THINKING & REASONING, 2011, 17 (4), 424–439

Psychology Press
Taylor & Francis Group

Interactive insight problem solving

Anna Weller, Gaëlle Villejoubert, and Frédéric Vallée-Tourangeau

Department of Psychology, Kingston University London, Kingston-upon-Thames, UK

Insight problem solving was investigated with the matchstick algebra problems developed by Knoblich, Ohlsson, Haider, and Rhenius (1999). These problems are false equations expressed with Roman numerals that can be made true by moving one matchstick. In a first group participants examined a static two-dimensional representation of the false algebraic expression and told the experimenter which matchstick should be moved. In a second group, participants interacted with a three-dimensional representation of the false equation. Success rates in the static group for different problem types replicated the pattern of data reported in Knoblich et al. (1999). However, participants in the interactive group were significantly more likely to achieve insight. Problem-solving success in the static group was best predicted by performance on a test of numeracy, whereas in the interactive group it was best predicted by performance on a test of visuo-spatial reasoning. Implications for process models of problem solving are discussed.

Keywords: Distributed cognition; Interactivity; Individual differences; Problem solving.

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. Insight problems are different in that the goal state, or

Correspondence should be addressed to Frédéric Vallée-Tourangeau, Department of Psychology, Kingston University, Kingston-upon-Thames, KT1 2EE, UK.
E-mail: f.vallee-tourangeau@kingston.ac.uk

We thank Susan Cook for her help designing and producing the material used in this experiment, and two anonymous reviewers for helpful comments on a previous version of this manuscript.

http://www.psypress.com/tar http://dx.doi.org/10.1080/13546783.2011.629081

resolution, is initially not visible or imaginable. With insight problems most participants initially experience an impasse from which they may or may not emerge to formulate a solution to the problem. 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 to anticipate the solution. Overcoming an impasse is understood to be driven by a representational change that re-casts the relationship among the elements of the representation or that redefines 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, which recruits long-term memory 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. Constraint relaxation plays an important role in the solution of the nine-dot problem (Maier, 1930; although a number of other factors have also been implicated in impasse resolution with this problem, see Kershaw & Ohlsson, 2004). In this problem the goal is to link all nine dots with four continuous lines without lifting the pen from the paper. The perceptual configuration of the dots imposes an implicit constraint that the pen can only draw lines within the projected perimeter delineated by the dots. Relaxing that constraint, allowing the possibility of extending the lines beyond the frame, can foster an appreciation of how such four lines can be drawn to link all nine dots.

In turn, a well-known 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). In this problem the reasoner must calculate the area of a 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 realise that the square–parallelogram perceptual 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.

MATCHSTICK ALGEBRA

Constraint relaxation and chunk decomposition are important drivers of representational restructuring, which is deemed necessary to overcome impasse and achieve insight. These 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 written with Roman numerals. Participants are required to move (by moving to a different location, rotating, sliding) one stick to make the equation true, with the "V" and "X" numerals each consisting of two slanted sticks (the movement of which can transform one into the other and vice versa). 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 the equal sign 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 that form numerals and symbols (operators).

To test the importance of constraint relaxation Knoblich et al. developed three types of false statements the solution for which required relaxing constraints of different scopes (see Table 1). Solving Type A problems involves 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 involves relaxing a constraint with a broader scope; that is, one including the constraint on manipulating operators (operator constraint). Solving Type C problems involves relaxing a constraint with an even

TABLE 1
The four matchstick algebra problem types developed by Knoblich et al. (1999)

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

Solutions for problems for Type A through C require relaxing constraints of increasing scope, while solving problems of Type D involves decomposing a tight perceptual chunk.

broader scope, namely the constraint that people rarely communicate in tautological terms (tautology constraint). Hence, to solve these problems, participants must realise that tautologies are acceptable. Knoblich et al. predicted that the solution rate 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; the hardest problems were Type C.

The authors also tested the importance of chunk decomposition in solving these problems. They developed a fourth type of problems, Type D, by taking Type A problems with their narrow value constraint but used Roman numerals that are more perceptually complex, forming tighter perceptual chunks (see Table 1; in this example the solution involves decomposing the "X" perceptual chunk into a "V"). Knoblich et al. predicted that problems of Type D would be 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 voiced their proposed solution, which was then noted by the experimenter. (In an eye-tracking follow-up study immobility was ensured with a bite bar that participants had to stop biting to announce their proposed solution out loud; Knoblich, Ohlsson, & Raney, 2001.) Yet, from a distributed cognitive system perspective, thinking is the product of an interactive assemblage of resources internal and external to the reasoning agent (Cowley & Vallée-Tourangeau, 2010, Giere, 2006; Wilson & Clark, 2009). Outside the cognitive psychologist's laboratory, the environment and its content can be exploited to facilitate reasoning and problem solving in a variety of ways. People naturally seek to augment their thinking abilities by constructing cognitive extensions; what Wilson and Clark (2009, p. 58) refer to as a "kind of intellectual niche construction" (cf. Laland, Odling-Smee, & Feldman, 2000). To this end a diverse range of actions are performed, some to reorganise the environment, some to recruit or design artefacts, while others complement thinking (e.g., muttering to oneself, pointing; Kirsh 1995a, 2009). External artefacts and representations are employed as vehicles for ideas and hypotheses, lightening cognitive load (Zhang & Norman, 1994). However, these external resources do not simply function as a means to offload memories and reduce cognitive strain. 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 opportunities to guide behaviour. Spatial rearrangement not only aids problem solving by literally changing the face of the problem representation, but it can also make aspects or objects within the task more perceptible (Kirsh, 1995b). 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 computational effort, 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).

Fioratou and Cowley (2009) used the cheap necklace problem to explore the role of artefacts in contributing to insight. The researchers split participants into static and interactive groups, for both of whom the aim was to construct a long chain from four shorter chains consisting of three links. The cost structure was such that it would cost participants 2 cents to open a link and 3 cents to close one, with the aim to complete the task having spent no more than 15 cents. Participants in the static group were shown a diagram of the four chains, while participants in the interactive group were given actual link chains to freely manipulate. Interactive participants performed significantly better on the task than static participants. Fioratou and Cowley suggest that participants in the interactive group can capitalise on felicitous but largely non-strategic manipulations of the artefacts to determine how to proceed to a solution. That is, an action may not always be premeditated, but may still guide attention and thinking along productive paths. Artefacts themselves are not part of cognition, but form part of a wider, extended, cognitive system.

THE PRESENT STUDY

We investigated 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; Bright, Jaldow, & Kopelman, 2002) and visuo-spatial reasoning abilities with the Beta III test (Kellog & Morton, 1999). 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

A total of 50 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 static group or the interactive group. In the static group participants were presented matchstick algebra problems on a piece 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 1-minute period. This test consisted of simple arithmetic questions (such as $18 - 6 = ?$). 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 (on the basis of a pronunciation guide provided with the test manual). Participants were then asked to complete the Roman numerals test in which they were required to translate a series of 24 simple Arabic numbers (ranging between 1 and 12) into their Roman numeral equivalent within a 1-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; see Table 2). The order of presentation was randomised for each participant. Each equation was printed in large black font in the centre of a sheet of white A4 paper held in a ring binder

TABLE 2
Matchstick algebra equations presented to participants in terms of equation type, number, and solution

Type	*Equation*	*Solution*
A1	VI = VII + I	VII = VI + I
A2	IV = III + III	VI = III + III
A3	II = III + I	III = II + I
A4	IX = VIII + III	XI = VIII + III
B1	I = II + II	I = III − II
B2	IV = III − I	IV − III = I
B3	III = V + III	III = VI − III
B4	V = III − II	V − III = II
C1	III = III + III	III = III = III
C2	IV = IV + IV	IV = IV = IV
D1	XI = III + III	VI = III + III
D2	VI = VIII + III	XI = VIII + III

Taken from Knoblich et al. (1999).

with the following instruction at the head of each page, "Move **ONE** stick to make the equation **TRUE**". Participants in the static group were asked to solve the equations using these sheets of paper only. For the interactive group, we designed a magnetic board (27 cm × 21 cm) on which participants created and modified Roman numerals and algebraic statements using magnetised matchsticks (.5 cm × 4.5 cm). 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 while reasoning about the problems. As in the original study by Knoblich et al. (1999), all participants were informed that they could move the stick to a different location, rotate it or slide it to make the equation true. 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 Beta III test. Participants were given an answer booklet consisting of the five component tests (Coding, Picture Completion, Clerical Checking, Picture Absurdities, and Matrix Reasoning) completed under timed condition. Each section of the test contained some practice problems to familiarise participants with the test procedure. The Coding test required participants to match a series of symbols to numerals demonstrated at the top of the page (test duration: 120 s). The Picture Completion section consisted of a series of pictures with aspects/items missing that participants must complete (180 s). The Clerical Checking test displayed a number of paired lists of symbols or numbers, some of which were identical and others not: Participants were required to identify the identical and non-identical pairs (120 s). The Picture Absurdities test consisted of a series of panelled picture sets, and required participants to identify which picture showed something absurd or illogical (180 s). Finally, in the Matrix Reasoning test participants chose a picture from a selection of five to complete sequences of abstract shapes (300 s).

Measures. Both the maths and Roman numerals tests were expressed in terms of percent correct answers. The NART was reported as the raw error score. 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 and out of the possible 12. Each element of the Beta III test was first scored individually, by summing the correct answers. These scores were then converted into age corrected scaled scores (ACSS; Kellog & Morton, 1999).

RESULTS

The mean percent solution for each problem type is displayed in Figure 1. Solution rates appeared marginally greater in the static 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 static counterparts. A 4 (problem type: A, B, C, D) by 2 (group: static, interactive) mixed

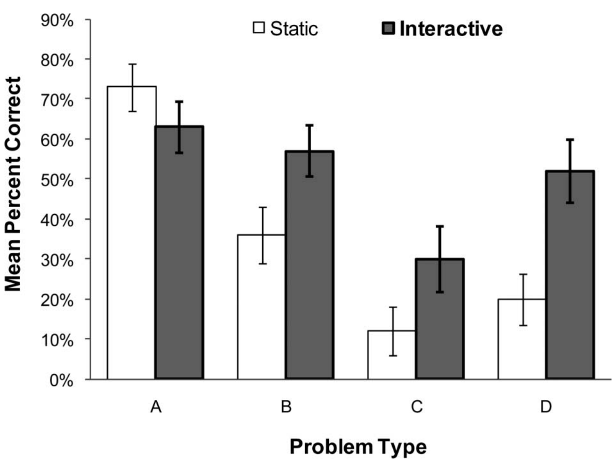

Figure 1. Mean percentage of correctly solved problems of Types A, B, C, and D in the static group (open bars) and the interactive group (dark bars). Error bars are standard errors of the mean.

analysis of variance (ANOVA) revealed a significant main effect of problem type, $F(3,\ 144) = 24.6$, $MSE = .079$, $p < .001$, a significant main effect of group, $F(1,\ 48) = 5.06$, $MSE = .230$, $p = .029$, and a significant interaction between problem type and group, $F(3,\ 144) = 5.03$, $MSE = .079$, $p = .002$.

Separate ANOVAs were conducted for the paper and interactive groups. In the paper group the overall problem type main effect was significant, $F(3,\ 72) = 29.2$, $MSE = .063$, $p < .001$. 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 ($p = .002$). As for the interactive group, the overall problem type main effect was also significant, $F(3,\ 72) = 5.39$, $MSE = .096$, $p = .002$. However, 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 ($p = .03$).

Predictors of performance

Mean performance on the knowledge and cognitive abilities tests are reported in Table 3. Numeracy did not differ significantly between the static 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 static group ($M = 48.7$, $SD = 23.2$) and the interactive

TABLE 3
Means and standard deviations for scores various tests of skills and knowledge

	Static		*Interactive*	
	M	*SD*	*M*	*SD*
Numeracy (%)	49.5	25.9	51.9	22.0
Roman numerals (%)	48.7	23.2	43.6	20.8
NART score	26.4	6.1	24.3	6.6
Beta III				
Coding	9.0	2.2	8.8	2.7
Picture completion	3.8	1.6	4.0	1.9
Clerical checking	9.6	2.4	8.8	2.4
Picture absurdities	7.9	2.2	8.2	2.6
Matrix reasoning	10.9	2.4	10.9	2.4

$N = 25$ in both groups. The results for each of the five component tests of the Beta III are reported as age corrected scaled scores.

Means and standard deviations for scores various tests of skills and knowledge, including a numeracy test, a roman numerals test, the National Adult Reading Test, and the five component tests of the Beta III.

group ($M = 43.6$, $SD = 20.8$), $t(48) = 0.83$, $p = .41$. Performance on the NART did not differ between participants in the static 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.

Performance on some of the cognitive abilities tests correlated with overall performance on the matchstick algebra task (see Table 4). For the static group participants, matchstick algebra performance was most strongly correlated with numeracy $r(23) = .51$, $p = .009$ and performance on the NART, $r(23) = -.45$, $p = .025$. A stepwise regression analysis produced a significant model, $F(1, 23) = 8.13$, $MSE = 6.81$, $p = .009$, with numeracy the sole variable entered in the model explaining 26% of the variance in matchstick algebra performance in the static 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, $F(1, 23) = 6.34$, $MSE = 7.56$, $p = .019$, with matrix reasoning the sole variable retained in the model, explaining 22% of the variance in matchstick algebra performance.

DISCUSSION

This experiment explored insight problem solving using the matchstick algebra problems presented in Knoblich et al. (1999) comparing participants who were either invited to manipulate matchsticks to solve the problems with those who had to reason about the problems by relying on nothing

TABLE 4
Pearson correlation coefficients

	Static	*Interactive*
Numeracy	.51**	.26
Roman numerals	.08	.25
NART score	−.45*	−.22
Beta III		
Coding	.11	.22
Picture completion	.07	.32
Clerical checking	.26	.38
Picture absurdities	.37	.46*
Matrix reasoning	.43*	.47*

*$p < .05$; **$p < .01$.

Pearson correlation coefficients for the relationships between tests of abilities and proportion of correctly solved matchstick algebra problems (out of 12).

other than their own internal cognitive resources (although we discuss below how participants in the static group also engaged in some complementary actions). Physically manoeuvring and manipulating the matchsticks within the equation transformed the abstract task representation of matchstick algebra into a simplified, more explicit form. For instance, seeing the Roman numeral "X" presented as two three-dimensional matchsticks as opposed to a printed, continuous numeral presents the information in a more cognitively convenient manner. Physical manipulation of the equations likely allowed participants to offload cognitive effort and memory onto the sticks and physically encode the progress of their problem-solving effort.

The pattern in the mean percentage success for each of the four types of problems in the static group closely replicated the pattern reported in Knoblich et al. (1999). That is, Type A problems were the easiest problems, fostering 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, consisting of the same type of constraint but with looser perceptual chunks. Note that this is exactly the pattern of solution rates observed in the static group. However, in the interactive group, the patterns in the solution rates departed substantially from those in the static 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 greater level of difficulty of Type C problems was observed in both groups. While the solution rate for Type C problems (30%) in the interactive group was 2.5 times the rate observed in the paper group (12%), type C problems were solved significantly less frequently than Type B problem in the interactive group. Finally, the largest absolute improvement in problem-solving performance with the interactive procedure was observed for Type D problems, which involve the decomposition of tight perceptual chunks. Clearly the impact of interactivity was greatest for Type C and D problems. Note that these were represented with two instances each, as in the original Knoblich et al. (1999) procedure. Future research efforts on interactivity and matchstick algebra should proceed on the basis of a larger set of matchstick problems, one in which the frequency of the four types is kept constant.

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, Type A, involving loose perceptual chunks and a relatively narrow constraint. Interactivity encourages spatial rearrangement of the matchsticks that generates configurations revealing novel affordances for action to the problem solver. For example, if a participant were to pick up the top horizontal stick of the equal sign, the remaining stick now forms a minus sign that may frame the action of where to place the stick in hand; therefore manipulation 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. Of course, movement may not be strategic but may inadvertently direct attention to hitherto unnoticed aspects of the problem representation, guiding problem-solving efforts along a more fruitful path (cf. Grant & Spivey, 2003). In turn, participants in the static paper condition are confronted with a permanent and perceptually immutable incorrect form of the equation that continually re-focuses attention, forcing them to attend to unhelpful information. The incorrect representation acts like a "rubber band" (Maglio, Matlock, Raphaely, Chernicky, & Kirsh, 1999, p. 2): 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 of the problem. This dynamic "perception-action cycle" (Wilson & Clark, 2009) inevitably transforms the perceptual representation of the problem, initiating different activation patterns in long term memory, cueing different knowledge structures that may encourage insight (Knoblich et al., 1999).

Predictors of performance

The insight problem-solving success for participants in the static 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 static 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 (Gottfredson & Saklofske, 2009; Hayashi, Kato, Igarashi, & Kashima, 2008) 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 psychologists investigate problem solving and explain performance. The development of process models of problem solving for insight as well as for non-insight problems (e.g., Jeffries, Polson, Razran, & Atwood, 1977) is inevitably predicated on a certain experimental procedure, which implicates different cognitive abilities and strategies.

Qualitative observations

Participants from both groups naturally employed complementary actions to reduce cognitive demands, weaken perceptual constraints, and decompose chunks to aid problem solving. 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 literally providing a cue that there is something to remember about that stick (Kirsh, 1995b). 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 and providing 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 in their original position. Spatial reconfiguration 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 static 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 continuously hover over the numerals, as though close proximity to the numerals was necessary. Others would tap the printed numerals. The use of hands in this condition may be an attempt by participants to materialise mental projections; a materialised projection can after all be manipulated and tested to reveal opportunities to aid insight (Kirsh, 2009).

A minority of participants in this condition would sit back and stare at the paper, making no contact with it unless indicating a potential solution. Yet even these candidates were noted to have, at some point, tapped or rested their fingers on the sticks. A few of these candidates would still speak out loud, relying less on physical movement but still on verbalisation to shoulder some of the cognitive demands of the task. One participant reached for a nearby pen, and started moving it between the sticks to demonstrate potential movements before it was confiscated. Several other candidates also reached for a pen, but were informed it was forbidden. This reflects the natural inclination to utilise physical resources to shoulder cognitive strain, and generate external representations that can be subjected to manipulation to guide hypothesis testing.

A representative window onto problem solving

The experimental procedure developed in the study reported here coupled people with artefacts in the process of thinking. In the interactive group the nature of the artefacts, and the movements they afforded, configured what Wilson and Clark (2009, p. 65) refer to as a "transient extended cognitive system", a "soft-assembled" system that interweaves participants' cognitive resources with external artefacts, fostering a broader range of ideas and actions, and augmenting problem-solving performance. 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: It remains a productive characterisation of the processes involved in overcoming impasse and achieving insight.

However, what the data presented here strongly suggest is that such a theory must be examined with experimental procedures that encourage the construction of distributed problem representations. These distributed representations engage thinking processes that recruit resources that are internal and external to the thinking agent. As a result, the control over behaviour 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, the concreteness of these tasks and the necessary interactivity engage 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 representative of the manner people think and behave. To be sure, interactivity introduces a large number of degrees of freedom, which reduce the psychologist's control over the experimental environment, but it also offers a much richer set of data from which to infer the reasoning mechanisms at play when solving problems, mechanisms that can inform how reasoning outside the laboratory environment proceeds.

Finally, interactivity and the distributed nature of the thinking processes cast neuroscience efforts in a new light. The impressive spatial resolution obtained with some neuroimaging technology, such as with functional magnetic resonance imaging, is predicated on an immobile agent, with interactivity kept to a minimal level for fear of contaminating the data. There is thus an important question concerning the representative nature of cognition when thinking is de-coupled from a physical environment, particularly when investigating problem solving.

Manuscript received 10 January 2011
Revised manuscript received 9 June 2011

REFERENCES

Bright, P., Jaldow, E., & Kopelman, M. D. (2002). The National Adult Reading Test as a measure of premorbid intelligence: A comparison with estimates derived from demographic variables. *Journal of the International Neuropsychological Society*, *8*, 847–854.

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

Fioratou, E., & Cowley, S. (2009). Insightful thinking: Cognitive dynamics and material artifacts. *Pragmatics & Cognition*, *17*, 549–572.

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

Gottfredson, L., & Saklofske, D. H. (2009). Intelligence: Foundations and issues in assessment. *Canadian Psychology*, *50*, 183–195.

Grant, E. R., & Spivey, M. J. (2003). Eye movements and problem solving: Guiding attention guides thought. *Psychological Science*, *14*, 462–466.

Hayashi, M., Kato, M., Igarashi, K., & Kashima, H. (2008). Superior fluid intelligence in children with Asperger's disorder. *Brain and Cognition*, *66*, 306–310.

Jeffries, R., Polson, P. G., Razran, L., & Atwood, M. E. (1977). A process model for Missionaries–Cannibals and other river-crossing problems. *Cognitive Psychology*, *9*, 412–440.

Kellog, C. E., & Morton, N. W. (1999). *Beta III manual*. New York: The Psychological Corporation, A Harcourt Assessment Company.

Kershaw, T. C., & Ohlsson, S. (2004). Multiple causes of difficulty in insight: The case of the nine-dot problem. *Journal of Experimental Psychology: Learning, Memory and Cognition*, *30*, 3–13.

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 Associates Inc.

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.

Knoblich, G., Ohlsson, S., & Raney, G. E. (2001). An eye movement study of insight problem solving. *Memory & Cognition*, *29*, 1000–1009.

Laland, K. N., Odling-Smee, J., & Feldman, M. W. (2000). Niche construction, biological evolution, and cultural change. *Behavioral and Brain Sciences*, *23*, 131–175.

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 Inc.

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

Nelson, H. E. (1991). *National Adult Reading Test, second edition*. Windsor, UK: 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 & 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.

Wilson, R. A. (1994). Wide computationalism. *Mind*, *103*, 351–372.

Wilson, R. A., & Clark, A. (2009). How to situate cognition: Letting nature take its course. In P. Robbins & M. Aydede (Eds.), *The Cambridge handbook of situated cognition* (pp. 55–77). Cambridge, UK: Cambridge University Press.

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