Constraints and their relaxation in the processes of insight

Hiroaki Suzuki
Department of Education
Aoyama Gakuin University
4–4–25 Shibuya, Shibuya-Ku
Tokyo, 150, JAPAN
`susan@ri.aoyama.ac.jp`

Kazuo Hiraki
ETL, MITI / Presto,JST
1-1-4 Umezono Tsukuba-shi
Ibaraki, 305 Japan
`khiraki@etl.go.jp`

TR-97-13

Abstract

Research on insight has accumulated experimental evidence on its cognitive processes. However, there is little agreement on what problem–solvers learn from their initial failure and at what point an insight actually takes place. To explore these issues, we first propose a general framework that involves three constraints, *object–level*, *relational*, and *goal*. The object–level and relational constraints represent people's natural preference of how objects and relations in a given problem are represented. The goal constraint evaluates a degree of match of the current state to the goal, and forces problem–solvers to select specific combinations of the representations of objects and relations. In the processes of insight, these constraints operate simultaneously and are gradually relaxed by repeated impasses. Using a kind of geometric puzzle problem, we empirically tested hypotheses derived from the framework. Experimental results revealed that the initial persistence in a wrong approach could be explained by the object–level and goal constraints, and that subjects could reach an insight by relaxing the object–level constraints as well as having the goal constraint easy to operate.

1 Introduction

Insight, one of the most mysterious cognitive activities, has been getting closer and closer to the scope of cognitive science. The processes of insight have been explored, using various kinds of approaches, such as experimental, protocol analysis, historical, etc (Sternberg & Davidson, 1995).

However, there still exists a mystery in insight. It comes to mind at a certain point during problem–solving processes. But, at what point? This problem (hereinafter, *at-what-point problem*) becomes more mysterious, considering that crucial information for insight is often available a lot before insight actually takes place. However, these neglected cues unexpectedly become illuminative. Ohlsson (1992) pointed out a related problem, by stating "why does a person encounter an impasse on a problem which he or she is competent to solve?"(p.6).

Although Ohlsson attributed the cause of impasse mainly to failure of operator retrieval, a counter-example is found in the Kaplan and Simon's study (1990). They empirically studied search constraints that lead problem–solvers to shift to another problem space, using the multilated checkerboard (MC) puzzle. It is well known that realizing parity of the differently colored squares is crucial for solving the puzzle. Some subjects in their experiment were given a special board where a word, Bread or Butter, was printed on each square, instead of colored black or white. Kaplan and Simon hypothesized that these subjects could notice parity more easily and solved the puzzle more quickly, since bread and butter connote parity. Their hypotheses were confirmed. These subjects actually noticed the parity and solved the puzzle more quickly.

However, they reported one puzzling result. The times from their first mention of parity to times of the final solution were longer for these subjects than those who were given a standard checkerboard or blank one. While subjects given a Bread–Butter board took on average 653 s to solve the puzzle from their first mention of parity, those given a standard checkerboard took 110 s.

Why could not the subjects in the Bread–Butter condition make use of parity information immediately after they noticed it? It seems inappropriate to explain this result in terms of retrieval failure, because noticing parity is supposed to activate related operators. Kaplan and Simon suggested that the parity cue does not work effectively until subjects abandon the covering problem space where they attempt to manually cover the board by dominos. It means that recognition of failure of the initial attempts is the key to use of parity information.

A similar view with stronger emphasis on the roles of failure was proposed by Seifert and her colleagues (Seifert, et al., 1995). According to their opportunistic assimilation hypothesis, when people find that a standard approach does not work, they generate failure indices that mark an initial problem solving attempt as an impasse. These failure indices are supposed to have special status in long–term memory, in the sense that they keep activated for a longer period than other types of memory traces. In the incubation phase where people stop their initial attempts and are engaged in other activities, relevant cues are sometimes provided externally, which remind them of their initial failure and lead them to an AHA experience.

Realizing failure and resulting memory traces undoubtedly play key roles in insight. It sometimes causes the shift from one problem space to another and/or generates a failure index that reminds problem–solvers of a past episode.

However, the roles of failure must be analyzed more carefully if the at-what-point problem is the issue. It is not always the case that people switch from one problem space to another, immediately after noticing failure. Such a "digital" switch is rather rare. Also known as (functional) fixation, people usually recognize the failure of a standard approach to an insight problem, but it is awfully difficult to escape from it. Kaplan and Simon reported an episode where a graduate student stuck in a wrong problem space of the MC puzzle, spent 18 hours, leaving a 61 pages long note without success.

Furthermore, explicit noticing of failure may not be a cause for the shift, but a result in some

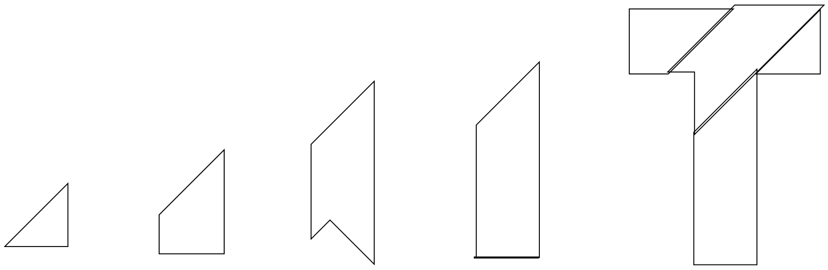

Figure 1: The T puzzle: Construct a shape of "T," using four pieces above.

cases. Even before noticing it, people could reach insight if a very relevant cue is self–generated or externally provided. In such cases, the failure of the initial attempt can be recognized only retrospectively.

Explaining insight in terms of a failure or its index, as employed by Seifert et al. has several difficulties. One of the problems is what is regarded as failure. More specifically what information is included in an episodic memory trace? Needless to say, mere recognition of the failure does not suffice. Suppose someone tries to cover the checkerboard by placing dominos vertically and finds that it does not work. If he interprets his unsuccessful attempts as the impossibility to cover by placing dominos vertically, his next attempt is likely to place them horizontally.

Moreover, failure indices become useful only when it is coupled with a proper solution strategy. If the same type of a failure occurs between the current problem and the past one, it is likely that these problems are isomorphic (Schank, 1982). However, understanding the isomorph itself does not enable problem solvers to reach the solution (Gick, 1992).

Thus, it is dubious that failure "triggers" an insight. Rather, the effects of failure seem to be indirect and gradual. It might be that accumulated experiences of failure cause a cognitively unstable state which in turn cause people to be sensitive to useful information previously unnoticed or neglected.

In terms of the at-what-point problem, there is another phenomenon to be noted, that is, large individual differences. For example, a subject in Kaplan and Simon's experiment spent only two minutes to cover the board by dominos, before shifting to another problem space, whereas, another subject took 18 hours only to find his attempts wrong. These two are quite different in recognizing the failure of their initial attempts. What determines the persistence of the first subject and the quick shift of the second subject? This is another problem to be explored.

2 The T Puzzle

We used the T puzzle, similar to the tangram, as material. The goal of this puzzle is to construct the shape of a "T"[1], using four pieces depicted in Figure 1.

At first glance, it appears to be very easy to solve, since there are only four pieces and one can easily identify possible positions that some of them should be placed. However, a pilot study, in addition to our own experiences, showed that no one, without having experience with this kind of puzzle, could solve it within five minutes. It usually took more than half an hour to solve it

---

[1] "T" is not a sole shape that can be made using these pieces. One can construct more than a dozen of well–formed shapes.

spontaneously. Furthermore, not a few claimed that it was impossible, and some gave up their attempts.

The reason for the difficulty can be attributed mostly to people's tendency to fill the swallow–tailed part (the notch) of the pentagon by the other pieces. However, this part must be kept so as to form one of the right–angled (or 270 degrees) corner that the horizontal and vertical bars of "T" form.

As Weisberg strongly claimed, one has to be very careful in selecting material when analyzing insight (Weisberg, 1995). One of the most important criteria he advocated is *restructuring*, a change of representation.

Does the T puzzle meet this criterion? The answer is positive. A pilot study revealed that, before insight, all the subjects spent most of their time to fill the notch of the pentagon. These fruitless attempts probably results from their belief that the "T" consists of a pair of vertical and horizontal bars. If the "T" shape is partitioned as such, people were strongly inclined to set "construct a pair of bars" as the subgoal. Consequently, they make every effort to fill the notch by the other pieces, since there are no such parts in each bar.

On the other hand, when an insight takes place, people seldom took such actions. Instead, they tried to extend the pentagon horizontally and vertically by attaching the other three pieces. This is probably because they realize that the pentagon is a part of the horizontal *as well as* vertical bars.

Obvious from the above, there are substantial differences in the representations before and after the insight. First of all, the main subgoals are different between the two. While people aim at constructing two bars before the insight, vertical and horizontal extensions of the pentagon become their main subgoal after the insight. Second, the objects and their relations are differently represented. Before the insight, each piece is represented as a part of either the vertical or horizontal bar. In contrast, the pentagon is represented as occupying a part of both in the horizontal and vertical bars. Due to these differences, people take different strategies before and after the insight. While one makes every effort in filling the notch before the insight, such an effort is never made in the latter. Instead, people try to make horizontal and/or vertical parts, letting the notch be itself.

Therefore, when the insight occurs in solving this puzzle, restructuring at a profound level is required.

3 Framework

Our interest is not to sketch the representational change in the way described above, but to explore the general cognitive architecture that produces initial impasses and illumination.

In order to do this, we first outline the general framework to explore the at-what-point problem. In order for the framework to be general enough, we use Duncker's "Candle" problem (Duncker, 1945) to exemplify them, instead of using the "T" puzzle. Next, we derive from it three constraints specific to the "T" puzzle and finally form hypotheses to be tested empirically.

Our strategy is basically in line with Ohlsson's (Ohlsson, 1992). But, we aim at developing a more coherent framework, so that the mechanisms of impasse and illumination are treated in a uniform manner.

3.1 General Framework

Whether a given problem needs an insight or not, the representation of a problem involves, at least, objects, their relations, and a goal.

When representing objects, we have a natural tendency to recognize it at the basic level (Rosch, 1978). This tendency sometimes becomes an obstacle for insight. For example, in the "Candle" problem, it is well known that people do not notice a pasteboard box of tacks as a holder of the candle. This is because the basic level of a box is "box," not a "solid body"(more abstract) or a "pasteboard box" (more concrete).

Relations define the ways how the object relates to or interacts with the other objects. They assign a specific role to objects. Usually, there is more than one relation that can connect objects. Tacks in the candle problem can interact with others in the ways of tacking, standing on something, being thrown, etc. However, the objects and goal force us to select a specific relation to be considered. One reason why the relation that tacks are connected to other objects is tacking is that the basic level recognition of the tack prohibits a relation such as standing on something. Another reason is that the goal prevents them from being thrown.

The representation of a goal involves the desired state and evaluation functions. If the current state is evaluated to increase the degree of match to the desired state, the processing proceeds to the succeeding stage. Otherwise, the current operation is, if possible, canceled and the process goes back to the preceding stage. The goal greatly constrains how objects and relations are represented. Although a relation of a candle to other objects is, by default, to light something, a relation such as to glue something by its wax is likely to be selected by the goal.

We will call them *object–level*, *relational*, and *goal constraints*, respectively. These constraints sometimes operate in harmonious ways to select one interpretation of the problem in some cases, they contradict each other in other cases.

According to our framework, it is important to note here that the strength of each constraint dynamically changes in the course of problem–solving. The shift from an impasse to an insight is explained by the relaxation of the constraints due to recognizing failure. In the course of problem–solving, the strengths of the initially dominant constraints gradually decrease , while less dominant constraints become to operate more frequently. As a result, less dominant constraints come into effect more frequently and lead to an insight. This changing process can be seen as constraint relaxation among object–level, relational, and goal constraints. However, we regard an insight as the product of a specific combination of constraints that are initially less dominant, relaxing a single constraint does not necessarily lead a person directly to an insight.

3.2 Constraints in the "T" puzzle

How are these constraints instantiated in the "T" puzzle? First, the object–level constraints for the puzzle can be expressed as those concerning how a single piece is placed. For example, people usually draw a trapezium with its longest side at the base Although we cannot specify one default way of piece placement, we can rule out many unnatural (but sometimes crucial, especially in insight problems) placements, such as a trapezium standing at one of its apexes.

Relational constraints in the "T" puzzle are concerned with how one piece is physically connected to others. This kind of puzzle has an infinite number of relations, because one can produce different patterns by sliding a side of a piece that is touching another. However, when solving the puzzle, people seem to follow, at least, two rules. The first one is that two pieces are joined so that the apex of each piece forms a single apex. Thus, a connection in Figure 2 (a) will be avoided. The second is that two pieces are aligned in a way such that a side of each piece forms a single side. Thus, a connection in Figure 2 (b) will be preferred to (c). These constraints greatly reduce the number of possible connections.

The goal evaluates particular placement and combinations of pieces. The greater a placement or combination overlaps the "T" shape, the higher value is assigned. For example, Figure 2 (b) will be rated high because it overlaps the horizontal bar of the "T."

As we noted above, constraints cannot exist independently. In the course of problem–sovling they interact each other and induce an insight. An instantiation of the interaction between constraint types for the "T" puzzle is summarized as follows:

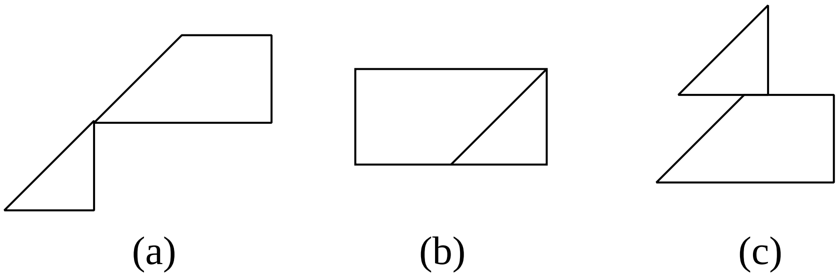

Figure 2: Connections of two pieces.

1. Failure is recognized with goal constraints by evaluating the rate of matching.
2. With accumulated experiences of failure, the strengths of object–level constraints of placing the pentagon horizontally or vertically decrease gradually.
3. The relative strengths of object–level constraints of placing the pentagon diagonally gradually increase.
4. With changes of object–level constraints, the strength of relational constraints of filling the notch decreases.
5. As a result, the rate of matching goes up.

Ohlsson restricts the constraint relaxation only to the goal (Ohlsson, 1992). On the contrary, our framework supposes that all the constraints be relaxed by the recognition of failure. Another difference is that Ohlsson emphasizes the change of the goal by constraint relaxation, while we suppose the goal itself to be constant. The difference lies in the fact that his argument is concerned with the subgoals or the interpreted goal, rather than *the goal* of the problem. We completely agree that subgoals must be changed in solving insight problems. In our framework, however, subgoals are implicitly represented as a product of interacting constraints, and their change is caused by the relaxation of these constraints due to the recognition of failure.

4 Study 1

From the general framework and its derivatives, we form four hypotheses to be tested empirically.

Placing the pentagon horizontally and vertically. Since the object–level constraints force subjects to place the pentagon with either the longest or the second longest side as the base, subjects' use of the pentagon will mainly be horizontal before insight. This is enforced by the goal constraint, because the pentagon placed in this way will produce relatively a high degree of match to the horizontal bar of the "T" shape. However, it is also expected that the pentagon is placed vertically, due to the relational constraint. Especially when the triangle is attached to the top side of the pentagon depicted in the Figure 1, the resulting shape matches well to the goal. Thus, placing the pentagon horizontally should be equally frequent [2].

Filling the notch. It is also predicted that people will spend most of their time filling the notch of the pentagon. Given, as mentioned above, horizontal or vertical placement of the pentagon

---

[2] This hypothesis at first, appears to be implausible, considering that pieces may be placed either horizontally or vertically. However, it is reasonable, because there are numerous other ways to place pieces, for example, placing it aslant with 10, 30, 45 degree, etc.

is greatly enforced, people should easily find the only non–overlapping part with "T" to be the notch. In such a case, the goal constraint will strongly motivate subjects to fill the notch.

Ignoring useful information. Because of the high pressures from the object–level and the goal constraints, subjects in the early stage of their processing cannot reach an insight even when the pentagon is occasionally placed in the proper way. In addition, such a placement itself does not produce a high degree of match, because it is difficult to find the overlapping part in the "T" shape.

Relaxing constraints by failure. Repeated failures should relax all the constraints that have enforced subjects to place the pentagon in wrong ways and to fill the notch. As a result, the relative frequency of the proper placement of the pentagon will increase gradually.

4.1 Method

Subjects Six undergraduate students were participated in the study. Since one subject was suspected of having prior experience on this puzzle, data from remaining five were analyzed.

Procedure Subjects were just asked to place the pieces to form the shape of "T." If the subjects could not solve it within 15 min., they were given a hint not to fill the notch. A further hint was prepared, but was not used, because those who could not have solved the puzzle before the first hint solved it immediately after the hint. Although the subjects were required to think aloud, very few did it. Consequently, we did not to perform the protocol analysis.

4.2 Results and Discussion

We sliced an entire problem–solving process of a subject to segments. A segment roughly corresponds to an attempt that begins with connecting pieces and ends up with noticing failure or achieving the goal. A segment is operationally defined as a series of actions that was initiated by joining two pieces and terminated by their separation. Thus, an action such as adding another piece to joint pieces does not constitute a segment, but is regarded as an action within a segment. An exception of the operational definition above is concerned with the rotation of joint pieces. We regarded such an action as a beginning of a new segment.

Table 1 shows the time to solve the puzzle and the number of segments for each subject. Of the five subjects, two could solve it without the hint, the others received the hint. As expected, there are large individual differences: the slowest subject took about seven times longer than the fastest, and the number of the segments varied from 19 to 90. Thus, we will present individual data, rather than aggregated ones, and accordingly perform no statistical analysis.

Table 1: Solution time and the number of segments of each subject. Subjects K, O and A were given the hint when about 900 seconds had passed.

	Subjects				
	H	M	K	O	A
Solution time (Sec.)	140	553	1025	1027	936
No. of Segments	19	47	90	75	83

Placing the pentagon horizontally. In order to test the first hypothesis, we compared the proportions of the types of pentagon placing. When calculating the proportions, we excluded segments where the pentagon was not involved, of which percentages ranged from 4 to 18%.

Table 2 shows the types of placing the pentagon. We can easily find that the subjects placed the pentagon horizontally or vertically more often than in the other ways.

Table 2: Percentages of the types of the pentagon placement. H/V denotes the placement of pentagon horizontally or vertically.

	Subjects				
	H	M	K	O	A
H/V	62	66	85	91	76
Other	38	34	15	9	24

There is one thing to be noted, the difference between those who could solve the puzzle spontaneously and those who could not. Although we did not perform any statistical tests due to the large individual differences, it seems that the subjects who needed the hint tended to place the pentagon horizontally or vertically more often than those who did not. This may indicate that insightful problem–solvers are more free, though limited, from the object–level constraints.

Filling the notch. The proportions of segments where the subjects filled the notch were .79, .62, .78, .68 and .60 for each subject respectively, in the order same as Table 1 and 2. These data again support our hypothesis that the main task for the subjects is to fill the notch.

Ignoring useful information. Table 2 also shows that the pentagon was placed in a way that it violated the initial object–level constraints. The number of such placement is relatively small, but non–negligible. This clearly shows that the subjects failed to recognize important cues, although these cues were generated a lot before the insight took place.

Relaxing constraints by failure. One of the most important thing that our framework proposes is that the constraints are gradually relaxed due to the recognition of the repeated failures. We predicted that the number of pentagon placements that violate the initially dominant object–level and relational constraints should increase.

In order to test this hypothesis, we simply divided an entire problem–solving process of each subject into the first and second halves and compare the frequencies of placements where the initial constraints are violated. Obvious from Table 3, the number of such placing dramatically increased in the second half. This strongly supports our claim that the object–level and relational constraints be relaxed in the course of problem–solving.

Table 3: Frequencies of the constraint–violating placement of the pentagon. Cases occurred after the hint are excluded for the subjects, K, O, A.

	Subjects				
	H	M	K	O	A
First	2	5	4	0	5
Second	4	10	8	5	10

Another thing to be noted is that, although the proportion of the constraint–violating placement increases, this type of placement is distributed over the process even in the second half. Figure

3 illustrates the fastest subject's performance in the second half. The second row shows how the pentagon is placed, H is the abbreviation for horizontal placement of the pentagon, V for vertical, D for diagonal, and X for other placement which did not involve the pentagon. This subject placed the pentagon diagonally at from the segment 12 to 14, but these attempts were interrupted by other types of placement, and finally reached an insight by placing the pentagon properly. This indicates that "switching by failure" was not the case even for the most insightful subject.

Segment	10–11	12–14	15–16	17	18	19
Pentagon	H	D	X	V	H	D
Notch	F	F	X	F	F	

Figure 3: Subject H's performance in the second half.

5 Study 2

The results of the previous study indicate that the three constraints are concerned with four characteristics of the subjects' performance. In the study 2, we try to validate our interpretations of the study 1, by giving hints to facilitate subjects to reach insight. First, we conjectured that the reason why subjects made fruitless attempts could be attributed to the object–level constraints that strongly enforce the particular placement of the pentagon. Second, the data suggested that, in the early stages of problem–solving, difficulties in matching to the goal prevent subjects from getting insight even when the pentagon is occasionally placed in the proper way.

To give further evidence to our interpretations above, one group of subjects were given the puzzle, with fixing the pentagon at the proper position and told to solve it without moving the pentagon. Since this manipulation does not allow the object–level constraints that are initially dominant to operate, subjects in this group should be able to reach an insight more rapidly. To validate the second interpretation, another group of subjects were given a sheet of paper where "T" was printed in the real size, as well as the four pieces of the puzzle. Since the presence of the template is expected to facilitate the evaluation of the match to the goal, subjects in this group should find the proper position for the pentagon to be placed.

5.1 Method

Subjects Eight undergraduate students were participated in the study 2. They were randomly assigned to one of the two conditions: fixed or template condition.

Procedure Subjects in the fixed condition were given the puzzle, with placement of the pentagon at the proper position. They were requested to solve it, without moving the pentagon. Subjects in the template condition were given a printed template described above as well as the four pieces and asked to place them in the way that the printed "T" was covered.

5.2 Results

Table 4 shows the solution time for each subject in the fixed and template conditions. The performance in the fixed condition is improved dramatically. No one needed the hint. In addition, of the four, three subjects could solve the puzzle within a minute. Furthermore, even the slowest subject required the smaller number of the segments than the fastest subject in the study 1.

Table 4: Solution time and the number of segments of each subject. Subject Tb was given the hint when about 900 seconds had passed.

	Fixed				Template			
	Fa	Fb	Fc	Fd	Ta	Tb	Tc	Td
Time	10	17	337	42	584	957	65	797

In contrast, the performance in the template condition shows ambiguous patterns. Three out of four subjects solved the puzzle without receiving the hint, while the solution time as well as the number of segments were not substantially different form the ones in the study 1.

These results leave two possibilities: one that, against our framework, the goal–constraint does not operate, the other that the template used was not facilitative enough for the goal constraint to operate. Since, as described earlier, the goal constraint operates together with the others, performance of the template condition may be improved under the condition where the other constraints are relaxed. Hence, it is safe to conclude at the moment that there is no direct evidence that denies the roles the goal constraint plays.

6 General Discussion

We suppose that the processes of insight have dynamic nature, in the sense that a relatively small number of constraints interact in a specified way but produce enormously complex patterns appeared as individual differences. In order to explore processes of insight, more specifically what problem–solvers do in the prepartion phase and what happens at the moment of insight, we propose a general framework. This framework involves three constraints, *object–level*, *relational*, and *goal* that operate simultaneously and are gradually relaxed by the recognition of failure. A triplet of specific constraints derived from the framework and their relaxation provides reasonable accounts for the performance in solving the “T” puzzle.

One of the most important task for our framework is to develop a concrete computer program that can be use to simulate the process of insight. With such a simulation, we can verify and refine our framework in more detail. To do this, a computer program is currently being constructed based on the Boltzmann Machine model (Ackley et al, 1985). The results of the simulation will be reported at the Cognitive Science meetintg.

References

Ackley, D.H., Hinton, G.E., Sejnowski, T.J. (1985) A learning algorithm for Boltzmann Machines, *Cognitive Science*, **9**, 147 – 169.

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

Gick, M. L. (1992) Learning from mistakes: Inducing analogous solution failures to a source problem procedures later success in analogical transfer. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, **18**, 623 – 639.

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

Ohlsson, S. (1992) Information processing explanations of insight and related phenomena. In M. T. Keane and K. J. Gilhooly (Eds.) *Advances in the Psychology of Thinking, vol. 1*, Hertfordshire, UK: Harvester.

Rosch, E. (1978) Principles of categorization. In E. Rosch and B. B. Lloyd (Eds.) *Cognition and Categorization.* Hillsdale, NJ: Erlbaum.

Schank, R. C. (1982) Dynamic Memory: A Theory of Reminding and Learning in Computers and people. Cambridge University Press.

Seifert, C. M., Meyer, D. E. Davidson, N., Patalano, A. L., and Yaniv. I. (1995) Demystification of cognitive insight: Opportunistic assimilation and the prepared-mind perspective. R. J. Sternberg and J. E. Davidson (Eds.) *The Nature of Insight.* MIT Press.

Sternberg, R. J. and Davidson, J. E. (Eds.) (1995) *The Nature of Insight.* MIT Press.

Weisberg, R. W. (1995) Prolegomena to theories of insight in problem solving: A taxonomy of problems. R. J. Sternberg and J. E. Davidson (Eds.) *The Nature of Insight.* MIT Press