful application of hill-climbing. The result supports Weisberg's (1996) contention that the problem is a hybrid rather than a pure insight problem, and that solutions may be found either through restructuring from a lateral to a rotational view of the movement of coins, or through other means. However, the results are slightly inconsistent with Weisberg's taxonomy, in that the majority of solutions did not appear to involve any discontinuity in approach. Rather, the ten-coin problem appears to be an extreme type of hybrid, which can be solved either through the continuous application of a hill-climbing heuristic or through restructuring. Interestingly, Isaak & Just (1996) have also claimed that the ten-coin problem requires an insight for solution, but the illustration they provide of the moves required to solve (p.320) is not the "rotational insight" solution but is in fact identical to the application of hill-climbing to one of the goal conceptions we described above. The apparent inconsistency highlights again the definitional difficulties and the absence of clear criteria for specifying insight solutions. However, as discussed below, the present results suggested one new factor that could be helpful in distinguishing between insight and non-insight problem-solving.

Although the evidence reviewed so far suggests that performance on the ten-coin problem was governed in many ways by the same kind of move selection heuristic as the six-coin problem, performance did differ on the two problems in one important respect: only 2 of the 10 participants who solved the six-coin problem in Experiment 4 were able to recreate their successful solution on the subsequent attempt. In contrast, 35 of the 36 participants who solved the ten-coin problem replicated their solution immediately. We propose that participants identified some kind of solution principle for the ten-coin problem that allowed them to recreate the solution without a requirement to remember a sequence of moves in its entirety. A number of solution principles might serve this purpose, including the 'rotate around a rosette insight' that has been attributed to the problem (Metcalfe & Wiebe, 1987). In contrast, it is difficult to conceive of a principle that captures in a single clause the path to solution of the six-coin problem. We propose that, while both problems are solved initially by the discovery of a sequence of moves selected under the application and subsequent relaxation of a hill-climbing heuristic, the ten-coin problem can be reproduced because its solution can be described as a single executable concept, whereas the solution to the six-coin problem cannot easily be reproduced because it is not amenable to a simple recoding. Insight into a single, executable concept may occur either prospectively, where it guides the solution, or retrospectively, we propose, as the rapid recoding of a solution that has been revealed through other processes. The latter may include hill-climbing, chance or demonstration.

Further evidence in support of the role played by solution recoding comes from the solution reproduction data of Experiment 3. Participants from conditions in which the two-group example was not shown but who nonetheless discovered the two-group solution were able to reproduce the solution despite a 180 degree inversion of the start state. Moreover, typically they reproduced the solution using a different sequence of moves, suggesting a conceptual rather than sequential encoding of the solution. In contrast, participants who saw a demonstration of the ring solution were unable to reproduce it from an identical start state. The procedure of Experiment 2 precluded collecting data on solution reproducibility. Instead, 14 additional participants were tested using the partial-ring condition only. After the solution had been demonstrated once, all fourteen were able to reproduce it without error from a start state inverted through 180 degrees. Thus, while the same ‘ring’ goal conception inhibited solutions in both the three-move ring problem of Experiment 1 and the two-move partial-ring problem of Experiment 2, its inhibiting effects on solution reproducibility appear to have been overcome by solution recoding in the latter but not in the former problem.

In contrast to the ring version of the six-coin problem, anecdotal reports given by five participants who took part in experiments on the eight-coin problem (Ormerod et al, 2002) indicate that they remembered the necessary “insight” to move coins in three dimensions many months after participating in our studies. Moreover, they successfully reproduced the solution at their first attempt. Knoblich et al (1999) report a similar ability of participants to remember, and to transfer to new problems, a conceptual insight into the solution principles underlying a range of matchstick algebra problems.

The present results raise a number of issues about the nature of insight problems and about the usefulness of the defining criteria that have been proposed. A summary is provided in Table 1. The first column of the table provides some of the criteria that have been proposed for insight, together with the criterion identified here, of solution recoding. The remaining columns of the table summarize the corresponding results obtained here for the three problems, and contrast them with the profile that would be expected for the "ideal" insight problem.

Although we did not formally measure the step-function emergence of solutions, some clues were provided by the empirical probabilities of correct moves across the first, second and third moves of each problem. These were obtained as described in the Results section of Experiments 1 and 4, and provided, for the three problems, the observed conditional probabilities of correct first, second and third moves. The results for the six-coin ring problem were 0.10, 0.33 and 1.00 for the three moves, respectively (Experiment 4 data). The combined probability of executing a correct first and second move was therefore very low, at 0.03, but if a person did so, the probability of finding the correct third move -- and solving the problem -- jumped to 1.00. This suggests a step-function component to the solution process, where the solution was suddenly obvious after correctly executing the first two moves. By the same measure, the conditional probabilities of correct moves showed no step-function pattern for the two other problems. For the abstract six-coin two-group solution the probabilities of correct moves were relatively flat across the three moves, at 0.57, 0.44 and 0.56 respectively (Experiment 1 data). For the ten-coin problem, the probabilities increased across moves, but in a progressive fashion, at 0.36, 0.60 and 0.95. (This pattern was the same whether or not participants showed the rotational insight.) Therefore, if we employed this step-function criterion alone, the conclusion would be that only the six-coin ring is an insight problem.

In terms of discontinuity, the six-coin ring solution does not appear to require a change in the moves that are sampled. In contrast, the two-group solution demonstrates a clear discontinuity in the resulting array, if not in the sampled moves themselves, whether we interpret it as relinquishing the ring hypothesis, or of separating the pieces into two groups. The ten-coin problem is a mixed case—the majority of solutions indicated no discontinuity while a minority did. The conclusion by this first of Weisberg's criteria is therefore that the six-coin ring version is not an insight problem and the ten-coin problem was not an insight problem for the majority of participants.

It seems clear that the two-group solution to the abstract six-coin problem may potentially involve restructuring if the representation of the solution changed from a single ring to two separate rings. However, we have no evidence that anyone ever found the solution in this way, while the results suggest that some, perhaps all, solutions came about by chance. This classifies the problem as "hybrid", in Weisberg's taxonomy. Similarly, the ten-coin problem can be solved by restructuring, if a shift in focus occurs from translating coins across a lateral axis to rotating them around a central axis. In this case, there was evidence that participants solved in both ways, making this problem a "hybrid" also, since it was solved both by restructuring and by other means.

The weight of evidence suggests that the six-coin ring is not an insight problem, and that, while the others have the potential to be insight problems, they may be solved by other means. The results provided empirical support for Weisberg's taxonomy, and caution against the use of any single criterion in diagnosing insight problems and problem-solving. The lack of commonality among any of the problems reviewed in Table 1 suggests that defining insight problems purely on phenomenology is of limited value. Moreover, the absence of any obvious conceptual change that might lead to solution of the ring-version of the six-coin problem, along with limited evidence of solutions to the ten-coin problem consistent with a conceptual change, raise doubts about the generality of definitions based upon conceptual restructuring alone.

We have tentatively introduced a new process-related criterion for insight, based on the recoding of a solution, which is distinct from the defining criteria previously proposed for insight. From the analysis offered in Table 1, solution recoding is not associated with a consistent combination of matches against other criteria for the problems reported here. An important distinction between this account and the traditional Gestalt account of insight is that the emergence of a new conceptual principle is not necessarily the precursor to solution: instead, such principles may be a product of solution discovery that enables future reproduction of solutions without extensive search for moves. It is evident that solution recoding has not yet been examined with problems other than the knowledge-lean, multi-step problems we discuss here. It is plausible and empirically testable suggestion, however, that the way in which a solution can be recoded may relate to whether or not complex and/or knowledge-rich problems are perceived as "insight problems".

We recognize that our solution recoding hypothesis is preliminary. Nonetheless, we believe that the results of the four experiments reported here undermine the view that a class of insight problems can be distinguished from other types of problems purely on the basis of phenomenology and/or processes that occur during solution discovery. Implications of the solution recoding hypothesis go beyond distinguishing between different definitions of insight. Our view of the processes of solution discovery in insight problem-solving indicates linkages between insight and conventional problem-solving, suggesting that accounting for insight lies within the scope of unitary cognitive architectures such as SOAR (Newell, 1991) and ACT-R (Anderson, 1993). Distinguishing between processes of solution discovery and solution recoding also has implications for neuropsychological studies that associate creative problem-solving with specific cortical regions (e.g., Carlsson, Wendt, & Risberg, 2000). Furthermore, an appropriate focus on solution recoding may help resolve the difficult question of why it appears to be so difficult to transfer or train creative or "insightful" thinking (e.g., Sternberg & Bhana, 1986; Davidson, 1995). Rather than relying on generic instructions to think "outside the box", it may be productive to encourage strategies for recoding, remembering and reusing solutions.

References

Anderson, J. R. (1993). *Rules of the mind*. New York, LEA.

Anderson, J. R. (2000). *Cognitive psychology and its implications* (5th ed.). New York: W. H. Freeman.

Carlsson, I., Wendt, P. E., & Risberg, J. (2000) On the neurobiology of creativity. Differences in frontal activity between high and low creative subjects, *Neuropsychologia*, 3, 873-885.

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

Davidson, J.E. (1995). The suddenness of insight. In R.J. Sternberg & J.E. Davidson (Eds.), *The nature of insight* (Paperback ed., pp. 125-156). Cambridge, MA: MIT Press.

Dominowski, R. L., & Dallob, P. (1996). Insight and problem-solving. In R. J. Sternberg & J. E. Davidson (Eds.), *The nature of insight* (Paperback ed., pp. 33-62). Cambridge, MA: MIT Press.

Gardner, M. (1977). *Mathematical carnival*. New York: Vantage Books, Random House.

Greeno, J. G. (1978). Natures of problem-solving abilities. In W.E. Estes (Ed.), *Handboook of Learning and Cognitive Processes, Volume 5: Human Information Processing (*pp. 239-270). Hillsdale, New Jersey: Lawrence Erlbaum Associates.

Isaak, M. I., & Just, M. A. (1996). Constraints on thinking in insight and invention. In R. J. Sternberg & J. E. Davidson (Eds.), *The nature of insight* (Paperback ed., pp. 281-325). Cambridge, MA: 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 & Cognition*, *25*, 1534-1555.

Knoblich, G., Ohlsson, S., & Raney, G. E. (2001). An eye movement study of insight problem solving. *Memory & Cognition*, *29*, 1000-9.

Lovett, M. C., & Anderson, J. R. (1996). History of success and current context in problem solving. *Cognitive Psychology*, *31*, 168-217.

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 & Cognition*, *27*, 176-201.

Metcalfe, J. (1986). Premonitions of insight predict impending error. *Journal of Experimental Psychology: Learning, Memory & Cognition*, *12*, 623-634.

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

Newell, A. (1991). *Unified theories of cognition*. Cambridge MA, Cambridge University Press.

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

Ormerod, T. C. and Gross, H. (2003). What toy problems can tell us: A comparison of the insight and creative problem-solving skills of designers and financiers. Paper presented at the Sixth International Conference on Naturalistic Decision-Making, Pensacola, Florida, May 2003.

Ormerod, T. C., MacGregor, J. N., & Chronicle, E. P. (2002). Dynamics and constraints in insight problem solving. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, *28*, 791-799.

Schooler, G. W., Ohlsson, S., & Brooks, K. (1993). Thoughts beyond words: When language overshadows insight. *Journal of Experimental Psychology: General*, *122*, 166-183.

Seifert, C. M., Meyer, D. E., Davidson, N., Patalano, A. L., and Yaniv, I. (1996). Demystification of cognitive insight: Opportunistic assimilation and the prepared-mind perspective. In R. J. Sternberg & J. E. Davidson (Eds.), *The nature of insight* (Paperback ed., pp. 65-124). Cambridge, MA: MIT Press.

Simon, H. A. (1986). The information processing explanation of Gestalt phenomena. *Computers in Human Behavior*, *2*, 241-255.

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

Smith, S. M. (1996). Getting into and out of mental ruts: a theory of fixation, incubation and insight. In R. J. Sternberg & J. E. Davidson (Eds.), *The nature of insight* (Paperback ed., pp. 229-251). Cambridge, MA: MIT Press.

Sternberg, R. G., and Bhana, K. (1986) Synthesis of research on the effectiveness of intellectual skills programs: Snake-oil remedies or miracle cures? *Educational Leadership,* 44, 60-67.

VanLehn, K. (1989). Problem solving and cognitive skill acquisition. In M. I. Posner (Ed.), *Foundations of cognitive science*. Cambridge, MA: MIT Press.

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* (Paperback ed., pp. 157-196). Cambridge, MA: 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. (1985). A Gestalt perspective on computer simulations of cognitive processes. *Computers in Human Behavior*, *1*, 19-33.

Author Note

Edward P. Chronicle, Department of Psychology, University of Hawaii at Manoa. Thomas C. Ormerod, Department of Psychology, Lancaster University, UK. James N. MacGregor, School of Public Administration, University of Victoria, British Colombia.

This work was partially supported by a National Science and Engineering Research Council research grant to J. MacGregor, and by a Leverhulme Trust Study Abroad grant to T. Ormerod. We are grateful to Simon Slavin for his help in determining the state spaces of the 6 and 10 coin problems, and to Faye Skelton for testing participants in Experiment 4. We also thank William Batchelder, Günther Knoblich and two anonymous reviewers for their helpful comments on drafts of this paper.

Correspondence concerning this article should be addressed to Tom Ormerod, Department of Psychology, Lancaster University, Lancaster LA1 4YF, UK. Email T.Ormerod@lancaster.ac.uk

Table 1.

Criteria for insight problems, and their match with characteristics of the 6-coin and 10-coin problems

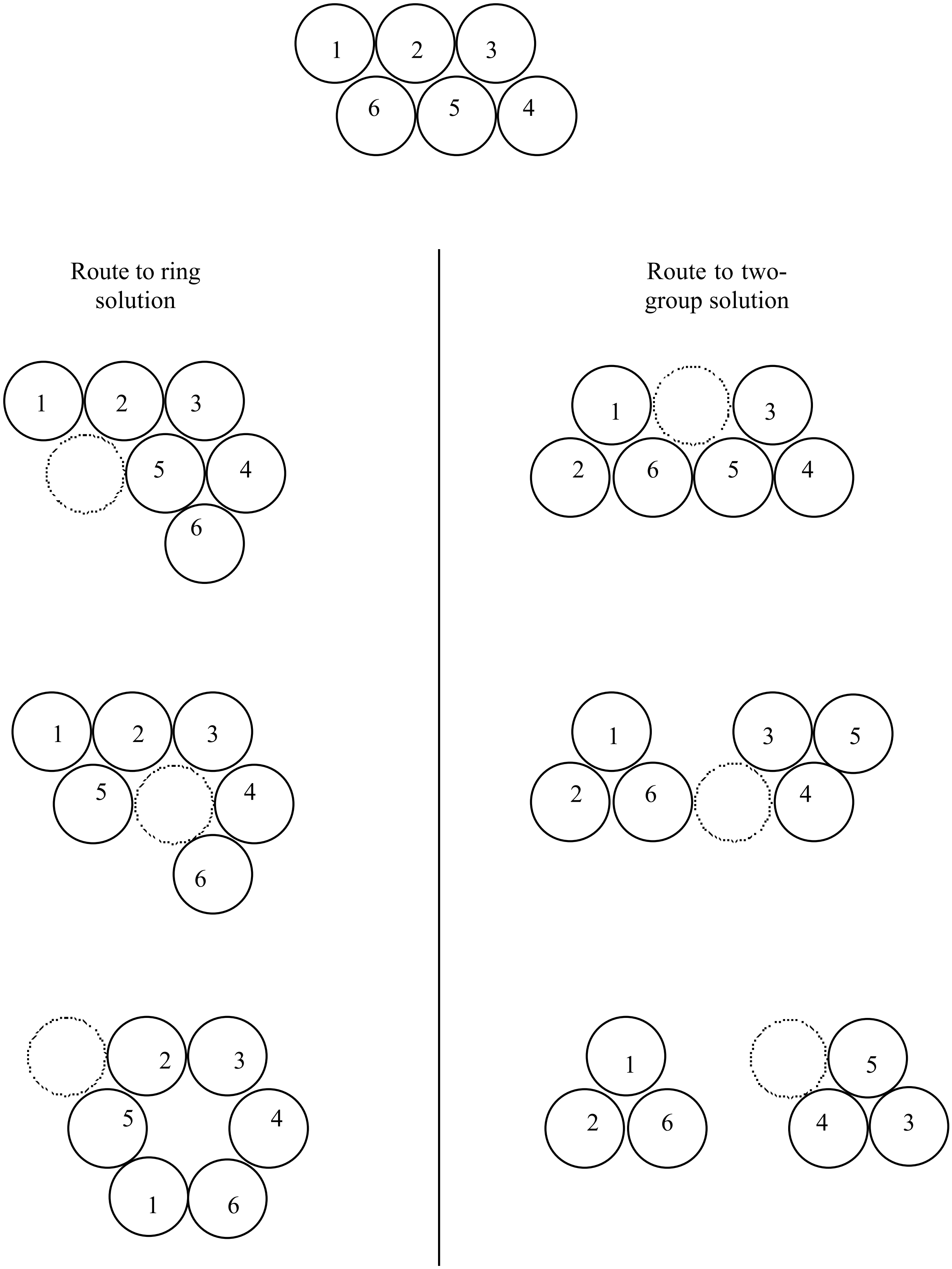
Criterion	Ideal	Six-coin Ring	Six-coin Abstract (2-group solution)	Ten-coin
Step-function	Yes	Yes	No	No
Discontinuity	Yes	No	Yes	Only in minority of cases
Restructuring (potential)	Yes	No	Yes	Yes
Restructuring (actual)	Yes	No	No	Only in minority of cases
Solution recoding	High	Low	High	High

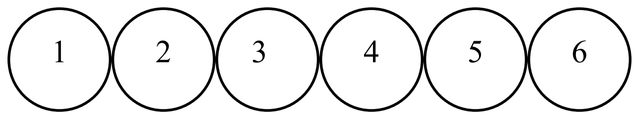
Figure Captions

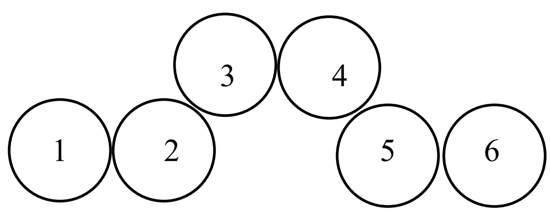
Figure 1. The six-coin problem, with initial state shown top center. In the ring version, shown in the left column, the task is to transform the initial state into the ring goal state (bottom left), moving only three coins. Each move involves sliding a coin (in two dimensions) to a position where it touches exactly two others, without nudging or displacing any other coin. There are only two correct sequences of moves that will reach the ring goal: 6 to 5 & 4, 5 to 1 & 2, 1 to 5 & 6, or the mirror image 3 to 1 & 2, 2 to 5 & 4, 4 to 2 & 3. The problem also has an abstract version, with the same rules and operator, in which the goal is given verbally as “each coin must touch exactly two others”. In the abstract version, an alternative, two-group goal is available. There are 176 paths to correct two-group arrangements. One example is shown in the right column.

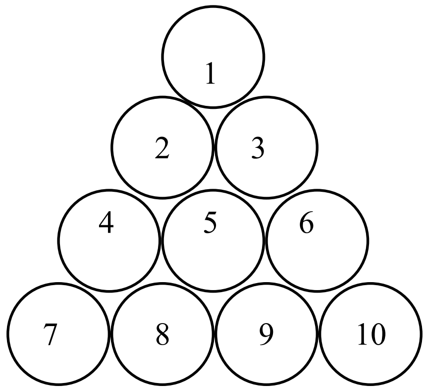
Figure 2. The starting arrays for the Straight-line (upper panel) and Partial-ring (lower panel) conditions of Experiment 2.

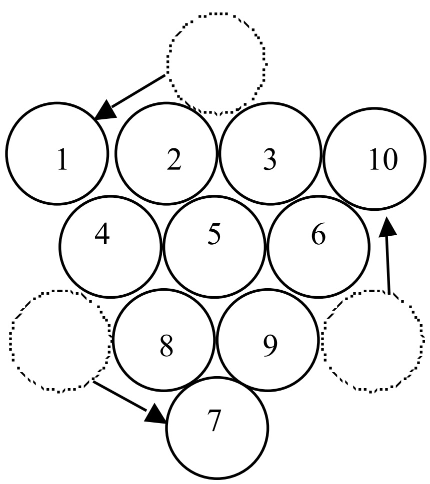
Figure 3. The ten-coin problem (upper panel), together with a “rotational” sequence of three moves that reach the goal state (lower panel).


Route to ring solution
Route to two-group solution


1
2
3
4
5
6


3
4
1
2
5
6


1
2
3
4
5
6
7
8
9
10


1
2
3
10
4
5
6
8
9
7