

An investigation of search strategies for hypothesis generation using eye movement data

Miki Matsumuro (muro@cog.human.nagoya-u.ac.jp)
Kazuhisa Miwa (miwa@is.nagoya-u.ac.jp)

Graduate School of Information Science, Nagoya University Fro-cho, Chikusa-ku, Nagoya, Japan

Abstract

The hypothesis generation process is divided into two phases: the search phase and the decision phase. In the former phase, many possibilities for the hypotheses are considered; and in the latter phase a focal rule is decided upon. In the current study, we investigated the search phase using eye movement data. We detected two types of search strategy: one strategy was a spread-attention search in which participants spread their attention over a broad area in rule space; and the other was a focused-attention search in which participants concentrate their attention on a narrow area in rule space. Results of our experiment showed that the focused-attention search strategy was more effective for rule discovery than the spread-attention search strategy.

Keywords: Rule discovery; hypothesis generation; eye-tracking; search strategy

Introduction

Various aspects of the processes of reasoning and discovery have been focused on in a considerable number of studies (e.g., Klahr, 2000; Mynatt, Doherty, & Tweney, 1977; Qin & Simon, 1990; Haverty, Kiedinger, Klahr, & Alibali, 2000). The process of discovery is divided into two phases: hypothesis generation and hypothesis testing. Although there have been many studies about hypothesis testing, only a few studies have focused on the mechanisms of hypothesis (rule) generation. In this study, we investigate the hypothesis generation process, dividing it into the search phase and the decision phase. The decision phase is the final part of the hypothesis generation process. In this phase, participants produce a focal rule by which they predict an instance for verifying the hypothesized rule. The search phase precedes the decision phase. In this phase, participants search in rule space checking various possibilities for rules before producing their focal rule. Almost all studies about hypothesis generation have investigated the focal rule produced in the decision phase. Only a few studies have examined the search process in the search phase, due to the difficulties of observing and recording this process.

One index for observing the search process is participants' protocols during hypothesis generation (e.g., Mynatt et al., 1977; Qin & Simon, 1990; Haverty et al., 2000). Verbal protocols are actually useful for understanding participants' thoughts during the discovery processes. However such protocol data are very coarse-grained and contain only conscious thoughts. Participants may say nothing or make long pauses because they cannot externalize all thoughts while engaging in a task. In this study, we use eye movements as a way to record the participant's search process in the search phase. In recent studies of category learning or insight problem solving,

researchers have begun to use eye movement data as a useful index (Knoblich, Ohlsson, & Raney, 2001; Rehder & Hoffman, 2005). We assume that eye movements provide good information about the participants' attention during a task. Compared with other indexes such as verbal protocols, eye movements give us more fine-grained data about the search process without hindering the participant's activity.

We focus on the spread of attention in the search phase. First, let us consider the relation between the spread of attention and search strategies in the search phase. A rule is generated through a search of rule space that contains all available rules (Simon & Lea, 1974). These rules are subdivided into some sets, the members of which are rules that have the same or similar attributes (Figure 1). In this study, we define each set as a "hypothesis set." Similar rules characterized by identical attributes are contained in an identical hypothesis set.

We assume that two search strategies in the search phase are defined based on to what degree the participants spread their attention when searching in rule space. One strategy is the spread-attention search strategy in which the participant's attention is broadly spread in rule space. Participants who use this strategy search in more than one hypothesis set before they finally produce a focal rule (Figure 1(a)). They search various rules belonging to various hypothesis sets and then produce a focal rule(s). Another strategy is the focused-attention search strategy in which the participant's attention is concentrated on one (or a few) hypothesis set in rule space. Participants who use this strategy search in only one hypothesis set before they produce a focal rule (Figure 1(b)). They selectively search similar rules belonging to an identical hypothesis set and then produce a focal rule(s).

Our question is which of these strategies is effective for

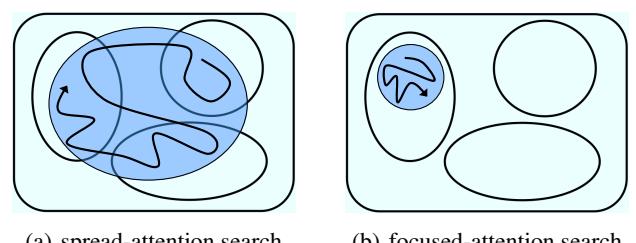


Figure 1: Concept diagrams of two search strategies
A square is a hypothesis space and each non-filled circle in a square is a hypothesis set. (a) Participants who use the spread-attention search strategy search for hypotheses across multiple hypothesis sets. (b) Participants who use the focused-attention search strategy search for hypotheses within one hypothesis set.

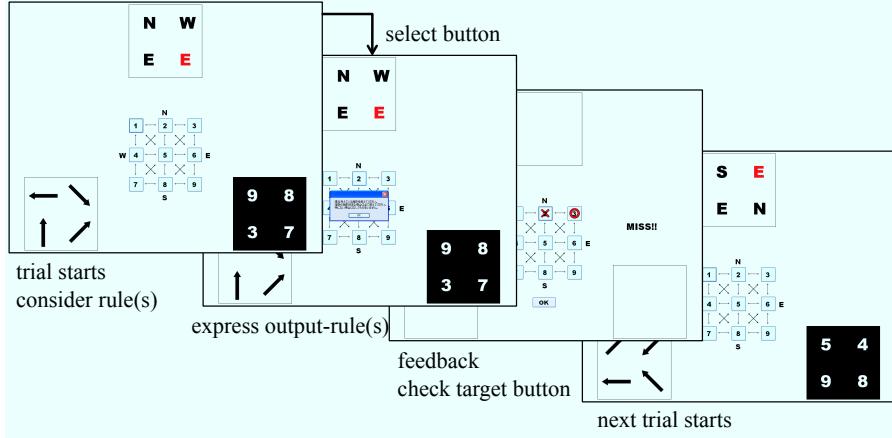


Figure 2: Example screen shot of task display in this study

The upper panel is the arrow panel, the lower left panel is the compass panel, and the lower right panel is the number panel. The nine squares with numbers in the center are buttons for prediction. The rule was consistent through the task.

rule discovery. In terms of producing a focal rule in the decision phase, some studies using Wason's 2-4-6 task suggested that a variety of generated rules was important for rule discovery (Gale & Ball, 2009; Vallée-Tourangeau, Austin, & Rankin, 1995). Generation of a variety of rules means that the produced focal rules are scattered in a broad area in rule space. Turning now to the search phase, which strategy should be used to find the correct rule: the spread-attention or the focused-attention search?

Two opposite answers are possible. The first answer is that the spread-attention search will be more effective. Participants who use the spread-attention search will be able to reach a hypothesis set that contains a correct rule more quickly because they search a broader area of rule space at a time. Additionally participants who use the focused-attention search may fall into an area of restricted interest, searching in an irrational hypothesis set that involves irrelevant rules. This possibility will be minimized if participants search more than one hypothesis set at a time.

In contrast, the alternative answer is that the focused-attention search will be more effective. Participants who use the focused-attention search can systematically search each hypothesis set one by one. Furthermore due to the limit of cognitive resources, even though participants use the spread-attention search, they might not be able to use all of rich information drawn from searches in various sets. We refer again to this point in detail later. To verify which of these two possibilities is correct, we observe the degree of spread of participants' attention using the eye-tracking method and investigate the relation between the search strategies and possibilities of reaching successful rule discovery.

To investigate the effectiveness of the spread-attention search, we also analyze the relation between the hypothesis sets searched in the search phase and the sets containing the produced focal rule(s) in the decision phase. From a normative point of view, if participants search multiple hypothesis sets using the spread-attention search, they should produce various types of focal rules from all of the hypothesis sets they search. In contrast, if participants search only one hy-

pothesis set using a focused-attention search, they can be expected to produce a single or multiple focal rules from the one hypothesis set they search.

In summary, we have three main objectives in the current study. First, by using eye movements as an index for observing the search process of hypothesis generation, we expect to detect two types of search strategies characterized by the degree to which attention is spread in a rule space search. One strategy is the spread-attention search in which participants spread their attention over a broad area in rule space. They search more than one hypothesis set before they produce a focal rule. Another strategy is the focused-attention search in which participants concentrate their attention on a narrow area in rule space. Typically, they search only one hypothesis set before they produce a focal rule. The second goal is to examine which strategy is effective for rule discovery. To do so, we investigate the relation between the search strategy and successful rule discovery. The third goal is to explain the effectiveness of the two search strategies based on the relation between the focal rule(s) produced in the decision phase and the hypothesis sets searched in the search phase.

Task

To observe the spread of attention in the search phase by eye-tracking, we developed a new rule discovery task in which the search in each hypothesis set as a mental process corresponds to the focusing of eye movements on a local space in the task display. Figure 2 shows an example screen shot of the task display in this study. The task display consists of three panels (arrow, compass, and number panels) and nine buttons. One of the nine buttons is a target button. Which of the nine buttons is the target button can be correctly inferred from the objects displayed on one of three panels. The participants are asked to find a rule, as quickly as possible, which correctly predicts the target button.

The experimental procedures are as follows. First, the participants predict the target button from the objects on the panels and select one of the nine buttons. Second, as soon as they select the button, they are prompted to report the rule(s) they

used for their prediction. We refer to this rule as an “output-rule.” Third, after they report their output-rule(s), the correct target button is presented. At this time, the objects on all panels are erased so that the participants cannot refer to them. The procedures above constitute one trial in the experiment. After the participants check the correct target button, the next trial starts. For the next trial, the objects on each panel are newly constructed and changed from those of the previous one. The location of the target button is also changed regardless of the success or failure of the participants’ prediction. The amount of trials is not controlled. When the participants make four successive references to the correct rule and correctly predict the target button, the task is terminated. The participants are given seven minutes to find the rule. The period during which the participants express their output-rule(s) is eliminated from this time. One block starts when the task starts and ends when the participants find the rule or seven minutes elapsed.

The following are example rules that the participants may infer.

In the number panel, one can predict the target button by calculating the sum or difference of two numbers selected on the panel. One example rule in Figure 2 may be: use the sum of the two numbers shown in upper left and upper right positions. In this case, the target button is Button 5 because that is the sum of 3 (upper left) and 2 (upper right).

In the compass panel, one can predict the target button by combining two directions selected on the panel. Each of Buttons 2, 6, 8, and 4 corresponds to one of four directions, north, east, south, and west, respectively. An example rule in Figure 2 is: upper right and lower right directions on the panel predict the target button. The target button is the southeast (i.e., Button 9) because the upper right and lower right buttons are east and south. On the compass panel, one direction colored with red sometimes plays a specific role for describing a rule.

In the arrow panel, one can predict the target button by the transition of a pointer guided by the arrows on the panel. The pointer starts at the center of the nine buttons as an initial position (i.e., Button 5). One example rule in Figure 2 is: use four arrows in a clockwise direction from the upper right arrow. In Figure 2, the pointer moves from Button 5 to Button 6 (by right pointing), stays at Button 6 (by upper right pointing), moves to Button 4 (by left pointing), and moves to Button 1 (by up pointing). Therefore the target button is Button 1.

In this study, we assume that each output-rule reflects a focal rule produced in the decision phase and eye movements before reporting the output-rule(s) reflect the search process in the search phase. Rules drawn from the objects in an individual panel are characterized by similar attributes and construct a single hypothesis set: i.e., rules in the arrow panel construct the arrow hypothesis set, the rules in the compass panel construct the compass hypothesis set, and the rules in the number panel construct the number hypothesis set in rule

space. Eye movements on a certain panel are interpreted as a search of the hypothesis set corresponding to that panel.

When the participants use the spread-attention search strategy, they search more than one hypothesis set. In this case, eye movements focusing on more than one panel are observed before participants report an output-rule(s). When the participants use the focused-attention search strategy, they search only one hypothesis set. In this case, eye movements are fixed on a single panel.

Experiment

Method

Participants Twenty-two undergraduates participated in the experiment.

Apparatus We presented the task display on a 17 in. monitor with a resolution of 1280×1024 pixels. The participants were seated approximately 60 cm away from the monitor. The panels equaled approximately $7.9^\circ \times 7.9^\circ$ of visual angle. The center of each panel was set 21.7° away. The participants’ eye movements were recorded using the Tobii T60 eye tracker. The participants were allowed to move their heads naturally.

Procedure Before starting the reasoning task, all participants received some basic instructions about the operations of the experimental system and example rules. They sufficiently learned example rules for complete understanding of the structures of rules used in the experiment. All participants engaged in a total of five blocks. Table 1 shows the rules used in the experiment. The easy rules were similar to the example rules learned in the instruction phase and the difficult rules were relatively different from the example rules and more complex than the easy rules.

The training phase consisted of three blocks. The rules used in this phase were three easy rules. Each rule was incorporated into one of the three hypothesis sets respectively. After a brief break, the test phase consisting of two blocks followed. Two rules were selected randomly from the three difficult rules. In both training and test phases, the order of the rules and the positions of the panels were decided randomly. Calibration for recording eye movements was conducted before each block started. After each block, we told the participants the correct rule regardless of whether they had been able to find the rule or not.

Table 1: The rules used in this study

hypothesis set	easy	difficult
arrow	clockwise from upper-right arrow	N shape from lower-left arrow
compass	upper-left and lower-left	red letter and next to it
number	difference between lower-left and lower-right	sum of upper-left and lower-right

Results

Two Strategies in the Search Phase First, we investigated what strategy participants used in the search phase before they decided on a focal rule(s), was effective for rule discovery. For categorization of the participants' search strategies in the search phase, we defined an index that represented to what degree the participants spread their attention. We used eye movement data from each participant for this index. We used only the data from the beginning of one trial starting after the participants failed to predict the target button through the point they reported an output-rule(s). When the participants predicted the target button correctly, we excluded the data from the analysis because after their correct prediction they did not generate a new rule so that their attention was fixed to one hypothesis set containing the maintained rule for verification.

The adjusted entropy was defined as follows where p_i ($i = 1, 2$, and 3) means the proportion of fixation time of eye movements in each of the three panels.

$$\text{adjusted-H} = -\frac{\sum p_i \cdot \log p_i}{\log 3} \quad (1)$$

A large value for the adjusted-H means that the participants searched multiple panels; in other words, they spread their attention over a broad area in rule space. A small value for the adjusted-H means that the participants concentrated their attention on one panel, i.e., one hypothesis set. The adjusted-H increased to the maximum value 1 when the participants spread their attention equally over three panels. The adjusted-H decreased to the minimal value 0 when the participants concentrated their attention on only one panel.

First of all, we verified the consistency of the search strategies used by the participants. Table 2 shows correlations of the averages of the adjusted-H among five blocks. All values between any set of two blocks were higher than .5, suggesting that each participant used a consistent strategy throughout the five blocks. In particular, the correlations among the third block in the training phase and two blocks in the test phase exceeded .8, meaning a very strong consistency. Since each participant seemed to use a consistent strategy, we categorized the participants into two groups using the mean value of adjusted-H for the two blocks in the test phase.

The criterion for categorization was the median of the mean values. The median value was 0.216. The focused group consisted of the participants whose mean adjusted-H values were below 0.216; the spread group consisted of those whose values were above 0.216. The participants in

Table 2: Correlations between the mean of adjusted-H

	training 1	training 2	training 3	test 1	test 2
1	1.000				
2	.729***	1.000			
3	.838***	.661**	1.000		
1	.752***	.620**	.889***	1.000	
2	.636**	.668**	.809***	.896***	1.000

** : $p < .01$, *** : $p < .001$

Table 3: Discovery rate of each group in each phase

		spread	focused	all
training	arrow	0.18 (2)	0.64 (7)	0.41 (9)
	compass	0.36 (4)	0.91 (10)	0.64 (14)
	number	0.91 (10)	0.73 (8)	0.82 (18)
	all	0.48 (16)	0.76 (25)	0.62 (41)
test	arrow	0.13 (1)	0.17 (1)	0.14 (2)
	compass	0.13 (1)	0.63 (5)	0.38 (6)
	number	0.83 (5)	0.88 (6)	0.79 (11)
	all	0.32 (7)	0.55 (12)	0.43 (19)
overall		0.42 (23)	0.63 (37)	0.56 (60)

We show the number of participants who found the rule in parentheses. The all number of each block was eleven in training phase and six or eight in test phase.

the focused group tended to use the focused-attention search strategy; those in the spread group used the spread-attention search strategy. For confirmation, the mean adjusted-H of the participants in the spread group was 0.477, significantly higher than 0.078 in the focused group ($t(11) = 6.730$, $p < .001$).

Discovery Rate The discovery rate means the proportion of the blocks in which the rules were successfully found (Table 3). The discovery rate through all five blocks was .56. For each phase, the discovery rate was .62 in the training phase and .43 in the test phase. The difference between the two phases revealed marginal significance ($\chi^2(1) = 3.819$, $p = .051$). As expected, the proportion of successful rule discovery in the test phase during which relatively difficult rules were used, was lower than the proportion in the training phase. The effect of the rule type also reached significance. The discovery rate for the number rule was the highest and the discovery rate for the arrow rule was the lowest in each phase. The discovery rates for arrow, compass, and number rules in each phase are as follows respectively; .41, .64, and .82 in the training phase ($\chi^2(2) = 18.236$, $p < .001$), .14, .38, and .79 in the test phase ($\chi^2(2) = 7.856$, $p = .020$), and .31, .56, and .81 in whole blocks ($\chi^2(2) = 12.122$, $p = .002$). In particular, the discovery rate for the arrow rule in the test phase was only .14, meaning that the rule was extremely difficult.

Table 3 shows the discovery rates for each group. Through five blocks the participants in the focused group found the rules more frequently than the participants in the spread group ($\chi^2(1) = 7.187$, $p = .007$). The same tendency was observed in the training phase in which the discovery rate was spread in the focused group than that in the spread group ($\chi^2(1) = 5.216$, $p = .022$). However, there was no significant difference between the two groups during the test phase ($\chi^2(1) = 2.316$, $p = .128$). For a detailed analysis, we compared the discovery rate for each rule in the test phase. For the compass rule, the difference between the spread and focused groups reached marginal significance ($p = .059$, Fischer's exact test). For the arrow and number rules, there was no difference between the two groups ($ps > .50$). The arrow rule was too difficult; only one participant in each group could find the rule. On the other hand, the number rule was too easy; five out

of six participants in the focused group and six out of eight participants in the spread group found the rule. In summary, the participants in the focused group were more successful in finding the rule than those in the spread group, strongly suggesting that the focused-attention search strategy in the search phase was a more effective strategy for rule discovery than the spread-attention search strategy.

Search Strategy and Output-Rule Finally for the third goal of this study we investigated whether the participants' focal rules in the decision phase normatively corresponded to their search in the search phase. The analysis was performed based only on the data from the test phase because of its good consistency. The output-rules in each trial were divided into two types: a single set and multiple sets. Regardless of the number of the output-rules they produced, when the output-rules in a trial belonged to an identical hypothesis set, they were categorized as the single set type. When the output-rules in a trial belonged to different hypothesis sets, these were categorized as the multiple sets type. We grouped the participants into those who always produced the single set type output-rules and those who produced the multiple sets type output-rules.

If their focal rules during the decision phase corresponded to the searched hypothesis sets in the search phase, then (1) the participants in the spread group who tended to search more than one hypothesis set in the search phase would be categorized as the latter participants, and their focal rules would be expected to come from the hypothesis sets that they had searched during the search phase, and (2) the participants in the focused group who searched only one hypothesis set in the search phase would be categorized as the former participants whose focal rule corresponded to the one hypothesis set that they searched in the search phase.

The number of participants who produced the multiple sets type output-rules was significantly larger in the spread group than the number in the focused group (5/11 and 0/11; $p = .018$, Fischer's exact test). In the focused group, all participants produced only a single set output-rule(s) in the decision phase. However, contrary to our expectation, in the spread group, only half of the participants produced output-rules in multiple sets.

Discussion and Conclusion

Search Strategy and Eye Movement

In this study, by using eye-tracking data, we succeeded in detecting two search strategies during the search phase. We performed further analysis of eye movement data to obtain more detailed patterns of each participant's search process. The participants in each of the spread and focused groups could be subdivided into two types based on the transition pattern of their adjusted-H values. The categorization was performed based on the degree of variation in the transition of the adjusted-H values. Some of the participants indicated consistent values for the adjusted-H, and others indicated fluctuant values in the transition. For this analysis, we

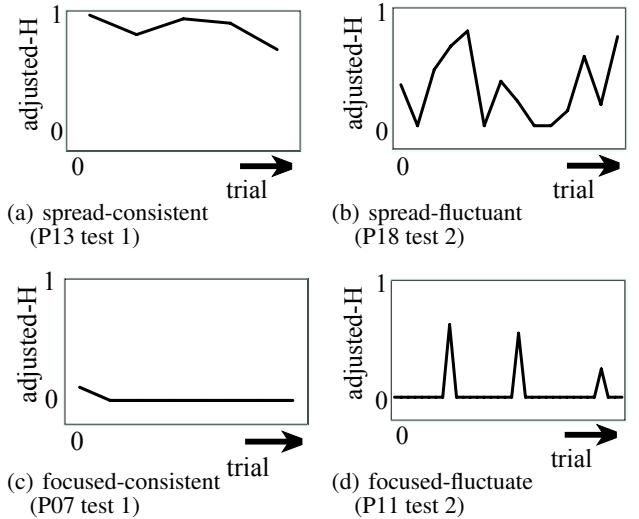


Figure 3: The transition pattern of adjusted-H
The pattern of spread group is presented in upper row and focused group is presented in lower row. The consistent type is presented in left column and the fluctuant type is presented in right column. Each panel is the actual transition of the participants who showed most typical pattern. The x and y axes indicate the trials and the adjusted-H respectively. We show the ID number of participants and the block number in parentheses.

used the transition patterns in two blocks in the test phase. Two participants in the spread group and one in the focused group were excluded from this analysis because two transition patterns in the two blocks were inconsistent. See Figure 3. Five participants in the spread group were categorized into the fluctuant subgroup (Figure 3(b)) and four were categorized into the consistent subgroup (Figure 3(a)). The latter four participants seem to have used the spread-attention search strategy throughout the blocks. Note that four of the five participants who produced multiple sets type output-rules were involved in this subgroup. On the other hand, the former five participants basically used the spread-attention search strategy but sometimes focused their attention into one or two target hypothesis sets. In the focused group, seven participants were categorized into the fluctuant subgroup (Figure 3(d)) and three were categorized into the consistent subgroup (Figure 3(c)). The participants' search strategy in the search phase in the focused group was basically the focused-attention search. Actually, the latter three participants used the focused-attention search strategy throughout the blocks. However the former seven participants searched multiple hypothesis sets in some of trials of each block. By doing so, they may have been trying to find the next target hypothesis set for the following trials.

The above analysis using eye movement data promises high reliability. Analysis based on verbal protocol data has some limitations. For example, the amount of verbalization varies among participants, and no verbal data is produced when participants verbalize nothing. When we try to capture

the individuality of participants in the rule (hypothesis) space search, these limitations become crucial. We believe that eye movement analysis will play a central role in the further studies of hypothesis generation.

Search Strategy and Discovery

The results of this study showed that the focused-attention search strategy in which the participants concentrated their attention on one hypothesis set was more effective for rule discovery than the spread-attention search strategy in which the participants spread their attention to more than one hypothesis set. The reason for this result may be due to the limitation of participants' cognitive resources (Dougherty & Hunter, 2003). The half of the participants who used the spread-attention search strategy did not produce multiple sets type focal rules from all the hypothesis sets they had searched; their output focal rules were limited to a part of the multiple hypothesis sets. This result suggests that their focal rules did not reflect their search during the search phase. They had to temporally keep the assumed focal rules and the predictions from the rules in mind while searching in rule space. Their cognitive resources were insufficient to conduct such activities simultaneously; therefore the participants could not take advantage of the spread-attention search strategy. This limitation was also observed in the participants who generated multiple sets type focal rules belonging to different hypothesis sets. When participants test several rules or hypotheses at a time, a set of diagnostic hypotheses each of which predicts a different instance is effective for rule discovery. However, four of the five participants who expressed the output-rules from the multiple sets did not generate such hypotheses or rules. All of the output-rules they expressed in one trial made one identical prediction, thus saving their cognitive resources.

On the other hand, the participants who used the focused-attention search strategy effectively utilized their search in forming their focal rule(s). They searched each hypothesis set individually under the constraint of their cognitive resources. This systematic search in each hypothesis set allowed them to find a correct rule more effectively. Note that this result of our study and the findings of preceding studies are not mutually inconsistent. In preceding studies, the participants had to generate the focal rules from a broad area in rule space in the *decision phase* (Gale & Ball, 2009; Vallée-Tourangeau et al., 1995). Conversely in our study, the participants had to concentrate their attention on a narrow area in rule space in the *search phase*. Actually, participants could generate the focal rules spread in a broad area in rule space while accumulating the search results in the search phase by using the focused-attention search strategy. Through the results of our experiment, we propose an effective process of rule discovery. In the search phase, participants concentrate their attention on one perspective to search each hypothesis set individually and produce a focal rule(s) and test it in experimentation. Additionally, at an adequate timing, they should also shift their attention from one hypothesis set to another one to produce various types of focal rules from a broad area in

rule space. This perspective may be interesting because the effective strategy is different for each of the two stages of hypothesis generation: the search phase and the decision phase. Further studies are needed to understand in more detail the nature of hypothesis search strategies in the two stages.

Limitation and Future of This Study

The task used in this study had several limitations. The most crucial limitation was that the task was not a discovery task in which participants had to find out what perspective they should use to decide what hypothesis set they should search. The nature of our task was attributed to a search task where participants were given perspectives to search, and the search space was initially decided. The participants could, in principle, reach the correct rule by comprehensive search; but this may be impossible in practice because of the time limitation. The participants in our experiment completely learned three sets of hypotheses containing a correct rule. However a post interview indicated that the participants could not evaluate the number of rules involved in each hypothesis set. In the future, we would like to develop a discovery task that can be used under our experimental paradigm to try to understand the nature of search strategies in hypothesis generation processes using such an experimental task.

References

Dougherty, M. R. P., & Hunter, J. E. (2003). Hypothesis generation, probability judgment, and individual differences in working memory capacity. *Acta psychologica*, 113, 263–282.

Gale, M., & Ball, L. J. (2009). Exploring the determinants of dual goal facilitation in a rule discovery task. *Thinking and reasoning*, 15(3), 294–315.

Haverty, L. A., Kiedinger, K. R., Klahr, D., & Alibali, M. W. (2000). Solving inductive reasoning problem in mathematics: not-so-trivial pursuit. *Cognitive Science*, 24, 249–298.

Klahr, D. (2000). *Exploring science*. the MIT Press.

Knoblich, G., Ohlsson, S., & Raney, G. E. (2001). An eye movement study of insight problem solving. *Memory and cognition*, 29(7), 1000–1009.

Mynatt, C. R., Doherty, M. E., & Tweney, R. D. (1977). Confirmation bias in a simulated research environment: an experimental study of scientific inference. *The quarterly journal of experimental psychology*, 29, 85–95.

Qin, Y., & Simon, H. A. (1990). Laboratory replication of scientific discovery processes. *Cognitive Science*, 14, 281–312.

Rehder, B., & Hoffman, A. B. (2005). Eyetracking and selective attention in category learning. *Cognitive psychology*, 51, 1–41.

Simon, H. A., & Lea, G. (1974). Problem solving and rule induction: a unified view. In L. W. Gregg (Ed.), *Knowledge and cognition*.

Vallée-Tourangeau, F., Austin, N. G., & Rankin, S. (1995). Inducing a rule in wason's 2-4-6 task: A test of the information-quantity and goal-complementarity hypotheses. *Thinking and reasoning*, 84A, 895–914.