Distributed Cognition and Insight Problem Solving

Frédéric Vallée-Tourangeau, Anna Weller, and Gaëlle Villejoubert
Department of Psychology, Kingston University
Kingston-upon-Thames UNITED KINGDOM   KT1 2EE
f.vallee-tourangeau/a.weller/g.villejoubert@kingston.ac.uk

Abstract

Problem solving from a distributed cognitive system perspective is an emergent product of the strategic and opportunistic manipulation of artefacts populating a physical space. In the present study, insight problem solving was investigated with matchstick algebra problems. These problems are false equations expressed with Roman numerals transformed into true equations by moving one matchstick. Participants were split in two groups. In the first, the paper group, they examined a static two-dimensional representation of the false algebraic expression and told the experimenter which matchstick should be moved. The non-interactive procedure was similar to the one employed in Knoblich, Ohlsson, Haider, and Rhenius (1999). In the second group, the interactive group, participants manipulated a concrete three-dimensional representation of the false equation. Success rates in the paper group for different problem types closely replicated the pattern of data reported in Knoblich et al. In turn, participants in the interactive group were significantly more likely to achieve insight. Problem solving success in the paper group was best predicted by performance on a numeracy test, whereas in the interactive group, it was best predicted by performance on a visuo-spatial reasoning test. Different types of resources and skills were involved in the different versions of the task. Implications for process models of problem solving are discussed.

Keywords: Problem solving, interactivity, individual differences, distributed cognition, education.

Introduction

Transformation problems such as the Tower of Hanoi or river crossing problems are structured in terms of a well defined space of intermediate states linked by simple discrete moves, with the goal state clearly visible or imaginable. Their solution rarely involves 'aha' moments. Insight problems on the other hand are different in that the goal state, or resolution, is initially not visible or imaginable. With insight problems most participants initially experience an impasse from which they may or may not emerge. The impasse is experienced as a result of a problem representation that is driven by 'organizing assumptions' (Segal, 2004, p. 142) that mislead the reasoner and prevent him or her from anticipating the solution. Overcoming an impasse is understood to be driven by a representational change that re-cast the relationship among the elements of the representation or that redefine the role of these elements. This representational perspective on insight has roots in Gestalt psychology (e.g., Wertheimer, 1959) and has been formulated in information processing terms by Ohlsson (1984, 1992).

The initial representation of the problem is based on the manner with which the reasoner configures perceptual elements that compose the problem (how these elements are 'chunked') and reflects the reasoner's comprehension based on his or her knowledge and expertise. Thus this initial representation, structured by perceptual chunks and conceptual assumptions, guides how the reasoner will attempt to solve the problem. However that guidance may also constrain and impede successful problem resolution. Certain assumptions of the problem representation need to be relaxed in order for the reasoner to solve the problem. A classic example of the importance of constraint relaxation in problem solving is offered by Maier's (1930) 9-dot problem. The task is to link all 9 dots with four continuous lines without lifting the pen from the paper. The perceptual configuration of the dots imposes an implicit constraint that the lines can only be drawn within the projected perimeter delineated by the dots. Insight for this problem involves relaxing that constraint.

In turn, a well-known Max Wertheimer problem illustrates how the segmentation of visual information into chunks is an important determinant of the ensuing problem representation and the ease with which a reasoner can solve the problem (see Ohlsson, 1984; Segal, 2004; Fig. 1). In this problem, the reasoner must calculate the area of the composite figure involving a square and a parallelogram. This initial problem representation specifies certain operators that must be retrieved from long term memory (such as the formula to calculate the area for parallelograms). It may be that given this initial representation the reasoner is unable to retrieve the appropriate operators and hence may experience an impasse. The reasoner may seek to restructure the problem representation by decomposing the perceptual chunks at the heart of it. Some people may realize that the square -parallelogram configuration can be decomposed in terms of two overlapping triangles. This new chunking arrangement may encourage a more fruitful representation in terms of a rectangle (once the triangles no longer overlap) that would cue much simpler operators to solve the problem.
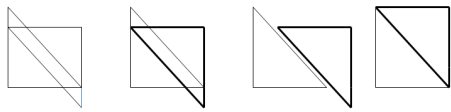
Figure 1: Max Wertheimer's square and parallelogram problem. The goal is to find the surface area of the composite figure. A decomposition of the perceptual chunks that configured the initial problem representation (involving a parallelogram on top of a square) leads to a representation involving overlapping triangles which in turn leads to a representation of a much simple perceptual chunk: a rectangle.

Matchstick Algebra

Constraint relaxation and chunk decomposition are important drivers of representational restructuring. The reconstruction of a problem representation is deemed necessary to overcome an impasse and achieve insight. Constraint relaxation and chunk decomposition were explored in a series of elegant experiments with matchstick algebra problems developed by Knoblich, Ohlsson, Haider, and Rhenius (1999). A matchstick algebra problem is a false statement expressed with Roman numerals. Participants are required to move, but not remove, one stick to make the equation true, with the 'V' and 'X' numerals each consisting of 2 slanted sticks. For example, 'VI = VII + I' is a false statement that can be transformed into a true one by moving a single stick from the '7' on the right of the equal sign to the '6' on the left of it such as to yield 'VII = VI + I'. To achieve insight, participants must relax constraints that reflect knowledge and assumptions concerning algebraic transformations, and decompose familiar perceptual chunks in the form of numerals and symbols (operators).

Table 1: The four matchstick algebra problem types developed by Knoblich, Ohlsson, Haider, and Rhenius (1999). Solutions for problems for Type A through C require relaxing constraints of increasing scope, while solving problems of Type D requires decomposing a tight perceptual chunk.

Type	Equation	Solution
A	VI = VII + I	VII = VI + I
B	I = II + II	I = III - II
C	III = III + III	III = III = III
D	XI = III + III	VI = III + III

Using matchstick algebra Knoblich et al. explored the importance of constraint relaxation and chunk decomposition in achieving insight. To test the importance of constraint relaxation, they developed three types of false statements the solution for which required relaxing constraints of different scopes (see Table 1). Solving Type A problems involved relaxing a relatively narrow constraint that numerals cannot be decomposed (value constraint). Relaxing that constraint enables participants to transform a numeral to make the statement true. Solving Type B problems involved relaxing a constraint with a broader scope, that is one including the constraint on manipulating operators (operator constraint). Solving Type C problems involved relaxing a constraint with an even broader scope, namely the constraint that people rarely communicate in tautological terms (tautology constraint). Hence to solve these problems, participants must realize that tautologies are acceptable. Knoblich et al. predicted that the solution rate for these three types of matchstick algebra problems would be a function of the scope of the constraint to be relaxed, with the narrow constraint of Type A problems the easiest to relax and hence to solve, and the broad constraint of Type C problems the hardest to relax and solve. Knoblich et al. observed the highest rates of problem solving success for Type A problems, followed by Type B problems, and the hardest problems were Type C. Problems of Type D involved relaxing the value constraint (like problems of Type A) but the solution necessitated decomposing a numeral that formed a much tighter perceptual chunk. Knoblich et al. predicted that Type D problems would be much harder to solve than problems of Type A, which is what they observed.

Interactive Problem Solving

We pause here to note with interest a key feature of the Knoblich et al. experimental procedure: Participants were never invited to manipulate matchsticks as such in solving these algebra problems. The so-called matchstick algebra problems did not involve actual matchsticks. Rather the false arithmetic statements were presented on a computer screen and participants announced their proposed solution, which was then noted by the experimenter. Yet, from a distributed cognitive system perspective (Giere, 2006; Cowley & Vallée-Tourangeau, 2010), thinking is the product of an interactive assemblage of resources internal and external to the agent. The environment and its content can be exploited to facilitate reasoning and problem solving in a variety of ways. To this end, a diverse range of actions are performed, including reorganisation of the environment, muttering to oneself, pointing and making notes (Kirsh, 2009).

Kirsh suggests that it is through these actions and interactions that thought is externalised, with external artefacts and representations employed as vehicles for ideas and hypotheses, lightening cognitive load. But these externalisations do not merely function as a means to offload memories and reduce cognitive demands. Rather the generation and, importantly, manipulation of these representations facilitate understanding by reordering the original representation into one that may be more cognitively congenial (Kirsh, 1996). The transformed representation may potentially reveal affordances and new

opportunities to guide behaviour. Spatial rearrangement may modify the problem so that it becomes more visually compelling, allowing the perception of task elements that were hitherto invisible to the reasoner. Spatial rearrangement may also conserve internal computing resources, as executing tasks externally (such as an object rotation) may be quicker and require less effort than if performed mentally, thereby increasing task efficiency (Kirsh, 1995b).

The Present Study

We sought to investigate problem solving in a context where matchstick algebra problems were expressed in a physical representation that could be manipulated by participants. We sought to determine the degree to which constraints of different scopes and the tightness of perceptual chunks remained important obstacles to insight in an interactive version of this problem solving task. Interactivity inevitably engages a broader range of cognitive, perceptual and motor processes and hence problem solving success may well implicate different skills in interactive and non-interactive contexts. In an attempt to gauge the importance of different cognitive skills in these two versions of the task, we profiled participants' numeracy, knowledge of Roman numerals, traditional verbal intelligence (as measured with the National Adult Reading Test; Nelson, 1991 which correlates positively with the Wechsler Adult Intelligence Scale full scale IQ) along with the Beta III test (Kellog & Morton, 1999). Separate elements of the test assess aspects of non-verbal intelligence including spatial reasoning, visual information processing and the speed and accuracy of processing. We were then in a position to identify the better predictor(s) of performance in interactive and non-interactive versions of insight problem solving using matchstick algebra.

Method

Participants

Fifty participants were recruited among students and administrative staff on the campus of Kingston University. Mean age was 27.84 (SD= 12.11) and the majority of participants were female (N= 30).

Procedure

Participants were allocated on a random basis to one of two experimental groups, the paper group or the interactive group. In the paper group participants were presented matchstick algebra problems on a sheet of paper and informed the researcher which 'matchstick' could be moved to transform the expression into a true equation. Participants in the interactive group manipulated artefacts to create and modify the false expressions into true ones. All participants were presented with the four types of problems (A though D) and hence the experimental design was a 2 (group) by 4 (problem type) mixed design. The dependent measure was the percentage of problems of different types solved by the participants.

Participants were tested individually in a quiet room. Participants first completed a numeracy test during a one minute period. This test consisted of simple arithmetic questions. They then completed the NART which involved reading aloud a series of 50 words, the pronunciation for each categorised as correct or incorrect by the experimenter. Participants were then asked to complete the Roman numerals test, in which they were required to translate a series of simple Arabic numbers into their Roman numeral equivalent within a one minute period. No feedback on performance was given on any of these tests.

Participants from both groups were shown 12 incorrect matchstick algebra equations; these equations were the same as those developed by Knoblich et al. (1999; Experiment 1).These 12 problems were composed of four of each of Types A and B, and two of each of Types C and D. The order of presentation was randomised for each participant. Each equation was printed in the centre of a sheet of white A4 paper in large, bold, black font held in a ring binder with the following instruction at the head of each page, "Move **ONE** stick to make the equation **TRUE**". Participants in the paper group were asked to solve the equations using these sheets of paper only. For the interactive group, we designed a magnetic board (27cm x 21cm) on which participants created and modified Roman numerals and algebraic statements using magnetized matchsticks (.5cm x 4.5cm). Participants in the interactive group were first asked to recreate the incorrect form of the equation as presented to them on paper and then to solve the equation by moving one stick to make the equation read true. They were encouraged to touch and manipulate the matchsticks in reasoning about the problems. Participants in both groups were given a maximum of 3 minutes to solve each equation, after which they were presented the next problem.

The experimental session concluded with the five components of the Beta III test: (i) The Coding test required participants to match a series of symbols to numerals (test duration: 120 seconds); (ii) the Picture Completion section consisted of a series of pictures with aspects/items missing that participants must complete (180s); (iii) the Clerical Checking test displayed pairs of symbols or numbers and participants were required to judge whether the pairs were identical or not (120s); (iv) the Picture Absurdities test consisted of a series of panelled pictures and required participants to identify which of a set of pictures showe something absurd or illogical (180s); (v) finally, the Matrix Reasoning test asked participants to choose a picture from a selection of five pictures to fill in a gap in a sequence (300s).

Measures. Both the Maths and Roman numerals tests were expressed in terms of percent correct answers. The

NART score was reported as the number of correctly pronounced words. Matchstick algebra performance for each participant was scored in terms of the percentage of equations correctly solved for each of the four types of problems. Each element of the Beta III test was first scored individually, by summing the correct answers. These scores were then converted to age corrected scaled scores (ACSS; Kellog & Morton, 1999).

Results

Cognitive Profiles

Numeracy skills did not differ significantly between the paper group ($M = 49.5$, $SD = 25.9$) and the interactive group ($M = 51.9$, $SD = 22.0$), $t(48) = 0.35$, $p = .73$. Knowledge of Roman numerals was equivalent in both the paper group ($M = 48.7$, $SD = 23.2$) and the interactive group ($M = 43.6$, $SD = 20.8$), $t(48) = 0.83$, $p = .41$. Performance on the NART did not differ between participants in the paper group ($M = 26.40$, $SD = 6.11$) and those in the interactive group ($M = 24.32$, $SD = 6.59$), $t(48) = 1.16$, $p = .25$. Finally, participants did not differ significantly on any of the Beta III component tests; largest non-significant $t(48) = 1.19$, $p = .24$ for clerical checking.

Matchstick Algebra Performance

The percentage of correct solutions for each problem type for each participant was calculated. The percent correct solution averages in both groups are displayed in Figure 2. Solution rates appeared marginally greater in the paper group compared to the interactive group for Type A problems, but the interactive participants solved more of types B, C and D problems than their paper counterparts. A 4 (problem type: A, B, C, D) by 2 (group: paper, interactive) mixed analysis of variance (ANOVA) revealed a significant main effect of problem type, $F(3, 144) = 24.6$, $p < .001$, a significant main effect of group, $F(1, 48) = 5.06$, $p = .029$, and a significant interaction between problem type and group on problem solving performance, $F(3, 144) = 5.03$, $p = .002$.

Separate ANOVAs were conducted for the paper and interactive groups. In the paper group, the problem type main effect was significant, $F(3, 72) = 29.2$, $p < .001$. Post hoc tests using the Bonferroni correction revealed that the solution rates for Type A problems were higher than for Types B ($p < .001$), C ($p < .001$), and D ($p < .001$), while the solution rates for Type B problems were greater than for Type C problems ($p = .002$). In the interactive group, the problem type main effect was also significant, $F(3, 72) = 5.39$, $p = .002$. Bonferroni corrected post hoc tests revealed that the only significant differences in the solution rates were observed between Types A and C ($p = .02$), and Types B and C ($p = .03$): Thus, in the interactive group, the solution rates for Type A, B, and D did not differ statistically
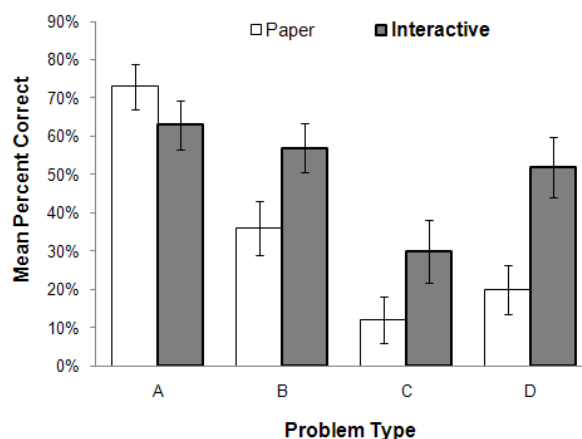

Figure 2: Average solution rates for Types A, B, C and D in the paper group (open bars) and the interactive group (dark bars). Error bars are standard errors of the mean.

Predictors of Performance

Performance on some of the cognitive abilities tests was correlated with overall performance on the matchstick algebra task. For the paper group participants' numeracy was most strongly correlated with matchstick algebra performance, $r(23) = .51$, $p = .009$, and performance on the NART, $r(23) = .45$, $p = .025$. A stepwise regression analysis produced a significant model, $F(1, 23) = 8.13$, $p = .009$, with numeracy the sole variable entered in the model explaining 26% of the variance in matchstick algebra performance in the paper group. For the interactive group, performance on two Beta component tests, picture absurdities, $r(23) = .46$, $p = .021$, and matrix reasoning, $r(23) = .47$, $p = .019$, were most strongly correlated with matchstick algebra performance. A stepwise regression analysis produced a significant model, $F(1, 23) = 6.34$, $p = .019$, with matrix reasoning the sole variable retained in the model, explaining 22% of the variance.

Discussion

The average solution rate for each of the four types of problems in the paper group closely replicated the solution rates reported in Knoblich et al. (1999). That is, Type A problems were the easiest problems, producing significantly higher rates of success than Types B and C problems; Type C problems were the hardest. While Knoblich et al. did not formulate a prediction concerning the level of difficulty of Type D problems relative to Types B and C, they did predict that Type D problems, involving tighter perceptual chunks, would be harder to solve than Type A problems, involving the same type of value constraints but with looser perceptual chunks. Note that this is exactly the pattern of solution rates observed in the paper group. However, in the interactive group, the patterns in the solution rates departed substantially from

those in the paper group and from those reported in Knoblich et al. (1999). For one, solution rates for Type A problems were identical to the solution rates for Type B problems. Remarkably, the solution rates for Types A and D did not differ significantly in the interactive group. The tautology constraint in Type C problems was the hardest to relax in both groups.

Interactivity Matters

Interactivity encouraged a much higher rate of insight problem solving for all types of problems, with the exception of the easiest type of problems, problem A involving loose perceptual chunks and a low level constraint. Interactivity encourages the rearrangement of the matchsticks which generates configurations revealing novel affordances for action. For example, picking up the top horizontal stick of the equal sign creates a minus sign that may frame the action of where to place the stick in hand. Manipulation thus leads to opportunities that would otherwise require cognitive effort to identify. Key abilities for the purpose of this task may therefore involve the strategic manipulation of the sticks, and the ability to perceive and act upon affordances in that space. In turn, participants in the static paper condition are confronted with a permanent and perceptually immutable incorrect form of the equation, continually re-focusing attention and forcing the problem solver to attend to unhelpful information. The incorrect representation acts like a "rubber band" (Maglio, Matlock, Raphaely, Chernicky, & Kirsh, 1999): no matter how far participants can mentally morph the visual representation, the physical information exerts a form of conceptual gravity that pulls these mental efforts back to their starting point.

Physically moving a matchstick helps deconstruct chunks by creating opportunities to perceive the elements that make up the numerals. It also facilitates constraint relaxation by revealing opportunities for action that the new physical representation may afford. This in turn may encourage additional manipulation of the physical representation. Inevitably the physical representation of the problem will be modified from its original form. Changes in the problem representation initiate different activation patterns in long-term memory, cue different knowledge, and better position the reasoner to overcome an impasse.

Predictors of Performance

The insight problem solving success for participants in the paper group was best predicted by their level of numeracy assessed under timed conditions. The non-interactive nature of the task meant that participants in the paper group had to rely on their internal/mental computational abilities to simulate certain matchstick movements. The timed numeracy test likely used executive function capacity and, of course, arithmetic abilities, key mental resources to simulate algebraic transformations mentally.

In turn, performance in the interactive group was best predicted by the Matrix Reasoning component of the Beta III. This suggest that non-verbal, spatial and inductive reasoning aspects of fluid intelligence are important in determining matchstick performance in interactive insight problem solving; verbal and mathematical skills are no-longer the dominant predictors of success. Thus different contexts of reasoning engage different skills. These results invite a careful examination of the manner with which problem solving is investigated. The development of process models of problem solving for insight as well as for non-insight problems is inevitably predicated on a certain experimental procedure, which engages different cognitive abilities and strategies. The question becomes, which experimental procedure offers the more representative window onto problem solving occurring outside the laboratory? We believe one that fosters interactivity.

Complementary Strategies

Participants from both groups naturally employed complementary strategies to reduce cognitive demands and achieve insight. Interaction with both printed and physical numerals of the matchstick equations in both groups was rife. A large number of participants in the interactive group would be in constant contact with the sticks even when they were not being moved. Participants would rest their fingers on the magnetic sticks and run them across the sticks maintaining continuous contact. Tapping and touching of the sticks are examples of complementary strategies, focusing attention to the stick in question, like pointing a pen at an item on a written list (Kirsh, 1995a). Touching the sticks may also form a type of symbolic marking in which the contact is a concrete cue that there is something to remember about that stick (Kirsh, 1995a). Participants were also seen to pick a matchstick from the board and hold it in their hand for extended periods of time, potentially allowing them to predict the consequences of action from moving the stick, creating a new short term structure to the task (Kirsh, 1995b). Participants in the interactive group would also frequently move the matchsticks into novel positions, physically testing ideas before placing them back to their original position. Spatial re-configuration of the equations allowed participants to encode strategy, simplify the form of the equation, unveiling new affordances and opportunities to guide subsequent action.

Participants in the paper group also engaged in complementary strategies during problem solving. Participants would frequently be in contact with the printed Roman numerals: They would move their finger across the printed equation as if to guide or focus thought, often using their finger to represent a matchstick, mimicking rotations and movements to aid visualisation and test spatial configurations. Some would frequently hover over the numerals, as though a close proximity to the numerals was

necessary. Others would tap the printed numerals. The use of hands in the paper group may reflect an attempt to materialise mental projections.

How People Think

The experimental procedure developed in the study reported here coupled people with artefacts in the process of thinking. And while the problem solving task remained constrained and artificial, we would argue that the interactive methodology employed here offers a much closer approximation of real-world problem solving behaviour than the non-interactive procedure initially employed in Knoblich et al. (1999). Improved performance in the interactive group should not be interpreted to call into question Ohlsson's (1992) representational change theory of insight: It remains a productive characterization of the processes involved in overcoming an impasse and achieving insight.

However, what the data presented here strongly suggest is that such theories must be examined with experimental procedures that encourage the construction and modification of distributed problem representations. These distributed representations recruit resources that are internal and external to the thinking agent; the control over behavior is also distributed among internal and external factors. The patterns in the correlations between test of cognitive abilities and performance with the matchstick algebra problems converge on the notion that designing interactive versions of these tasks is not simply an exercise in making things more concrete to facilitate reasoning. Rather, making the task concrete to foster interactivity inevitably engages a different set of cognitive, perceptual and motor skills. The data reported here encourage the design of experimental environments that capture problem solving as situated, embedded, and embodied activities, which are likely more representative of the manner people think and behave. To be sure, interactivity introduces a large number of degrees of freedom which reduce experimental control. But it also offers much richer data from which to infer the reasoning mechanisms at play when solving problems, data that inform how reasoning outside the laboratory proceeds.

References

Cowley, S., & Vallée-Tourangeau, F. (2010). Thinking in action. *AI & Society*, *25*, 469–475.

Giere, R. (2006). *Scientific perspectivism*. University of Chicago Press.

Kellog, C. E., & Morton, N. W. (1999). *Beta III Manual*. The Psychological Corporation, A Hartcourt Assessment Company.

Kirsh, D. (1995a). Complementary strategies: why we use our hands when we think. In J. Moore, & J. Fain Leman (Ed.), *Proceedings of the Seventeenth Annual Conference of the Cognitive Science Society,* (pp. 212-217). Mahwah: NJ: Lawrence Erlbaum

Kirsh, D. (1995b). The intelligent use of space. *Artificial Intelligence*, *73*, 31-68.

Kirsh, D. (1996). Adapting the environment instead of oneself. *Adaptive Behavior*, *4*, 415-452.

Kirsh, D. (2009). Interaction, external representation and sense making. In N. A. Taatgen, & H. v. Rijn (Ed.), *Proceedings of the Thirty First Annual Conference of the Cognitive Science Society* (pp. 1103-1108). Austin, TX: Cognitive Science Society.

Knoblich, G., Ohlsson, S., Haider, H., & Rhenius, D. (1999). Constraint relaxation and chunk decomposition in insight problem solving. *Journal of Experimental Psychology*: *Learning, Memory and Cognition*, *25*, 1534-1555.

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

Maglio, P. P., Matlock, T., Raphaely, D., Chernicky, B., & Kirsh, D. (1999). Interactive skills in Scrabble. In M. Hahn, & S. C. Stoness (Eds.), *Proceedings of the Twenty-First Conference of the Cognitive Science Society* (pp. 326-330). Mahwah, NJ: Lawrence Erlbaum Associates, Publishers.

Nelson. H. E. (1991). *National Adult Reading Test* (2nd Ed.). Windsor, Berkshire: The NFER-NELSON Publishing Company Ltd.

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

Ohlsson, S. (1992). Information processing explanation of insight and related phenomena. In M. T. Keane and K. J. Gilhooly (Eds.), *Advances in the psychology of thinking, volume 1* (pp.1-44). London: Harvester Wheatsheaf.

Segal, E. (2004). Incubation in insight problem solving. *Creativity Research Journal, 16*, 141-148

Wertheimer, M. (1959). *Productive thinking* (enlarged edition). New York: Harper & Brothers Publishers.