Journal of Cognitive Psychology
Publication details, including instructions for authors and subscription information:
http://www.tandfonline.com/loi/pecp21

Cognitive change in insight problem solving: Initial model errors and counterexamples

M. Aisling Murray [a] & Ruth M. J. Byrne [a]
[a] Trinity College Dublin, University of Dublin, Dublin, Ireland

Version of record first published: 15 Jan 2013.

Journal of Cognitive Psychology, 2013
http://dx.doi.org/10.1080/20445911.2012.743986

Cognitive change in insight problem solving: Initial model errors and counterexamples

M. Aisling Murray and Ruth M. J. Byrne

Trinity College Dublin, University of Dublin, Dublin, Ireland

We report the results of four experiments that examined the cognitive changes that occur in problem solvers' mental models of insight problems. The experiments showed that participants produced more correct solutions to insight problems that required single steps than multiple steps. Experiment 1 showed that their diagrams and explanations corresponded to initial model errors. Experiment 2 found more correct solutions for problems reworded to enable the retrieval of counterexamples to common assumptions. Experiment 3 found more correct solutions when physical props enabled the construction of a counterexample to the initial erroneous model and also to subsequent erroneous models. Experiment 4 showed more correct solutions when physical props limited the subsequent possibilities. The implications of the results for alternative theories of insight problem solving are discussed.

Keywords: Cognitive change; Insight; Mental models; Problem solving.

When people solve novel problems they sometimes experience an "aha" moment of cognitive change, a subjective experience of "insight" that reflects a sudden transition from bafflement to enlightenment (e.g., Chu & MacGregor, 2011; Ohlsson, 2011; Sio & Ormerod, 2009). Consider this problem: Describe how to throw a ping-pong ball so that it will travel a short distance, come to a dead stop and then reverse itself. You are not allowed to bounce it off any surface or tie anything to it (e.g., Ansburg & Dominowski, 2000). Suggestions include putting a back-spin on the ball, which will not work, or throwing it to another person, which violates the requirement for the ball to stop and return by itself. The correct solution is to throw the ball up and gravity will cause it to stop for an instant before falling back down. The solution requires the insight that the ball can travel vertically. Once the insight is achieved, the solution follows readily. We consider these problems to be "single-step" problems.

We distinguish them from "multiple-step" problems, such as the following: Consider how to form four equilateral triangles using six matches of equal length. Each complete match must form one complete side of a triangle (e.g., Scheerer, 1963). Attempts to solve the problem include constructing triangles flat on a table in two dimensions. The solution requires the insight that the triangles must be formed in three dimensions. But even after the insight is achieved, the solution requires further steps to be executed accurately, e.g., one triangle can be constructed from three matches to form the base on the horizontal plane, the other triangles can be constructed from the remaining three matches to form the three sides of a pyramid shape in the vertical plane. Our aim is to report a series of

Correspondence should be addressed to Aisling Murray, School of Psychology and Institute of Neuroscience, Trinity College Dublin, University of Dublin, Ireland. E-mail: murray.aisling@gmail.com

The research was funded by the Irish Research Council for the Humanities and Social Sciences through a Government of Ireland PhD Scholarship and by Enterprise Ireland through a PhD award, both to MAM. We thank Marco Ragni for helpful comments on an earlier draft of the paper.

experiments that test novel proposals based on a mental model theory of insight problem solving.

INSIGHT AND INITIAL MENTAL MODELS

Our first proposal is that people solve insight problems by constructing an initial mental model of the problem. Their models are guided by a principle of truth—they mentally represent what is true rather than what is false, and a prinicple of parsimony—they represent few possibilities, perhaps because of working memory constraints (e.g., Johnson-Laird & Byrne, 2002). They envisage a possibility that corresponds to how the world would be, according to the problem description, e.g., they think about throwing a ping-pong ball. Their simulation incorporates assumptions about the normal way the event occurs, e.g., a ping-pong ball usually travels horizontally. The assumption of an object's usual function may be resistant to change and its mental representation may lead to "functional fixedness" (e.g., Duncker, 1945).

Insight problems are represented initially in a way that does not allow access to the knowledge needed to solve the problem (e.g., Weisberg & Alba, 1981). Assumptions are added to the initial model implicitly, rather than through explicit deliberative processes, e.g., the horizontal trajectory of the ping-pong ball is presupposed implicitly, rather than identifed through a process of reflection (e.g., Helie & Sun, 2010). The initial representation leads participants to encounter an "impasse" in their attempts to solve the problem (e.g., Ohlsson, 2011). Our first experiment examines the initial model errors that people made for a variety of insight problems.

INSIGHT AND COUNTEREXAMPLES

Our second proposal is that people modify their initial model when they encounter a counterexample. People rely on counterexamples to reason in many domains (e.g., Byrne, 2005; De Neys, Schaeken, & d'Ydewalle, 2005). If they are provided with a counterexample, they can change the assumptions in their initial model, e.g., about a ball's normal trajectory. Consider the problem of how to throw a *basketball* so that it will go a short distance, come to a dead stop, and then reverse itself. Basketballs often follow a vertical trajectory when in play and they are heavier than ping-pong balls, and so these counterexamples ensure that the unhelpful constraint of a horizontal trajectory is not added to the initial model.

A "representational change" is required for an impasse to be overcome (Ohlsson, 2011). Consider an equation made out of matchsticks: VII = V – II. How can the equation be balanced by moving just one matchstick and without taking any away? Attempts include moving one of the "I"s. The correct solution is VII – V = II, achieved through the insight that a matchstick comprising the "equals" sign can be moved (e.g., Knoblich, Ohlsson, & Raney, 2001). The initial model incorporates the assumption that only matchsticks comprising numbers can be moved. Restructuring the representation can occur through processes such as "constraint-relaxation"—removing unnecessary limits such as the assumption about moving numbers, and "chunk-decomposition"—breaking up elements of a perceptual chunk, e.g., the "chunk" of a "V". It can occur through "elaboration"—adding new information to the representation, and "re-encoding"—correcting a mistake (e.g., Knoblich, Ohlsson, Haider, & Rheinus 1999). Our second experiment shows that people make more correct solutions when they are provided with counterexamples to the usual initial model errors.

SINGLE- AND MULTIPLE-STEP INSIGHT PROBLEMS

Our third proposal is that for "single-step" insight problems, the modified initial model makes the solution readily available, but for "multiple-step" problems, further steps are necessary. Consider nine dots arranged in a formation of three rows of three across with equal spaces between them, a 3×3 matrix. How can you join all nine dots using just four connected straight lines, without retracing any part of any line or raising your pen from the page once you have started on the first line? Common attempts include maximising progress, e.g., joining the maximum number of dots (e.g., Chronicle, MacGregor, & Ormerod, 2004). The correct solution requires the insight that lines can be extended beyond the square shape. But the insight does not lead directly to a solution (Weisberg & Alba, 1981). There are multiple alternative possibilities that follow, and it may require further insights to choose between them (e.g., Suzuki, Abe, Hiraki, & Miyazaki, 2001).

The solution also requires searching through a problem space of possible moves, as well as the selection and evaluation of moves, just as for well-defined permutation problems (e.g., Kaplan & Simon, 1990; McGregor, Ormerod, & Chronicle, 2001). Individuals differ in their abilities to solve insight problems perhaps because of differences in working memory or attention-switching skills (Gilhooly & Fioratiou, 2009; Murray & Byrne, 2005). Participants engaged in "progress monitoring" do not simulate enough possibilities in a "look ahead" through moves, to foresee that maximum dot cancellation cannot be achieved within the square shape. They need to experience "criterion-failure" in order to relax constraints (e.g., Chronicle et al., 2004). Our third and fourth experiments show that people make more correct solutions when they are provided with physical props with which they can interact (Weller, Villejoubert, & Vallee-Tourangeau, 2011). The physical props enable them to encounter counterexamples to the subsequent alternative possibilities that follow from the initial model, and to choose between these alternative possibilities. All four experiments show that single-step problems are easier than multiple-step problems.

EXPERIMENT 1

The aim was to examine the models that people construct for single-step problems and multiple-step problems. For each problem we identified an "initial model" error that results from constructing a model of the situation as it normally occurs, which adds a constraint, e.g., the ping-pong ball must travel horizontally (see the Appendix). We examined initial model errors by requiring participants to solve insight problems and to construct external representations: Explain the problem to a friend and draw a diagram of it.

Method

Materials and design. Participants were given a set of eight problems (adapted from Ansburg & Dominowski, 2000; Duncker, 1945; Scheerer, 1963; see the Appendix). Each participant attempted to solve the problem and generated an external representation of it. For half the problems, participants were asked to draw a diagram and for the other half to explain it "as if to a friend", and to control for order effects; half the participants drew diagrams for the first four problems and provided explanations for the last four, and the other half did the tasks in the opposite order. In addition, half the participants were instructed to attempt to solve the problem first, and then to generate an external representation of it; the other half received the tasks in the opposite order.

Participants and procedure. The participants were 26 psychology undergraduates from Trinity College Dublin, six men and 20 women, aged between 18 and 45 with a mean age of 22, who volunteered to participate for course credits. They were tested in small groups of two or three individuals. They were instructed by way of an example. Each problem was displayed on a screen for at least 30 s. Participants had a further 1 minute, 30 s to solve the problem.[1] For the external representation task, they were instructed, "For some of the problems you will be asked to draw a simple diagram of the problem. For others you will be asked to explain the problem in writing as if you were giving it to a friend, giving as much detail as possible but without revealing the answer directly". They were given 1 minute to do so. They were asked to complete the problems in the order they were given, not to skip any, and not to change an answer once they had completed it. The problems were presented in a different randomised order in each testing session.

Results and discussion

Participants solved more single-step than multiple-step problems overall, 50% versus 24%, as shown by a Wilcoxon's signed ranks test, $z = 3.93$, $r = .77$, $p < .0005$.[2] Their external representations contained the conjectured initial model error for

[1] Each problem was presented for a period of time (range 30–70 s) determined in a pilot study where eight volunteers read each problem aloud slowly and carefully and their average reading time was doubled. Previous studies have shown that 97% of correct solutions occur within 2 minutes, including time to read the problem (Lockhart, Lamon, & Gick, 1988).

[2] All statistical tests in the experiments are one-tailed unless otherwise indicated.

87 out of the 96 incorrect solutions, and for only eight out of the 64 correct solutions.[3]

Mann Whitney U-tests showed no order effects for generating an external representation first, or attempting to solve the problem first, $z < 1$. There were no order effects for generating a diagram first or an explanation first, $z < 1$. The experiment shows that multiple-step problems are harder to solve than single-step problems, and that people produced diagrams and explanations that corresponded to initial model errors.

EXPERIMENT 2

The aim was to enable people to construct an accurate initial model of the problems, by ensuring that they encountered a counterexample to the initial model error, such as referring to a "basketball" instead of a "ping-pong ball". We expected that counterexamples would be more effective for single-step problems than multiple-step problems, because multiple-step problems require further work even from an accurate initial model. We also asked participants to suggest a hint for a friend attempting to solve the problem, after the participant knew the solution, to examine further their initial model errors.

Method

Design and materials. Participants were given two of the single-step and two of the multiple-step problems from the previous experiment. The control group received the standard wording; the experimental group received improved wording. For the ping-pong ball, the improved wording referred to "basketball"; for the window cleaner, the improved wording removed the reference to "high-rise" (describing the office building) and "60-foot" (describing the ladder); for the six matches, the improved wording was "show how a shape could be formed" instead of "show how they can be used to form a shape". For the nine-dot, the standard version referred to nine dots, the improved version referred to nine aliens in a computer game. Each participant completed the problems in a different randomised order. For each problem, participants attempted to solve it, and then were told the solution and asked to imagine they were to provide a suitable hint to a friend on a game show who is allowed to ring for help with the problem, but who cannot be told the answer directly.

Participants and procedure. The participants were 38 members of the Trinity College Dublin School of Psychology participant panel (members of the general public recruited through national newspaper advertisements), who were paid a nominal fee of 8 euro per hour. There were 28 women and nine men (and one did not provide information). They were aged between 27 and 74 years, with a mean age of 54 (and two did not provide information). They were assigned at random to the control group who received the standard wording ($n = 22$) or the experimental group who received the improved wording ($n = 16$). The procedure was the same as the previous experiment. Participants were allowed 2 minutes to read the problem and attempt a solution, and then a further minute to read the presented solution and write down a hint.

Results and discussion

Participants given the improved wording solved more problems compared to those given the standard wording, as shown by Mann Whitney U tests, for all problems overall, $z = 2.019$, $r = .328$, $p < .05$, as Table 1 shows. The effect occurred for single-step problems, $z = 1.863$, $r = .302$, $p < .05$, although not for multiple-step problems, $z < 1$, which were rarely solved in either condition perhaps because of the time limits, or because the improved wording for these problems simply was not good enough.[4]

Once again, participants solved more single-step than multiple-step problems, as shown by Wilcoxon's tests, for all problems overall, $z = 4.138$, $r = .672$, $p < .0005$; for the standard wording, $z = 2.762$, $r = .589$, $p < .01$; and for the improved wording, $z = 3.106$, $r = .777$, $p < .01$.

[3] We analysed only the 160 problems that participants indicated had not been seen before (of a potential set of 208 problems: 26 participants $\times$ 8 problems). An independent rater scored over one-third of the responses, and interrater reliability was 99%.

[4] We analysed the 139 problems that participants indicated had not been seen before (of a potential set of 152 problems: 38 participants $\times$ 4 problems). An independent rater classified the hints for over one-third of the responses and agreement was more than 90%.

TABLE 1
Percentages of correct solutions for single- and multiple-step problems in the control and experimental conditions in Experiments 2, 3, and 4

		Control	*Experimental*
Experiment 2			
Single-step	Ping-pong	29	53
	Window cleaner	36	60
Multiple-step	Nine dot	0	0
	Six matches	0	7
Experiment 3			
Single-step	Ping-pong	79	79
	Window cleaner	79	100
Multiple-step	Nine dot	38	25
	Six matches	35	78
Experiment 4			
Single-step	Horse-and-rider	95	100
	Equation	73	100
Multiple-step	T-puzzle	5	53
	Necklace chains	7	47

Most participants provided ‘‘hints to a friend’’ that targeted the initial model errors. The most frequent hints for the ping-pong ball referred to the ball travelling on the vertical rather than the horizontal plane, 42%, or to gravity, 39%. The most frequent hint for the washing windows referred to the man’s height off the ground, 78%. The most frequent hint for the six matches referred to the shape of pyramids, 43%, either by giving a clue about Egypt, 29%, or by mentioning pyramids, 14%, or by referring to three dimensions, 31%. Perhaps unexpectedly, in the nine dot, most of the hints referred to the general shape of the solution, especially the triangular portion, 63%; the next most frequent hint was to extend lines beyond the square, 23%.

The experiment shows that erroneous initial models can be remedied by information that allows participants to encounter counterexamples, such as a man being close to the ground when he falls from a ladder, and counterexamples helped single-step but not multiple-step problems.

EXPERIMENT 3

The aim was to test the effects of physical props that enable participants to encounter counterexamples to subsequent erroneous possibilities (Tsai, 1987; Weisberg & Alba, 1981; Weller et al, 2011). For example, the provision of six matches and blue-tack enables participants to encounter counterexamples not only to the initial model error that triangles are formed in two dimensions, but also counterexamples to subsequent possibilities, e.g., constructing an upright triangle without a base.

Method

Materials and design. We used the same four problems as in the previous experiment. The control group received the standard wording, and the experimental group received in addition physical props. For the ping-pong the prop was a standard white ping-pong ball; for the window cleaner the prop was a model wooden ladder (595 mm long and 40 mm wide) and a small figure of a man (70 mm in height). For the six matches, the prop was six wooden ‘‘matches’’ (rods 8 mm in diameter and 250 mm long) and a packet of blue-tack. For the nine dot the prop was a corkboard with nine red pins in the matrix formation, extra pins in a cup, a ball of string, and scissors. Each participant was given the four problems in a different randomised order.

Participants and procedure. The participants were 40 students from Trinity College Dublin who participated voluntarily and in some cases in return for course credits. There were 23 women and 17 men, and their ages ranged from 17 to 49 years, with a mean age of 19. They were assigned at random to the control group who received the standard wording ($n = 20$) or the experimental group who received physical props ($n = 20$). Participants were tested individually. Participants in the control group were allowed pen-and-paper; participants in the physical condition were not, instead the written description was placed in front of them along with the props for each problem and the comment ‘‘here are some materials you might like to work with’’. Participants were allowed to work on each problem for 10 minutes. When they thought they had the correct answer, they announced it; if the solution was incorrect, they were given this feedback and attempted to solve it again.

Results and discussion

Participants given physical props solved more problems than those given the standard verisons, as shown by a Mann-Whitney U-test, for all problems overall $z = 2.116$, $r = .335$, $p < .05$,[5] as Table 1 shows. They did so for the single-step window cleaner, Fisher's exact test, $p < .05$, but not for the single-step ping-pong; and for the multiple-step six matches, $\chi^2 = 7.01$, $df = 1$, $p < .01$, but not the multiple-step nine dots, as Table 1 shows. Once again, participants solved more single-step than multiple-step problems, as shown by a Wilcoxon's test, $z = 3.231$, $r = .511$, $p < .01$.

The experiment showed that participants solved more problems when they were able to interact with physical props. The props enabled them to manipulate an external model, e.g., of the six matches. The props allowed them to correct their erroneous initial model by removing any assumption that the triangles were to be constructed in two dimensions only. Importantly, the props also allowed them to encounter a counterexample to any subsequent erroneous models, e.g., attempts to construct triangles in three dimensions but without a base triangle, for at least some of the insight problems. It is notable that the props did *not* help participants to solve some of the problems, perhaps because the pen-and-paper allowed in the standard condition was more familiar to use as an aid, e.g., in drawing dots and placing lines through them, compared to the props, e.g., the corkboard with pens.

EXPERIMENT 4

The aim was to examine further the finding that participants solve more problems when they encounter counterexamples through physical props, by extending it to a new set of insight "move" problems (see the Appendix) and by ensuring that the physical props *limited* the subsequent steps to helpful possibilities only. For example, for the "T-puzzle", there are four pieces to be arranged into the shape of a capital letter T, and participants were given a template into which to insert the broken pieces (e.g., Suzuki et al., 2001).

Method

Materials and design. We used four insight "move" problems (adapted from Knoblich et al., 2001; Suzuki et al., 2001; Weisberg, 1995; Weisberg & Alba, 1981; see the Appendix), comprising two multiple-step (T-puzzle, necklace chain), and two single-step problems (horse-and-rider, matchstick equation). Participants in the control group received physical props: the broken pieces of the T-shape, the horse-and-rider panels, a fully formed necklace chain, and the matches formed in the equation; participants in the experimental condition received constrained physical props that restricted the subsequent possibililities they could pursue: a cardboard cut-out of a T shape in which they were to insert the broken pieces, the horse-and-rider panels pinned through the centre so that they could be moved through rotation only, a fully formed necklace chain with three coloured links to encourage them to open and close only three links, three coloured matches forming the two mathematical symbols in the matchstick equation and the information that the crucial match was among them. Participants were given one multiple-step and one single-step problem each (at random either the two pattern problems, T-puzzle and horse-and-rider, or the two arithmetic-type problems, necklace and matchstick equation). The problems were presented in a different randomised order to each participant.

Participants and procedure. The participants were 65 members of the Trinity College Dublin School of Psychology participant panel. They volunteered to take part and were paid a fee of 8 euro. There were 46 women and 19 men and their ages ranged from 18 years to 73 years, with a mean age of 48 years. They were assigned at random to one of two groups, the control group who received physical props ($n = 35$) or the experimental group who received constrained physical props ($n = 30$). The procedure was the same as the previous experiment.

Results and discussion

Participants who received *constrained* physical props solved more problems than those who received the standard physical props, for all problems overall, $\chi^2 = 11.56$, $df = 1$, phi $= .422$, $p < .01$; for the T-puzzle, Fisher's exact test, $p < .005$, the necklace chains, Fisher's exact test, $p < .05$, the matchstick equation, Fisher's exact test, $p < .05$, but not for the horse-and-rider, perhaps because of a ceiling effect.

[5] One hundred and forty-two unseen problems (of a potential set of 160 problems) were analysed.

Once again participants solved more single-step problems than multiple-step problems as shown by Wilcoxon's tests for all problems overall, $z = 6.557$, $r = .813$, $p < .0005$; for the pattern-matching problems, i.e., horse-and-rider and T-puzzle, $z = 5.000$, $r = .870$, $p < .0005$, and the arithmetic-type problems, i.e., matchstick equation and necklace chains, $z = 4.243$, $r = .739$, $p < .0005$, as Table 1 shows.

The experiment showed that participants solved more problems when they had constrained physical props that limited the subsequent steps that could be taken from the initial model to helpful possibilities only.

GENERAL DISCUSSION

People solve insight problems by first constructing an initial model that corresponds to how the world would be, according to the problem description, e.g., they think of throwing a ping-pong ball. Their initial model incorporates assumptions about the normal way the event occurs, e.g., a ping-pong ball travels horizontally. Participants produced diagrams and explanations that corresponded to such initial model errors, and they failed to solve the problems when they did so, as Experiment 1 showed. Their hints to a friend also targeted the initial model error, as Experiment 2 showed.

However, people can modify their initial model when they encounter a counterexample, e.g., a *basketball*'s normal trajectory is vertical rather than horizontal. Participants solved more problems when the problems were reworded to make counterexamples available, as Experiment 2 showed. This "representational change" allows an impasse to be overcome, at least for single-step problems (e.g., Knoblich et al., 2001).

Single-step problems are easier than multiple-step problems, as all four experiments showed. Single-step problems can be solved once the initial model error is corrected, but multiple-step problems require further steps (e.g., Chronicle et al., 2004; Suzuki et al., 2001). Participants solved more problems when they had physical props that enabled them to correct their erroneous initial model, e.g., triangles are to be formed only in two dimensions, to a more accurate model, triangles can be formed in three dimensions; and also to encounter a counterexample to subsequent erroneous models, e.g., triangles are to be formed in three dimensions without a base, so that they chose a more accurate subsequent possibility, e.g., triangles are to be formed in three dimensions from a base of a two-dimensional triangle, as Experiment 3 showed. They solved more problems when they could interact with physical props that also limited the subsequent steps that could be taken to just the helpful steps, as Experiment 4 showed. This series of experiments corroborates the suggestion that the "aha" moment of insight that people experience when solving novel problems results from the cognitive change that occurs when they modify an initial model through the discovery of a counterexample.

REFERENCES

Ansburg, P. I., & Dominowski, R. L. (2000). Promoting insightful problem-solving. *Journal of Creative Behaviour*, *34*, 30–60.

Byrne, R. M. J. (2005). *The rational imagination: How people create alternatives to reality*. Cambridge, MA: MIT Press.

Chronicle, E. P., MacGregor, J. N., & Ormerod, T. C. (2004). What makes an insight problem? The role of heuristics, goal conception and solution re-coding in knowledge-lean problems. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *30*, 14–27.

Chu, Y., & MacGregor, J. N. (2011). Human performance on insight problem solving. *Journal of Problem Solving*, *3*, 119–150.

Duncker, K. (1945). On problem solving. *Psychological Monographs*, *58*(5) (Whole No. 270).

De Neys, W., Schaeken, W., & d'Ydewalle, G. (2005). Working memory and everyday conditional reasoning: Retrieval and inhibition of stored counterexamples. *Thinking and Reasoning*, *11*, 349–381.

Gilhooly, K., & Fioratou, E. (2009). Executive functions in insight versus non-insight problem solving: An individual differences approach. *Thinking and Reasoning*, *15*, 355–379.

Helie, S., & Sun, R. (2010). Incubation, insight, and creative problem solving: A unified theory and a connectionist model. *Psychological Review*, *117*, 994–1024.

Johnson-Laird, P. N., & Byrne, R. M. J. (2002). Conditionals: A theory of meaning, pragmatics, and inference. *Psychological Review*, *109*, 646–678.

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

Knoblich, G., Ohlsson, S., Haider, H., & Rheinus, 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 and Cognition*, *29*, 1000–1009.

Lockhart, R. S., Lamon, M., & Gick, M. L. (1988). Conceptual transfer in simple insight problems. *Memory and Cognition*, *16*, 36–44.
MacGregor, J. N., Ormerod, T. C., & Chronicle, E. P. (2001). Information processing and insight: A process model of performance on the nine-dot and related problems. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *27*, 176–201.
Murray, M. A., & Byrne, R. M. J. (2005). Attention and working memory in insight problem solving. In B. G. Bara, L. Barsalou, & M. Bucciarelli (Eds.), *Proceedings of the twenty-seventh annual conference of the Cognitive Science Society* (pp. 1571–1576). Mahwah, NJ: Lawrence Erlbaum Associates, Inc.
Ohlsson, S. (2011). *Deep learning: How the mind overrides experience*. Cambridge, UK: Cambridge University Press.
Scheerer, M. (1963). Problem solving. *Scientific American*, *208*, 118–128.
Sio, N. U., & Ormerod, T. C. (2009). Does incubation enhance problem solving? A meta-analytic review. *Psychological Bulletin*, *135*, 94–120.
Suzuki, H., Abe, K., Hiraki, K., & Miyazaki, M. (2001). Cue-readiness in insight problem-solving. In J. D. Moore & K. Stenning (Eds.), *Proceedings of the 23rd annual conference of the Cognitive Science Society* (pp. 1040–1045). Mahwah, NJ: Lawrence Erlbaum Associates, Inc.
Tsai, L. S. (1987). Influence of various tridimensional clues on solving a tetrahedron problem. *Perceptual and Motor Skills*, *65*, 672–674.
Weisberg, R. W. (1995). Prolegomena to theories of insight in problem-solving: A taxonomy of problems. In R. J. Sternberg & J. E. Davidson (Eds.), *The nature of insight* (pp. 157–196). Cambridge, MA: MIT Press.
Weisberg, R. W., & Alba, J. W. (1981). An examination of the alleged role of "fixation" in the solution of several "insight" problems. *Journal of Experimental Psychology: General*, *110*, 169–192.
Weller, A., Villejoubert, G., & Vallee-Tourangeau, F. (2011). Interactive insight problem solving. *Thinking and Reasoning*, *17*, 424–439.

APPENDIX

Experiments 1, 2, and 3

Here are the eight problems used in Experiment 1, their solutions, and conjectured initial model errors; the four problems used in Experiments 2 and 3 are marked with an asterisk.

Single-step problems

**Ping-pong problem.* Describe how to throw a ping-pong ball so that it will go a short distance, come to a dead stop, and then reverse itself. You are not allowed to bounce the ball against any object or attach anything to it.
Solution: Throw the ball straight up in the air.
Error: Represent the ball travelling along the horizontal and not the vertical plane.

**Window cleaner problem.* A man was washing windows on a high-rise office building when he slipped and fell off a 60-foot ladder onto the concrete pavement below. Despite having no safety equipment or anything to break his fall, he was not injured in any way. How was this possible?
Solution: The man was only at the bottom of the ladder when he fell.
Error: Represent the man at the top of the ladder when he fell and not lower down.

Camping trip problem. While out on a camping trip, Miss Jones woke one morning and felt something in the pocket of her shorts. It had a head and tail but no legs. When Miss Jones got up, she could feel it move inside her pocket. Miss Jones, however, was not concerned and allowed the thing to remain in her pocket all day. What was it?
Solution: The "thing" is a coin.
Error: Represent the "thing" as animate and not inanimate.

Woods walk problem. Two men get lost while walking in the woods. One starts walking northwards, whilst the other heads south. They bump into each other a quarter of an hour later. Presuming that neither man changed direction during that time, explain how such a situation could have arisen.
Solution: The men were not walking together at the outset.
Error: Represent the men walking together at the outset.

Multiple-step problems

**Nine-dot problem (alien version).* In a computer game, there are nine aliens to be captured. The aliens are standing in a formation of three rows of three across with equal spaces between them, in other words a 3×3 matrix. To capture an alien, you must walk through him. Show how you can trace a route that goes through all nine aliens using just four connected straight paths. Do not retrace any part of your route (or raise your pen

from the page) once you have started on the first path. You may use simple dots to represent the aliens.
Solution: The solution requires the extension of some lines beyond the matrix formed by the dots. The most common solution is roughly arrow-shaped, although other valid solutions with different shapes exist.
Error: Represent the lines within the 3×3 matrix and not outside it.

**Six-matches problem.* A teacher challenges her pupils to form four equilateral triangles using just six matches of equal length. They are also told that each complete match must form one complete side of a triangle. How can it be done?
Solution: Three-dimensional figure with one triangle forming the base, and the other triangles forming the three sides.
Error: Represent the triangles in two dimensions, and not three dimensions.

Candle problem. In a room there is a table pressed against a wall. On the table are a candle, a box of tacks, and a packet of matches. Your task is to attach the candle to the wall above the table in such a way that the wax from the lighted candle will not drip onto the table or the floor. Sketch or describe how this can be done.
Solution: Use the box containing the tacks as a candle-holder. Use some of the tacks to attach it to the wall, light candle and melt some of the wax into the box, secure candle in box.
Error: Represent the box as a container for the tacks, not as a platform for candle.

Radiation problem. A doctor has a patient with a malignant tumour in her stomach. It is impossible to operate but, unless the tumour is destroyed, the patient will die. There is a kind of ray that, if directed at the tumour at a sufficiently high intensity, will destroy the tumour. Unfortunately, at this intensity the healthy tissue through which the ray passes will also be destroyed. At lower intensities, the ray is harmless to the healthy tissue but will not affect the tumour either. What type of procedure might be used to destroy the tumour with such rays and at the same time avoid destroying the healthy tissue?
Solution: Multiple low intensity rays converge at the tumour to form a laser of sufficient intensity at that point only.
Error: Represent one high intensity ray at a time getting to tumour, not low intensity rays.

Experiment 4

Here are the four problems used in Experiment 4, their solutions, and the constrained physical props provided.

T-puzzle. The T-puzzle comprises four segments that can be used to form a "T" shape (see Suzuki et al., 2001 for an illustration). Participants were asked to arrange the four pieces of a puzzle to form a T-shape. The finished "T" measured 160 mm wide and 180 mm high. Each bar of the "T" was 60 mm wide. The puzzle pieces were cut from rigid foam 5 mm thick. The upper side was painted silver to differentiate it from the black underneath.
Solution: The solution is a complete T-shape.
Constrained physical prop: An actual-size cardboard cut-out of the finished "T" that could be used as a template was provided.

Horse-and-rider. The horse-and-rider puzzles used two light cardboard panels, one of which had an illustration of two horses and the other had two riders sitting on part of a horse (see Weisberg & Alba, 1981, for an illustration). Participants were asked to place the horses panel on the riders panel in such a way that both riders are properly astride their horses. The horses panel was 127 mm square and the riders panel measured 127 mm by 32 mm.
Solution: Rotate one of the panels through 180 degrees.
Constrained physical prop: The panels were pinned through the centre so that the riders panel only could be rotated into position over the horses panel.

Necklace chain. The necklace chain problem comprises 12 paper chain-links, divided into four sections of three links, as well as 15 cents worth of 1 and 2 cent coins. The task was to join all the links together to form a chain like a necklace. It costs 2 cents to open a link and 3 cents to close a link and you have a maximum of 15 cents to spend. The paper links can be opened and closed via blue-tack.

Solution: Open all the links in one section (2 cent $\times 3 = 6$ cent). Use the open links as joining links between the existing sections and close each one (3 cent $\times 3 = 9$ cent).

Constrained physical prop: A completed necklace-like chain was constructed with each of the three sections consisting of three white links joined with a blue link. A note informed participants that they could only afford to open and close three links. All the links given to participants to work with were white.

Matchstick equation. The matchstick equation used matches to form an unbalanced equation reading VII = V − II. Participants were asked to balance the equation by moving just one match and without taking any away.
Solution: VII − V = II.
Constrained physical prop: The three matches making up the equals and minus signs were coloured pink and participants were instructed that the key match was one of these matches.