

Rebus puzzles as insight problems

JAMES N. MACGREGOR AND J. BARTON CUNNINGHAM
University of Victoria, Victoria, British Columbia, Canada

Research on insight—the phenomenon of suddenly solving an apparently intransigent problem—has been hampered because stimulus problems have been few, ad hoc, heterogeneous, and difficult to solve. Responding to the need for a larger pool of problems of a similar type and of varying level of difficulty, we report an experiment testing the validity of rebuses as insight problems. A rebus combines verbal and visual clues to a common phrase, such as *PAINS* (“growing pains”). Solving a rebus requires breaking implicit assumptions of normal reading, similar to the restructuring required in insight. We hypothesized that, the more implicit assumptions are involved, the more difficult the solution. The results of a two-part experiment supported the hypothesis, with participants solving more problems involving one assumption than they did problems involving two or more. Also, rebus performance correlated significantly with self-rated insight and with scores on remote associates, but not with general verbal ability. The findings suggest that rebus puzzles may be a useful source of theoretically grounded insight problems.

A central question in cognitive psychology concerns the nature of *insight* (Mayer, 1996). The “insight experience,” of being perplexed by a problem whose answer later becomes suddenly clear, may be common, but agreeing on a psychological definition remains problematic (Weisberg, 1996). For some, the phenomenology of sudden illumination is the defining characteristic (Metcalfe & Wiebe, 1987). Others have focused on psychological process, such as restructuring or representational change following impasse (Knoblich, Ohlsson, Haider, & Rheinert, 1999; Ohlsson, 1992), satisfying progress toward solution (MacGregor, Ormerod, & Chronicle, 2001), or unblocking (Duncker, 1945; Smith, 1996).

Similar difficulties arise in defining the properties of *insight problems* (Chronicle, MacGregor, & Ormerod, 2004; Weisberg, 1996), and the status of some classic problems has been questioned. The textbook nine-dot problem, for example, is supposed to require one to “think outside the box” for the solution (Newell & Simon, 1972). However, a large body of research now casts doubt on this standard interpretation (Kershaw & Ohlsson, 2004; Lung & Dominowski, 1985; MacGregor et al., 2001; Weisberg & Alba, 1981).

Even if researchers agreed on what problems are likely to invoke insight, the field would suffer from a relative dearth of problems. Until recently, stimulus materials were limited to a small, heterogeneous collection of spatial puzzles and some dozens of verbal riddles (e.g., “A man was washing windows on a high-rise building when he fell from the 40-foot ladder to the concrete path below. Amazingly, he was unhurt. Why? [Answer] He slipped from the bottom rung!”). In addition to validity, the available problems raise other measurement issues: Sensitivity is one, because some problems are so difficult that virtually nobody solves them (MacGregor et al., 2001). In other cases, difficulty level

is unknown, except for a few published empirical norms (Ansburg & Dominowski, 2000). Assessing reliability may also be an issue. The heterogeneous nature of the problems rules out equivalent forms, whereas the relative memorability of insight solutions (Dominowski & Dallob, 1996; Knoblich et al., 1999; Ormerod, MacGregor, & Chronicle, 2002) rules out test/retest reliability. Fortunately, however, the situation may be improving, thanks to new sources of potential insight problems.

“Matchstick arithmetic,” developed by Knoblich and his associates (1999), represents one new source. The task consists of matchsticks arranged as an incorrect equation in Roman numerals, to be corrected by moving one matchstick (for example, $IV = III - I$; answer, $IV - III = I$). Matchstick problems address some of the measurement issues. First, theory and results suggest that they represent valid insight problems. Second, there is a theoretical basis for establishing ranked categories of problem difficulty. This helps deal with issues of measurement sensitivity and equivalence. Third, there is the practical advantage of a large pool of homogeneous instances (though a practical limitation may be that they require knowledge of Roman numerals).

A second relatively new source of potential insight problems comes from the Remote Associates Test (RAT), developed as a test of creativity (Mednick, 1962). The task is, given three words, to find a fourth that connects them; for example, *lick, mine, shaker*: answer, *salt*. This type of problem has been widely used to study insight and related phenomena (Ansburg, 2000; Bowers, Regehr, Balthazard, & Parker, 1990; Smith & Blankenship, 1991; Wiley, 1998). RAT performance has been reported to correlate with performance on standard insight problems (Ansburg, 2000; Dallob & Dominowski, 1993; Schooler & Melcher, 1995) and with ratings of insight (Bowden & Jung-Beeman, 2003a).

In addition, it has been argued that RAT problems involve the same solution processes as “classic” insight problems (Bowden & Beeman, 1998). However, although there are wide variations in the difficulty of RAT problems, and empirical performance norms exist (Bowden & Jung-Beeman, 2003b), there is no theoretical basis for determining level of difficulty. In this respect, they are inferior to matchstick arithmetic problems as stimuli for studying insight.

We proposed another pool of potential insight problems, in the form of the rebus (Cunningham & MacGregor, 2006). A rebus combines verbal and visual clues to a familiar phrase (e.g., *you just me* = “just between you and me”). Rebuses (or *wordies*) have been used to study fixation and incubation (Smith & Blankenship, 1989) but otherwise have been overlooked as candidate insight problems. Nevertheless, they may share properties attributed to insight problems. For example, Bowden and Beeman (1998) cite three properties: (1) Insight problems misdirect solution efforts away from a nondominant meaning; (2) solvers may be unable to report the processing that leads to solution; (3) solving is often accompanied by an “aha!” experience. Many rebus problems also exhibit these characteristics.

Insight is often considered to require “restructuring,” a radical shift in how problem elements are cognitively or perceptually represented (Knoblich et al., 1999). Past experience may be involved, by imposing constraints on search processes that must be “relaxed” for solution (Ohlsson, 1992) or by suggesting a compelling but wrong approach that must subsequently be changed (Ormerod et al., 2002). Rebus solutions typically exhibit these characteristics. From past experience with reading, we tacitly assume, for example, that word order has a syntactic significance but not a semantic one, whereas in the case of *you just me*, the solution ignores the grammar of word order and requires a meaningful interpretation. This is a common principle in rebus puzzles, where spatial relationships of problem components must be given a verbal interpretation, such as “over,” “after,” “beside,” and so on. Past experience also indicates that meaning is independent of print characteristics, such as font size, font style, and so on, whereas for some rebus solutions these are the characteristics that need to be attended to, as in *PUNISHMENT* (“capital punishment”) and *a front* (“a bold front”).

The foregoing considerations led to the hypothesis that rebus solutions may involve insight. Also, because a number of principles can be used to encrypt the meaning of a rebus, a successful solution may require more than one dimension of restructuring. This leads to the possibility that the difficulty of rebuses may depend, in part, on the number of overlearned constraints to be relaxed to decode the meaning. For example, to solve *SOMething* requires attention to a spatial relation and to a font property (“the start of something big”), whereas the example

$$\frac{\text{exit}}{\text{leg}}$$

involves three different encrypting devices—synonym substitution (exit = go out), spatial relation (on), and superordinate category recognition (leg = a limb). If in-

sight involves changing implicit assumptions, as is widely supposed, then the number of assumptions that must be changed may indicate a rebus problem’s difficulty. We refer to this as the “restructuring index.” We report an experiment assessing the validity of rebuses as insight problems by testing restructuring index as a measure of instance difficulty.

Restructuring Index

We obtained a pool of 100 rebus problems from a variety of Internet sources. One author examined each problem and its solution, to identify which encrypting devices were involved that countermanded the normal assumptions of reading. He identified a total of 16 types. The most frequent required the spatial relations of words to be interpreted, as in “under,” “through,” “beside,” and so on, which occurred in 61% of cases. Other principles included: trend (growing, decreasing, etc., as in *PAINS* = “growing pains”); counting (e.g., *poPPd* = “two peas in a pod”); and interpreting letters as words

(e.g., $\frac{\text{iiii}}{\text{ooo}}$ = “circles under the eyes”).

Individual rebuses included a minimum of one to a maximum of four principles. The same process was applied to a second set of 152 rebuses, resulting in two additional principles. As a reliability check, a second judge used the definitions of the 18 different encoding principles to count the number involved in each of 30 rebuses randomly selected from the second set. We compared these totals with those identified in the original classification. The number of agreements across problems divided by the number of agreements plus disagreements resulted in a level of agreement of 80%.

Other Measures

To estimate convergent and discriminant validity of rebuses as insight problems, the experiment measured several additional variables. For convergent validity, we included self-ratings of insight and performance on RAT items. For discriminant validity, we included two verbal analogy tests. If rebuses measure insight beyond general verbal fluency, we expected rebus performance to correlate more highly with self-ratings and RAT performance than with verbal analogy scores.

Previous research has used forms of insight ratings for specific problems, treating insight as a characteristic of a problem’s solution process (Bowden & Jung-Beeman, 2003a; Metcalfe & Wiebe, 1987). This is an appropriate approach, but practical considerations made us consider an alternative, to treat insightfulness as an individual characteristic. Previous results have indicated that there may be reliable individual differences in the ability to solve insight problems. For example, Jacobs and Dominowski (1981) reported moderately stable individual differences in performance across seven different object-use insight problems, and average performance on the seven problems correlated significantly with scores on the Gestalt Transformations Test. As mentioned above, Ansburg (2000) reported

a significant correlation between performance on verbal insight problems and RAT scores. Gilhooly and Murphy (2005) compared performance on 24 presumed insight problems with a variety of individual difference measures encompassing fluid and crystallized intelligence, working memory capacity, and ideational fluency and flexibility. Composite scores on verbal insight problems were significantly predicted by vocabulary and figural fluency scores, whereas spatial insight scores were significantly predicted by figural fluency scores only. Murray and Byrne (2005) compared performance on eight verbal insight problems with several measures of attention switching, selective attention, and working memory, and found significant positive relationships of insight scores with both attention switching and working memory capacity. Ash and Wiley (2006) found that average success rate across a range of insight problems correlated significantly with working memory span. All of these studies found that properties of insight can be measured within individual subjects.

Given the possibility of reliable individual differences in insight ability, we considered that people may be able to provide self-ratings of their insightfulness, both in terms of how they normally feel (potential "trait insight") and how they feel at the present time (potential "state insight"). The procedure we employed for self-ratings involved an applet showing a lightbulb whose color saturation could be varied by clicking and dragging a button (see Figure 1). We used this because the lightbulb metaphor of insight as illumination is pervasive, and this visual analogue might help define for our participants the construct we wanted to measure.

The experiment also used 20 RAT items and 20 verbal analogies. RAT items were randomly selected from 68 that are available from an online source (Kihlstrom, 2005). Verbal analogies were randomly selected from 28 used by Ansborg (2000), such as "Water is to steam just as fire is to ___?: A. chimney; B. wood; C. smoke; D. arson." Ten were presented in a multiple-choice format, with the answer being selected from four or five options. The remaining 10 were presented in an open-ended format, without the multiple-choice answer options. The experiment was conducted as part of a larger procedure requiring two 1-h sessions.

METHOD

Participants

A total of 40 volunteers were recruited at the University of Victoria. Participants were paid \$25 to participate in two 1-h sessions.

Materials

We obtained a pool of 21 rebuses from the first set of 100 by randomly sampling 7 each from the 73 with a restructuring index of one, and the 20 with an index of two, and using all 7 that had an index of three or four. These were printed on separate, letter-sized sheets. The order of the sheets was randomized, and they were then stapled into a booklet, preceded by a page of instructions and a page providing a practice example and answer. Similarly, we obtained a second pool of 24 from the second set of 152 rebuses by randomly selecting 8 each from those with a restructuring index of one, two, and greater than two.

The 20 RAT items were printed in a three-page booklet. The first page contained instructions and an example (with answer). The second and third pages contained 10 test items each. The multiple-choice

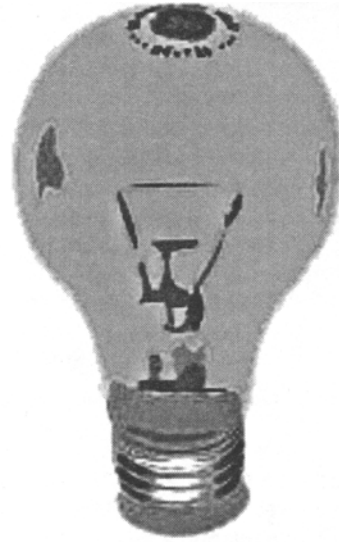


Figure 1. Lightbulb and scale to obtain insight ratings.

verbal analogies were printed in a four-page booklet. The first page contained instructions and two examples, with answers, and the remaining pages contained the analogies and multiple-choice answers. The open-ended analogies were printed in a two-page booklet, the first page providing instructions and the second the analogies.

Design and Procedure

Participants were tested individually in two sessions, separated by approximately 1 week. Session 1 started by describing insight in the context of invention and discovery and then suggested that, in a smaller way, most people have had the insight experience. "For example, most people have had the experience of trying to . . . solve a problem that stumps them for a while. Then, often out of the blue, an answer comes to them. Aha!" They were then asked to rate how insightful they usually were by adjusting the brightness of the lightbulb. Immediately following this, they provided a similar rating, of how they felt at the present.

Following the two ratings, participants performed a number of tasks, occupying approximately 30 min, and then they worked on the rebus problems for a maximum of 10 min. At the end of the procedure, the participants provided a final insight rating of how they felt at the moment.

Session 2 followed a similar format to that of Session 1. Participants began by providing self-ratings of their usual insight level and then of their level of insight at that time. They then completed a task occupying approximately 10 min, before working on the 24 rebus problems for a maximum of 12 min. This was followed by the multiple-choice verbal analogies (5 min), the open-ended verbal analogies (5 min), and the RAT (10 min). At the end, they provided an insight rating of how they felt at the moment.

RESULTS

Rebus Performance and Restructuring Level

We scored rebus answers as correct (1) or incorrect (0). The means (and standard deviations) of number correct for Restructuring Levels 1, 2, and 3 were 4.88 (1.09), 4.18 (1.57), and 3.00 (1.59), for Session 1, out of a possible 7. For Session 2, the corresponding results were 5.45 (1.01),

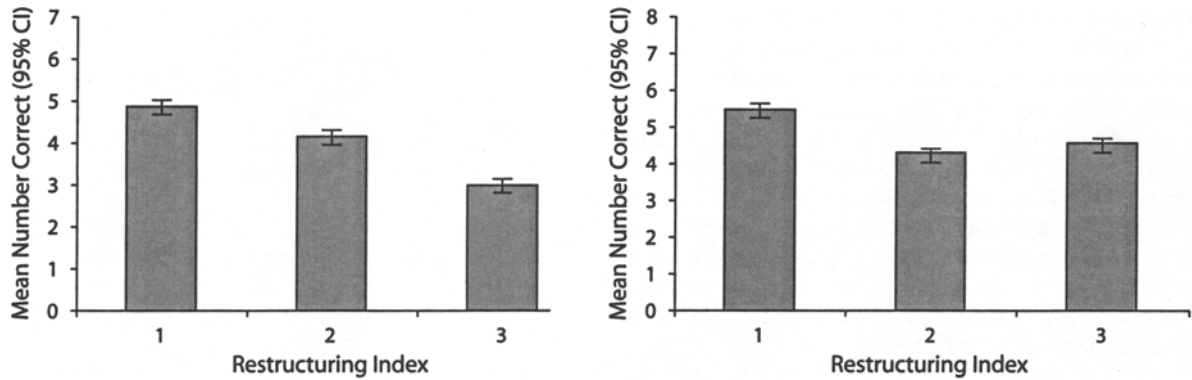


Figure 2. Mean number of rebus puzzles solved in Session 1 (left) and Session 2 (right) by restructuring index.

4.28 (1.55), and 4.55 (1.50), out of 8. The overall mean proportion correct was .57 (.17) for Session 1 and .59 (.13) for Session 2. The correlation across participants between the proportion correct in Session 1 and in Session 2 was $r = .43, p < .01$. Figure 2 shows the mean number correct by restructuring level for Sessions 1 and 2. For Session 1, mean performance levels decreased across increasing levels of restructuring. For Session 2, performance decreased between Restructuring Levels 1 and 2, but not between Levels 2 and 3.

We conducted within-subjects ANOVAs on rebus performance with index of restructuring as the independent variable and number correct as dependent variable. For Session 1, restructuring level had a significant effect on rebus solution [$F(2,78) = 35.52, p < .001, \eta^2 = .48$]. Performance decreased linearly across increasing levels of restructuring. Post hoc comparisons indicated significant differences between all three conditions ($p < .01$). For Session 2, there was again a significant overall effect of restructuring level [$F(2,78) = 13.22, p < .001, \eta^2 = .25$]. In this case, there was a significant quadratic trend, with performance decreasing from Level 1 to Level 2, and then increasing slightly at the highest level of restructuring. Post hoc comparisons indicated that performance at the first level of restructuring was significantly better than at the other two levels ($p < .01$). The difference between Level 2 and Level 3 was not significant ($p = .24$).

Secondary Measures

Insight self-ratings were recorded on a scale from 0 to 100. Answers on the verbal analogy tests and RAT were scored 1 if correct and 0 if incorrect, and total scores were expressed as the proportion correct. For the open-ended analogies, one of the authors and an independent judge scored the answers, resulting in a 92% level of agreement. For the multiple-choice analogies, scores were high, with little variability. The mean was .94 with a standard deviation of .10, and the majority of participants had perfect scores. Because of this truncated range, the variable was excluded from further analysis. Our principle interest was in the correlations of rebus performance with RAT performance, and also with the noninsight verbal analogy scores, shown in Table 1.

Rebus scores from Session 1 and Session 2 both correlated significantly with RAT scores from Session 2 ($r = .44$ and $.38$, respectively), whereas neither Session 1 nor Session 2 rebus scores correlated significantly with the Session 2 verbal analogy score ($r = .25$ and $.22$, respectively). However, whereas the correlations between both sets of rebus scores and RAT scores were greater than the correlations between rebus and analogy scores, the differences were not significantly greater.

Finally, we compared rebus scores with self-ratings of insight. The two ratings of normal level of insight (from the first and second sessions) were significantly correlated ($r = .82$), and we averaged the two to provide a single measure of putative "trait insight." The correlations of rebus scores with the self-ratings are shown in Table 2. Correlations of analogy scores and self-ratings are also included, for comparison.

With respect to the rebus scores, four of the six correlations with self-ratings were significant ($p < .05$). These were the correlations of Session 1 rebus scores with trait insight rating and state insight at the start of the session, and the correlations of Session 2 rebus scores with both state ratings. The analogy scores did not correlate significantly with any of the three ratings. All six correlations of

Table 1
Correlations of Rebus Scores With Other Performance Scores

Test	Rebus 2	RAT	Analogy
Rebus Session 1	.43*	.44*	.25
Rebus Session 2		.38*	.22

*Correlations $\geq .31$ significant at $p \leq .05$.

Table 2
Correlations of Insight Self-Ratings With Rebus and Analogy Scores

	Insight Self-Ratings		
	"Trait" Rating	"State" Rating at Start	"State" Rating at End
Rebus Session 1	.31*	.15	.35*
Rebus Session 2	.24	.40*	.36*
Analogy	.23	.13	.03

*Correlations $\geq .31$ significant at $p \leq .05$.

self-ratings with rebus scores were higher than the corresponding correlations with analogy scores, but again, the differences did not reach significance.

DISCUSSION

We conducted a preliminary investigation of rebuses as candidate insight problems. We developed and tested a theoretical measure of rebus difficulty based on the number of implicit assumptions in reading that would have to be relaxed to solve a rebus, and we compared the difficulty of rebus problems against this theoretical index. We also compared rebus performance with two other candidate measures of insight—self-rated insight level and RAT performance—as a preliminary assessment of convergent validity. We addressed discriminant validity by comparing rebus performance with verbal analogy scores.

The results provided initial support for the status of rebuses as insight problems in several respects. For Session 1, the solution rate for rebuses decreased systematically across three levels of increasing restructuring. The results for Session 2 were less clear, showing a decrease in performance across the first and second levels of restructuring, but not across the second and third. However, a partial explanation is that performance at the third level in Session 2 may have been artificially inflated because of the presence of two items highly similar to items seen in Session 1. Specifically, these were: *RIGHT your iiiiii* (“right before your eyes,” Session 1) and

EZ
iii (“easy on the eyes,” Session 2).

The former had a success rate of 20%; the latter, of 55%. Similarly, the Session 1 rebus,

posting
cccccc (“overseas posting”),

may have partially cued a similar Session 2 item,

sailing
cccccc (“sailing overseas”).

The success rates for these were 55% and 95%, respectively. However, omitting these two items from the Session 2 scores does not entirely reconcile the differences between the Session 1 and Session 2 results. For Session 1, the percentages correct were 70%, 60%, and 43% for Restructuring Levels 1, 2, and 3, respectively, whereas the corresponding recalculated proportions for Session 2 were 68%, 53%, and 51%. As can be seen, performance rates in Session 2 were lower for Level 2 and higher for Level 3 than in Session 1.

The pattern of correlations of rebus performance with other measures also provided partial, but not completely consistent, support for the rebus-as-insight hypothesis. Rebus scores from both sessions correlated significantly with RAT scores obtained in the second session, and did not correlate significantly with verbal analogy scores. The correlations of the two rebus scores with the RAT were essentially the same as their correlation with each other, indicating that the correlation with RAT was as high as rebus

reliability would allow. Although the correlations between rebus and RAT scores were not significantly greater than the correlation between rebus and verbal analogy scores, the direction of the differences provides some indication that the correlations between rebus and the RAT may not have arisen purely through their shared variance with verbal fluency in general.

The pattern of correlations of rebus performance with self-rated insight also provided some preliminary support for the rebus-as-insight proposal. If rebus scores and self-ratings are valid indicators of insight, then we would expect relatively high correlations between rebus performance and state insight ratings within each session, and lower correlations of trait insight ratings with rebus performance in both sessions. The observed correlations partially met these expectations. Consistent with the expectation, the correlation of rebus performance with the state insight ratings at the start and end of Session 2 were both significant ($r = .40$ and $.36$, respectively), whereas the correlation of rebus scores from Session 1 with trait insight was significant, but lower ($r = .31$). Inconsistent with expectations, the rebus scores in Session 1 failed to correlate significantly with state insight measured at the start of that session. A possible explanation is that approximately 30 min had passed between the rating and the rebus test, during which time state insight level may have changed. In contrast, in Session 2, where the correlation was significant, less time elapsed between the state rating and rebus test (approximately 12 min). In Session 2, the self-rating that failed to correlate with rebus score was that of trait insight level. However, anecdotal accounts suggest that insight tends to be serendipitous, and trait insight may be a sporadic predictor of specific performance. The pattern of self-ratings here may reflect this fluctuating nature of insight capacity, where the correlation between the two trait ratings, taken 1 week apart, was $.82$, whereas the correlation between the trait and the first state ratings, taken a few minutes apart, was only $.57$ in Session 1 and $.63$ in Session 2. We note that “inconsistency,” defined as transient fluctuations in states, has also been observed in other cognitive domains (Williams, Hultsch, Strauss, Hunter, & Tannock, 2005). Finally, the state ratings taken at the end of the session would presumably have been influenced by performance during a session. In Session 2, the analogies test was conducted after the rebus task and was therefore closer in time to the final rating. Nevertheless, the rating was more closely related to performance on rebuses ($r = .36$) than on analogies ($r = .03$). This suggests that participants may have been more likely to link insight and the subjective experience of solving rebuses than of solving analogies, perhaps because rebus solution was more likely to evoke the “Aha!” reaction. This will be investigated in future research.

The preliminary evidence reported here suggests that rebus puzzles may be a promising source of insight problems that offer a number of potential advantages over some of the ad hoc puzzles and riddles that have been used as stimuli in the past—there is an abundant supply, the problems are of a similar type, and they appear to vary in difficulty on a principled basis.

AUTHOR NOTE

The authors thank Michelle Arnold for sharing her collection of rebus puzzles. The research was partially supported by a research grant from the Natural Science and Engineering Research Council of Canada to the first author and a research grant from the Social Sciences and Humanities Research Council of Canada to both authors. Correspondence concerning this article should be addressed to J. N. MacGregor, School of Public Administration, University of Victoria, Victoria, BC, Canada, V8W 2Y2 (e-mail: jmacgreg@uvic.ca).

REFERENCES

- ANSBURG, P. I. (2000). Individual differences in problem solving via insight. *Current Psychology*, *19*, 143-146.
- ANSBURG, P. I., & DOMINOWSKI, R. L. (2000). Promoting insightful problem solving. *Journal of Creative Behavior*, *34*, 30-60.
- ASH, I. K., & WILEY, J. (2006). The nature of restructuring in insight: An individual-differences approach. *Psychonomic Bulletin & Review*, *13*, 66-73.
- BOWDEN, E. M., & BEEMAN, M. J. (1998). Getting the right idea: Semantic activation in the right hemisphere may help solve insight problems. *Psychological Science*, *9*, 435-440.
- BOWDEN, E. M., & JUNG-BEEMAN, M. (2003a). Aha! Insight experience correlates with solution activation in the right hemisphere. *Psychonomic Bulletin & Review*, *10*, 730-737.
- BOWDEN, E. M., & JUNG-BEEMAN, M. (2003b). Normative data for 144 compound remote associate problems. *Behavior Research Methods, Instruments, & Computers*, *35*, 634-639.
- BOWERS, K. S., REGEHR, G., BALTHAZARD, C. [G.], & PARKER, K. (1990). Intuition in the context of discovery. *Cognitive Psychology*, *22*, 72-110.
- CHRONICLE, E. P., MACGREGOR, J. N., & ORMEROD, T. C. (2004). What makes an insight problem? The roles of heuristics, goal conception, and solution recoding in knowledge-lean problems. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, *30*, 14-27.
- CUNNINGHAM, J. B., & MACGREGOR, J. N. (2006, September). An experimental test of rebus puzzles as insight problems. Paper presented at the Joint New Zealand Psychological Society/Australian Psychological Society Conference, Auckland, New Zealand.
- DALLOB, P. I., & DOMINOWSKI, R. L. (1993). *Erroneous solutions to verbal insight problems: Effects of highlighting critical material*. Paper presented at the annual meeting of the Western Psychological Association, Phoenix, AZ.
- DOMINOWSKI, R. L., & DALLOB, P. (1996). Insight and problem solving. In R. J. Sternberg & J. E. Davidson (Eds.), *The nature of insight* (pp. 33-62). Cambridge, MA: MIT Press.
- DUNCKER, K. (1945). On problem solving. *Psychological Monographs*, *58*(5, Whole No. 270).
- GILHOOLY, K. J., & MURPHY, P. (2005). Differentiating insight from non-insight problems. *Thinking & Reasoning*, *11*, 279-302.
- JACOBS, M. K., & DOMINOWSKI, R. L. (1981). Learning to solve insight problems. *Bulletin of the Psychonomic Society*, *17*, 171-174.
- 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, & Cognition*, *30*, 3-13.
- KIHLSTROM, J. F. (2005). Remote Associates Test. Online document retrieved on March 18, 2005, from socrates.berkeley.edu/~kihlstrm/RATest.htm.
- KNOBlich, G., OHLSSON, S., HAIDER, H., & RHENIUS, D. (1999). Constraint relaxation and chunk decomposition in insight problem solving. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, *25*, 1534-1555.
- LUNG, C.-T., & DOMINOWSKI, R. L. (1985). Effects of strategy instructions and practice on nine-dot problem solving. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, *11*, 804-811.
- MACGREGOR, J. N., ORMEROD, T. C., & CHRONICLE, E. P. (2001). Information processing and insight: A process model of performance on the nine-dot and related problems. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, *27*, 176-201.
- MAYER, R. E. (1996). The search for insight: Grappling with Gestalt psychology's unanswered questions. In R. J. Sternberg & J. E. Davidson (Eds.), *The nature of insight* (pp. 3-32). Cambridge, MA: MIT Press.
- MEDNICK, S. A. (1962). The associative basis of the creative process. *Psychological Review*, *69*, 220-232.
- METCALFE, J., & WIEBE, D. (1987). Intuition in insight and noninsight problem solving. *Memory & Cognition*, *15*, 238-246.
- 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 XXVII Annual Conference of the Cognitive Science Society* (pp. 1571-1575). Mahwah, NJ: Erlbaum.
- NEWELL, A., & SIMON, H. A. (1972). *Human problem solving*. Englewood Cliffs, NJ: Prentice-Hall.
- OHLSSON, S. (1992). Information-processing explanations of insight and related phenomena. In M. T. Keane & K. J. Gilhooly (Eds.), *Advances in the psychology of thinking* (Vol. 1, pp. 1-44). London: Harvester Wheatsheaf.
- ORMEROD, T. C., MACGREGOR, J. N., & CHRONICLE, E. P. (2002). Dynamics and constraints in insight problem solving. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, *28*, 791-799.
- SCHOOLER, J. W., & MELCHER, J. (1995). The ineffability of insight. In S. M. Smith, T. B. Ward, & R. A. Finke (Eds.), *The creative cognition approach* (pp. 97-134). Cambridge, MA: MIT Press.
- SMITH, S. M. (1996). Getting into and out of mental ruts: A theory of fixation, incubation and insight. In R. J. Sternberg & J. E. Davidson (Eds.), *The nature of insight* (pp. 229-251). Cambridge, MA: MIT Press.
- SMITH, S. M., & BLANKENSHIP, S. E. (1989). Incubation effects. *Bulletin of the Psychonomic Society*, *27*, 311-314.
- SMITH, S. M., & BLANKENSHIP, S. E. (1991). Incubation and the persistence of fixation in problem solving. *American Journal of Psychology*, *104*, 61-87.
- WEISBERG, R. W. (1996). Prolegomena to theories of insight in problem solving: A taxonomy of problems. In R. J. Sternberg & J. E. Davidson (Eds.), *The nature of insight* (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.
- WILEY, J. (1998). Expertise as mental set: The effects of domain knowledge in creative problem solving. *Memory & Cognition*, *26*, 716-730.
- WILLIAMS, B. R., HULTSCH, D. F., STRAUSS, E. H., HUNTER, M. A., & TANNOCK, R. (2005). Inconsistency in reaction time across the life span. *Neuropsychology*, *19*, 88-96.

(Manuscript received January 19, 2007;
revision accepted for publication May 22, 2007.)