Authors

Kershaw, Trina C.
Ohlsson, Stellan
Coyne, Colleen

The Fallacy of Single-Source Explanations:
The Multiple Difficulties of the Nine-Dot Problem

Trina C. Kershaw (tkersh1@uic.edu), Stellan Ohlsson (stellan@uic.edu), and Colleen Coyne
Department of Psychology
University of Illinois at Chicago
1007 W. Harrison St., Chicago, IL, 60607

Abstract

Single-source explanations of problem difficulty are common throughout the insight literature (cf. Dominowski & Dallob, 1995; Isaak & Just, 1995; Weisberg & Alba, 1981). However, many insight problems contain multiple difficulties. We propose that the nine-dot problem (Maier, 1930) is difficult because of the interaction between perceptual, knowledge, and process factors. Two experiments show how different types of training can have a statistically significant effect on the overall solution rate, yet produce a small effect size. A third experiment shows how the multiple difficulties of the nine-dot problem can be simultaneously addressed.

Sources of Difficulty in Insight

A main focus of the problem solving literature is determining what features of a problem lead to difficulty in achieving the solution. Insight problems, such as the nine-dot problem (Maier, 1930) and Duncker's (1945) candle problem, are defined by the constraints the problem solver attaches to the problem. As the solver encodes the insight problem, prior knowledge is activated that does not necessarily help the solver (Ohlsson, 1992; Wiley, 1998). This prior knowledge constrains the solver's representation of the problem, therefore leading to an impasse that may not be resolved. If the solver is able to relax these constraints, he or she exits the impasse and solves the problem (Knoblich, Ohlsson, Haider, & Rhenius, 1999).

In studying insight, researchers are often interested in how solvers are able to overcome an impasse. Knoblich et al. (1999) propose that constraint relaxation, and therefore impasse resolution, occurs as a result of spreading activation. However, other researchers believe that a conscious restructuring of the problem is needed to achieve solution. Sometimes researchers give hints to help solvers break out of an impasse, or, alternatively, train the solvers on helpful elements of the problem to prevent impasse (cf. Kershaw & Ohlsson, 2001). Within this general explanatory scheme, the difficulty of each insight problem is explained by identifying the particular difficulty that is operating in each.

The analysis of the difficulty of a problem is often based on the results of providing solvers with hints or training. Or, on the basis of limited evidence, a particular explanation is sometimes passed down from author to author. A striking feature of the majority of explanations offered for the difficulty of insight problems is that they attribute the difficulty to a single source.

The difficulty of the six matches problem (use 6 matches to form four equilateral triangles with each side being the length of a match) has often been explained by solvers' belief that the solution is two-dimensional (Scheerer, 1963; Weisberg & Alba, 1981). More recently, Ormerod, MacGregor, and Chronicle (2002) attributed this same failure to the eight-coin problem. Failure to solve in both Maier's (1931) two-string problem and Duncker's (1945) candle problem has been attributed to functional fixedness (Dominowski & Dallob, 1995). Kaplan and Simon (1990) allocate the difficulty of the Mutilated Checkerboard problem to the incompleteness of the initial problem representation. Isaak and Just (1995) list 20 insight problems and give only a singular explanation of the difficulty for each. The nine-dot problem (Maier, 1930) provides a multi-decade example of single-source explanation.

Single-Source Difficulties in the Nine-Dot Problem

Maier (1930) is the earliest source for the nine-dot problem (connect a 3 x 3 square of dots by using only four straight lines without lifting your pen from the page, see Figure 1) and gives the most popular reason for the problem's difficulty: People do not want to draw lines outside of the box which is set up by the dots (cf Chronicle, Ormerod, & MacGregor, 2001; Scheerer, 1963). Maier's claim has been tested by a series of experiments that proposed other, single-source explanations for the nine-dot problem's difficulty.

Burnham and Davis (1969) believed that the problem was difficult because people were unable to understand its abstract definition. Weisberg and Alba (1981) contended that the difficulty of the problem was due to lack of applicable prior experience. Lung and

Dominowski (1985) found that solvers were unwilling to begin and end lines where there was not a dot. More recently, MacGregor, Ormerod, and Chronicle (2001) proposed that level of mental lookahead (the number of lines that can be held in memory at any one time) predicts performance on the nine-dot problem. Chronicle, Ormerod, and MacGregor (2001) studied the influence of perceptual manipulations on performance.
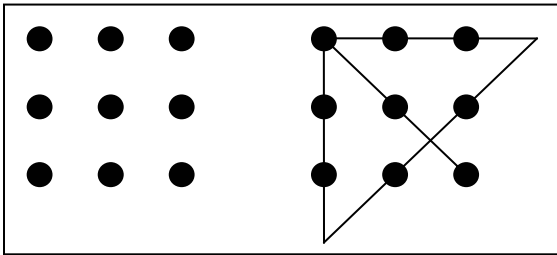

Figure 1: The Nine-Dot Problem and its Solution

In our own work, low solution rates have been attributed to the non-dot turns hypothesis (Kershaw & Ohlsson, 2001). Simply, people do not want to turn on a non-dot point, i.e., a space between two dots. Training participants to make non-dot turns led to increased performance relative to control (19/30 vs. 11/30) on a variant of the nine-dot problem. However, we have had mixed results applying the non-dot turns training to the nine-dot problem itself.

Problems with Single-Source Explanations

Relying upon a singular explanation of difficulty for an insight problem is not useful when the explanation does not fully account for the experience of solvers who attempt the problem. None of the explanations of the nine-dot problem's difficulty have fully accounted for the poor performance of subjects on the nine-dot problem (typical unaided solution rate = 0%, MacGregor et al., 2001). After reviewing the available research and watching people struggle while solving the nine-dot problem, we decided to work from the idea that the problem is composed of multiple difficulties. We propose that the multiple difficulties of the nine-dot problem contain perceptual, knowledge, and process factors. Perceptual factors include figural integrity and figure/ground relationships, knowledge factors include past experience with other dot problems, and process factors include the size of the problem space, the ill-defined goal, and mental lookahead.

If there are multiple sources of difficulty in the nine-dot problem, we should expect the following pattern: Each type of training or hint should (a) have a statistically significant effect on the solution rate, but (b) the effect size should be small, because taking care

of just one source of difficulty leaves the other sources in place. Consequently, it should be possible to get significant but small effects with a wide variety of training and hint types that may or may not have any conceptual relations to each other, depending on the configuration of difficulties that the solver faces. In two studies, we demonstrate precisely this pattern in the case of the nine-dot problem. In a third study, we show how applying training that accounts for multiple difficulties leads to an increased solution rate and an increased effect size.

Experiment I

In addition to non-dot turns, the nine-dot problem is difficult because of perceptual factors. The square set up by the dots creates a gestalt that solvers are unwilling to breach. Chronicle et al. (2001) attempted to increase solution through perceptual manipulations such as adding two additional unfilled circles next to the upper right and lower left filled dots. In this experiment, we attempted to increase non-dot turns through the use of training problems made of black dots presented on a grid of unfilled circles (see Figure 2).
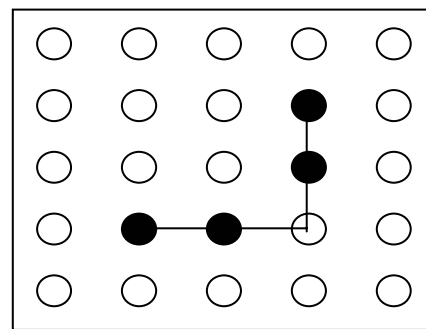

Figure 2: Facilitating Training Exercise with Solution
NB: Grid covers entire page in real training

In a 2 x 2 design, 160 subjects were assigned to one of four groups: facilitating grid, hindering grid, facilitating no grid, and hindering no grid. Participants in the facilitating groups received training problems that required non-dot turns, while participants in the hindering groups received training problems in which they always turned on a dot (for more examples of training problems, see Kersahw & Ohlsson, 2001). The training problems did not look like the nine-dot problem and required one, two, or three turns (see Figure 2). All participants completed 12 training problems, and had the nine-dot problem as their target problem.

There was a main effect of training, in that subjects in the facilitating training groups had a higher solution rate (12/80) than subjects in the hindering training groups (1/80), $\chi^2 (1, N=160) = 10.13$, $p < .05$, lambda = .12. However, only 13 subjects out of 160 solved, or

8%, and the effect size was small . There was no effect for the grid, the perceptual factor.

The significant effect of the non-dot turns training provides support for the knowledge factor. The increased performance by the facilitating groups on the nine-dot problem mirrors Kershaw and Ohlsson's (2001) finding of increased performance on variants of the nine-dot problem following training on non-dot turns. However, perceptual factors did not significantly impact subjects' performance.

Experiment II

Burnham and Davis (1969) believed that a consequence of the abstract definition of the nine-dot problem was that the sequence of lines needed to solve it was non-obvious. The nine-dot problem is ill-structured because solvers do not know what the shape of the solution will be. We devised two types of training that would familiarize participants with the solution shape of the nine-dot problem and explore the process factors of the nine-dot problem. One type of training used a motor manipulation to teach participants the order to connect the dots in, and the other type of training sought to teach participants how to perceptually distinguish the shape of the nine-dot solution from other shapes.

In a 2 x 2 design, 120 participants were assigned to one of four groups: motor relevant, motor irrelevant, perceptual relevant, and perceptual irrelevant. Participants in the relevant shape conditions learned the shape of the nine-dot solution, while participants in the irrelevant shape conditions learned a shape that did not correspond to the nine-dot solution. The motor training consisted of figures composed of numbered dots. Participants were instructed to connect the dots in number order, much like a child's connect-the-dots puzzle (see Figure 3).

In the perceptual training conditions, participants were shown a shape, and then rated a sequence of target and distractor shapes as being either the same as or different than the shape they were shown (see Figure 4). Participants in the perceptual training groups also rated how confident they were of their judgments on a scale of 1-5, 1 = not at all confident and 5 = completely confident. All participants attempted the nine-dot problem as their target problem.

There was a main effect of training shape, in that subjects in the shape-relevant groups had a higher solution rate (7/60) on the nine-dot problem than subjects in the shape-irrelevant groups (1/60), χ^2 (1, 120) = 4.72, $p < .05$, lambda = .03. However, only 8 out of 120 participants solved, or 7%, and the effect size was small. There was no effect for training type.

We infer that knowing the shape of the nine-dot solution is an important process factor, otherwise there would have been no increase in the solution rate above

the expected 0%. We also infer that there are other sources of difficulty, otherwise the increase would have been larger.
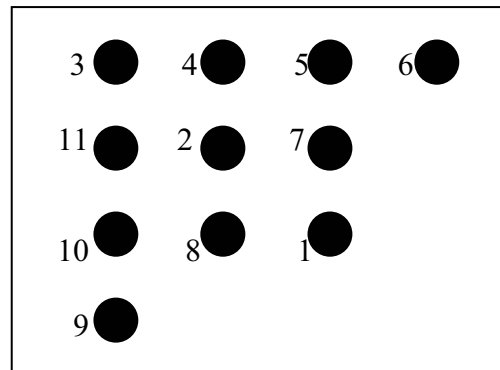

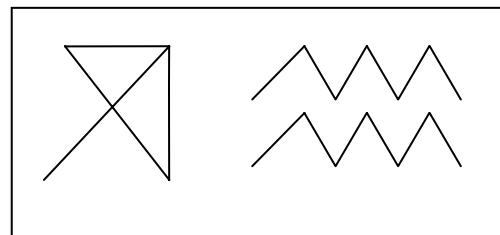Figure 3: Motor Relevant Training Exercise

Figure 4: Target and Distractor Shapes for the Perceptual Relevant Condition

Experiment III

In both Experiments I and II, we attempted to address one of the multiple difficulties of the nine-dot problem. In Experiment I, we found evidence for knowledge factors via the non-dot turns (facilitating) training. In Experiment II, we explored process factors, via learning how to distinguish the correct shape in the perceptual condition, and how to draw the correct shape in the motor condition. However, applying each of the multiple difficulties of a problem *individually* does not serve to increase the solution rate for the problem by a large amount. Training on any one of the individual difficulties can lead to a partial insight (Ohlsson, 1992), but will not be enough to allow an individual to solve the whole problem. For example, if a participant learns to make non-dot turns, he or she still does not know which dot to begin at, what order to draw the lines in, or when and where to draw lines that extend outside of the dots. If a problem has multiple difficulties, then these difficulties need to be attacked in a combinatorial fashion.

The obvious remedy here is to look for multiple sources of difficulty, and to test such hypotheses with experiments that aim to alleviate all of them. Experiments I and II identified several difficulties of the nine-dot problem. Experiment III combined all of

the elements that have been shown to increase the solution rate for the nine-dot problem, both from Experiments I and II, and from the nine-dot literature. The training used in Experiment III featured problems made of black, filled dots presented on a grid of other unfilled dots as well as problems made of black dots that were alone on the page (see Experiment I, and Kershaw & Ohlsson, 2001). The training also had a perceptual component in which participants learned to distinguish the shape of the nine-dot problem solution from other shapes (see Experiment II: perceptual relevant condition). In addition, the training contained a dialogue component in that participants were informed of the purposes of each training task. Specifically, participants were told that the shape they learned during the perceptual training was the shape that would be required to solve the target problem, and that it was necessary to draw lines outside the dots and turn in the empty space between dots. Finally, Experiment III contained a feedback component, in that participants were shown the correct answer for judging a shape or connecting dots for each judgment or problem that was completed.

In addition to combining elements from the previous experiments and the literature on the nine-dot problem, Experiment III compared several new variants of the nine-dot problem to the original as a further test of the non-dot turns hypothesis. Much of the nine-dot literature has shown that variants of the nine-dot problem (such as problems with additional dots) are easier to solve than the traditional problem (cf. Burnham & Davis, 1969; MacGregor et al., 2001); Weisberg & Alba, 1981), but none of this research has specified a testable theory as to why these variants are easier than the nine-dot problem. The non-dot turns hypothesis (Kershaw & Ohlsson, 2001) predicts that the more non-dot turns a problem requires, the more difficult the problem is. By counting the number of non-dot turns required, the relative difficulty of the variants of the nine-dot problem can be predicted.

The new variants of the nine-dot problem have the same solution shape, but require differing numbers of non-dot turns. The variants either contain more dots or have a dot moved to a place outside of the traditional nine-dot square to create a new figure. One variant is the 10-dot problem (see Figure 5), which has an additional dot at the lower left-hand corner, thereby requiring one turn on a non-dot point. Another variant is the displaced nine-dot problem (see Figure 5), in which the top line of dots in the nine-dot problem is shifted to the right by one dot. This version requires two non-dot turns. The third variant moves the dot in the upper left hand corner to the bottom right hand corner, as an extension to the right of the figure (see Figure 5). This variant requires three turns on a non-

dot point, and is therefore called the three-turn problem.

Different solution rates are expected for the different problem versions, based on the number of non-dot turns required. In general, the more non-dot turns required, the greater the difficulty, or the lower the solution rate, of the problem. However, the nine-dot problem should have the lowest solution rate because in addition to the difficulty of turning on a non-dot point, the traditional nine-dot problem forms a figure that discourages participants from extending lines beyond the boundaries of the figure.
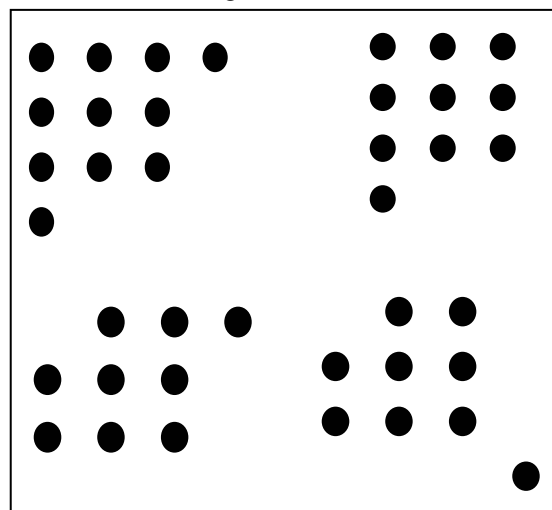

Figure 5: 11-Dot, 10-Dot, Displaced Nine-Dot, and Three-Turn Problems

The variant requiring the three non-dot turns should have a solution rate equal to or possibly lower than the nine-dot problem. However, due to the very low solution rate of the nine-dot problem, it is not expected that the variant requiring three non-dot turns will have a significantly lower solution rate than the nine-dot problem.

Based on studies that include an 11-dot variant (cf. Burnham & Davis, 1969; MacGregor et al., 2001; see Figure 5), the 10-dot problem is predicted to have the best solution rate of the three new problems, but the 11-dot problem should have the best solution rate overall because it requires no non-dot turns. The solution rate for the displaced nine-dot should be somewhere between the solution rates for the 10-dot and the traditional nine-dot. The reason the displaced nine-dot should have a higher solution rate than the traditional nine-dot is that the displaced nine-dot breaks up the gestalt of the traditional nine-dot. We believe that participants will be more likely to extend lines beyond the figure in the displaced nine-dot version.

There was a main effect of training, in that participants who received the training were more likely

to solve their target problem (89/150), than participants in the control group (24/150), χ^2 (1, N=300) = 59.98, p < .05, lambda = .25. In the control group, there was a significant difference between solution rates for the problem types, χ^2 (4, N=150) = 39.39, p < .05, lambda = .08. About half of the participants in the 11-dot group solved their target problem, compared to the low solution rates for each of the other problem types in the control group (see Figure 6). In the trained group, there was a significant difference between solution rates for the problem types, χ^2 (4, N=150) = 39.07, p < .05, lambda = .30.

The combination of multiple difficulties into one set of training led to the highest solution rate for the target problem, 59% (solution rate for nine-dot problem in experimental condition = 40%). There was a moderate effect size for both the effect of the training and the effect of problem type in the trained group. Additionally, the pattern of results in the trained group supports the non-dot turns hypothesis. As the number of non-dot turns increased, the solution rate decreased (see Figure 6).
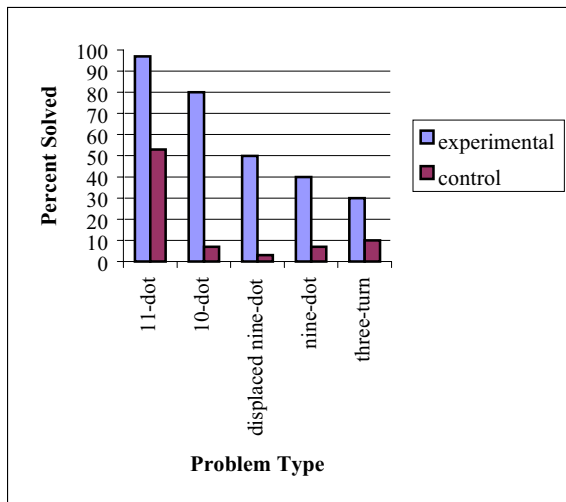

Figure 6: Experiment III: Solution Rate for Each Target Problem

The results also showed a distinction between Lung and Dominowski's (1985) conception of the main difficulty of the nine-dot problem as drawing lines that begin or end on a non-dot point, and Kershaw and Ohlsson's (2001) conception of non-dot turns as the key difficulty. We counted the number of participants, for each problem type that required non-dot turns and within each condition, who drew lines outside the box set up by the dots (see Table 1). The differences between the problem types within the control group were not significant, χ^2 (3, N = 120) = 5.35, p > .05, lambda = .05. In the experimental group, the majority

of participants drew lines outside of the dots even when they did not solve the problem. But when participants received the nine-dot problem, they were less likely to draw lines outside the dots, nine-dot problem vs. displaced nine-dot, χ^2 (1, N = 60) = 9.31, p < .05, lambda = .23. The problems have the same number of non-dot turns and require the same solution, but the displaced nine-dot breaks up the Gestalt formed by the dots. Although the solution rate for these two problems was not significantly different, the displaced nine-dot problem led to a greater incidence of drawing lines outside the dots than the traditional nine-dot problem.

Table 1: Experiment III: Number of Lines Drawn Outside for Each Group and Problem Type

Problem Type	Control		Experimental	
	Outside	Inside	Outside	Inside
10-dot	5	25	26	4
dis. nine-dot	10	20	28	2
nine-dot	5	25	18	12
three-turn	11	19	27	3

Note. The number of lines drawn outside was not counted for the 11-dot problem because the problem does not require any non-dot turns.

General Discussion

The nine-dot problem, like many insight problems, presents the problem solver with a small and artificially restricted problem solving environment, and the solution is short and undemanding. It is therefore tempting to think of the solution as encapsulated in a single, unitary idea or thought. This way of thinking naturally leads one to think of the difficulty of the problem in terms of a single, unitary blockage. The focus of research then becomes to identify that blockage, and the history of research on the topic becomes a history of one hypothesized source of difficulty being replaced by another when it is shown that removal of the first one does not boost the solution rate very much.

However, closer inspection and reflection reveals that even simple problems with short solutions can contain multiple sources of difficulty. The nine-dot problem provides a good example. The Gestalt psychologists focussed on the perceptual factors operating in this problem. The square shape of the nine dots forms a natural perceptual configuration that interferes with the invention of the solution, which breaks with that shape in several ways. Other researchers (cf. Chronicle et al., 2001) have confirmed that such perceptual factors are indeed operating in the

nine-dot problem, but their removal only increases solution rates with a minor amount.

Other factors are operating as well. The shape formed by the solution is not in any way familiar or natural or previously known. Thus, familiarizing the solvers with that shape also has an effect, again small in magnitude. Likewise, prior experience with dot problems does not apply to the nine-dot problem (Weisberg & Alba, 1981).

It is the combination of these perceptual, knowledge, and process factors that can lead to success on the nine-dot problem and its variants. However, the difficulty of turning on a non-dot point is still present, as seen in the results of Experiment III.

The implicit assumption of single-source difficulty that has characterized much research on insight problems leads to a non-cumulative research process. Each hypothesized difficulty is 'confirmed' in the eyes of its advocates, because it generates a statistically significant rise in the solution rate. However, each is also 'falsified' in the eyes of its detractors, because its removal does not lead to anywhere near 100% solution rate. So the history of the topic takes the form of a sequence of single-source hypotheses, none of which explain more than a small proportion of the difficulty. The nine-dot problem provides an example of a chain of non-building research. Instead of utilizing and extending previous explanations, researchers dismissed earlier claims in favor of establishing new explanations of the difficulty of the nine-dot problem. We hope to change the direction of research on the nine-dot problem, and on insight problem solving in general, by moving towards difficulty as a synthesis of factors, rather than a single element. In addition, it would be interesting to evaluate if multiple difficulties are operating in other creative tasks.

References

Burnham, C.A., & Davis, K.G. (1969). The nine-dot problem: Beyond perceptual organization. *Psychonomic Science, 17*(6), 321-323.

Chronicle, E.P., Ormerod, T.C., & MacGregor, J.N. (2001). When insight just won't come: The failure of visual cues in the nine-dot problem. *Quarterly Journal of Experimental Psychology, 54A*(3), 903-919.

Dominowski, R.L, & Dallob, P. (1995). Insight and Problem solving. In R.J. Sternberg and J.E. Davidson (Eds.), *The nature of insight* (pp. 33-62), Cambridge, MA: The MIT Press.

Duncker, K. (1945). On problem solving. *Psychological Monographs, 58*, whole no. 270.

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

Kaplan, C.A., & Simon, H.A. (1990). In search of insight. *Cognitive Psychology, 22*(3), 374-419.

Kershaw, T.C., & Ohlsson, S. (2001). Training for insight: The case of the nine-dot problem. In J.D. Moore & K. Stenning (Eds.), *Proceedings of the Twenty-third Annual Conference of the Cognitive Science Society* (pp. 489-493). Mahwah, NJ: Lance Erlbaum Associates.

Knoblich, G., Ohlsson, S., Haider, H., & Rhenius, D. (1999).Constraint relaxation and chunk decomposition in insight problem solving. *JEP: LMC, 25*, 1534-1555.

Lung, C.T., & Dominowski, R.L. (1985). Effects of strategy instructions and practice on nine-dot problem solving. *JEP:LMC, 11*(4), 804-811.

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. *JEP:LMC, 27*(1), 176-201.

Maier, N.R.F. (1930). Reasoning in humans: I. On direction. *Journal of Comparative Psychology, 10*, 115-143.

Ohlsson, S.(1992). Information processing explanations of insight and related phenomena. In M. Keane and K. Gilhooly (Eds.), *Advances in the Psychology of Thinking* (Vol.1, pp. 1-44). London: Harvester-Wheatsheaf.

Ormerod, T.C, MacGregor, J.N., & Chronicle, E.P. (2002). Dynamics and constraints in insight problem solving. *JEP:LMC, 28*(4), 791-799.

Scheerer, M. (1963) Problem solving. *Scientific American, 208*(4), 118-128.

Weisberg, R.W., & Alba, J.W. (1981). An examination of the alleged role of "fixation" in the solution of several "insight" problems. *Journal of Experimental Psychology: General, 110*(2), 169-192.

Wiley, J. (1998). Expertise as mental set: The effects of domain knowledge on creative problem solving. *Memory & Cognition, 26*(4), 716-730.