Visual discovery in mind and on paper

RITA E. ANDERSON
Memorial University, St. John's, Newfoundland, Canada

and

TORE HELSTRUP
University of Bergen, Bergen, Norway

The effectiveness of mental imagery with and without drawing support (perceptual assistance) in the visual synthesis of novel patterns was studied in three experiments. When the task was to create one recognizable pattern from three simple shapes in a 2-min assembly period, subjects were as likely to produce a recognizable or creative pattern per trial whether mental imagery was augmented by external drawing support or not. When the task was to create as many patterns as possible in a 3-min assembly period, more patterns were produced per trial with external drawing support than without; however, neither the recognizability nor the creativity of the patterns differed. Differences in performance in the visual synthesis task with and without external drawing support are interpreted in terms of limited cognitive resources.

Can people use mental imagery to make genuine discoveries? According to many, the answer is yes and is documented by anecdotes such as the celebrated story of Kekule, who upon imaging a snake-like chain of atoms twist upon itself so as to bite its tail, realized that the structure of benzene must also be a closed chain or ring (Shepard, 1978). In fact, Shepard's review suggested that mental imagery was central to the creations of people in fields as diverse as physics, literature, and psychology. However, until recently, there has been virtually no scientific documentation that genuine discoveries can be made in imagery. In contrast, there is ample evidence that imagery may not be as effective as the anecdotal record might suggest. For example, it is more difficult to find a hidden part in a mental image of a complex figure than in a percept of the same figure (Reed & Johnsen, 1975). Perhaps more importantly, Chambers and Reisberg (1985) reported that although subjects could not "see" the alternative interpretation of classical ambiguous figures in their mental images of the figures, they could see it as soon as they drew their images on paper. Although subsequent research has shown that subjects can reinterpret mental images of classical ambiguous figures under some conditions, it is generally more difficult to reinterpret images than it is to reinterpret pictures of the same figures (Hyman, in press; Kaufmann & Helstrup, in press; Peterson, in press; Peterson, Kihlstrom, Rose, & Glisky, 1992).

Finke and Slayton (1988) developed a visual synthesis task to determine whether untrained subjects could mentally synthesize novel patterns. Their stimuli were drawn from a pool of 15 simple geometric shapes and alphanumeric figures, such as a line, a square, and the capital letters L, D, and X. On each trial, the subjects had 2 min to try to combine three randomly selected parts into one recognizable pattern by using only mental imagery. At the end of the assembly period, they wrote a brief description of the pattern and then drew it. Three judges independently rated the correspondence between the descriptions and drawings, and classified the recognizable or good correspondence patterns (i.e., those with a mean correspondence rating of 4.0 or higher, hereafter referred to as *good patterns*) as creative or not. Across two experiments involving at least 150 different combinations of parts and over 100 subjects, good patterns were produced on 38%–44% of the trials. Furthermore, approximately 16% of those good patterns were also judged to be creative.

But, how effective is mental imagery in the synthesis of novel patterns that are recognizable and, perhaps, creative? In the three experiments reported below, we include an external representation condition (drawing support) as a baseline against which to evaluate the effectiveness of the use of internal representations alone (mental imagery). Note that the internal–external manipulation compares the use of purely mental representations and processes with the use of mental representations and processes augmented by pencil-and-paper support. Experiment 1 was modeled

after Finke and Slayton's (1988) first experiment. The subjects had 2 min to produce one recognizable pattern on each trial, sometimes using mental imagery alone and at other times with drawing support. In Experiments 2 and 3, the subjects had 3 min to produce as many recognizable patterns on each trial as possible, again under internal and external representation conditions.

Our initial expectation, based on data and theory, was that the provision of drawing support would facilitate performance. That is, perceptual tasks are often completed faster, more accurately, or more completely than imagery tasks (Chambers & Reisberg, 1985; Finke, 1989; Roth & Kosslyn, 1988; Thompson & Klatzky, 1978; Weber & Harnish, 1974). We (Anderson & Helstrup, in press; Helstrup & Anderson, 1991) hypothesized that the provision of drawing support would provide the perceptual assistance needed to relieve the resource demands on the controlled processes involved in image generation and manipulation, as well as the demands involved in maintaining the image in a visual buffer (Farah, 1984; Kosslyn, 1980; Logie & Marchetti, 1991; see also Reisberg, 1987). These freed resources could then be used to combine the parts and interpret the combinations more effectively, thereby leading to more discoveries. On the other hand, just as ideas often flow faster than one can describe them, potential patterns may be assembled more quickly in the mind than on paper. In fact, Roskos-Ewoldsen (1989) reported that subjects took longer to generate patterns when drawing than when imaging. In addition, as reported by Finke (1990), Neblett, Finke, and Ginsburg found that subjects were as likely to create a pattern when they used mental imagery alone as when they could physically manipulate parts provided as transparencies. Because arguments for the facilitating effects of drawing support could be countered by some evidence to the contrary, it was not clear a priori whether the provision of drawing support would enhance performance or not. Hence, we performed the following experiments.

EXPERIMENT 1

We designed Experiment 1 to evaluate the efficacy of the use of mental imagery with and without drawing support in the synthesis of one recognizable novel pattern from three randomly selected simple shapes. To minimize the effects of individual differences in the ability to synthesize novel patterns and differences between stimulus triplets in the ease of pattern construction, all subjects were tested under both internal and external representation conditions on the same stimulus sets. Because the basic experimental procedures were modeled after those used by Finke and Slayton (1988), the subjects were expected to synthesize good patterns on 30%–40% of the mental imagery trials. The important unknown was whether performance would be facilitated or not when mental imagery was augmented by external pencil-and-paper support.

Method

Subjects. Twenty-one students (14 females and 7 males) in a course on human cognition at Memorial University participated in the experiment as part of a laboratory project. They were naive with respect to the issues raised by this experiment.

Design. Representation condition (internal and external) was a within-subject factor. The 10 experimental trials were blocked by condition and counterbalanced so that 10 randomly selected subjects used mental imagery alone on the first 5 trials, while 11 subjects used drawing support on the first 5 trials.

Materials. The pool of 15 parts and the procedure for selecting the 10 triplets were based on those used by Finke and Slayton (1988). For each triplet, three different parts were randomly chosen from the pool. The simple shapes (square, circle, triangle, rectangle, and horizontal and vertical lines) and capital letters (D, L, T, and C) were sampled three times as often as the more complex alphanumeric forms (J, X, V, P, and 8). Because only 10 different triplets were used in this experiment, we required that each part had to be selected at least once.

The test booklets consisted of a cover page, two response sheets for the internal and external representation conditions in counterbalanced order, and a short questionnaire. The response sheets were divided into five sections, corresponding to the blocks of five trials in each representation condition. Each section contained a blank line for writing out a brief description of a pattern and a space for drawing the pattern. The appropriate cue word, "Image" or "Percept," was typed to the right of each section.

Procedure. The subjects were tested as a group. They were told that the experiment was designed to find out how people create new figures using mental imagery alone or with pencil-and-paper support. The experimenter named the 15 basic parts on the blackboard and then described a typical trial as follows. As soon as the subjects heard the names of three randomly selected parts, they were to try to assemble those parts into a recognizable pattern, either by using mental imagery alone or with external support provided by drawing or doodling on a separate work sheet. After a 2-min assembly period, the subjects were first to write out a brief description of any pattern that they had created and then to draw it on the response sheet. All three parts had to be used in the pattern; the pattern could have been anything—letters, numbers, objects, familiar shapes, or symbols—as long as it could be recognized from the brief description. Although the size, position, and orientation of the parts could be varied, the shape of the parts had to remain the same.

The subjects received two practice trials, the first with mental imagery alone, and the second, with pencil-and-paper support. In the internal representation condition, the subjects were to close their eyes during the assembly period and try to assemble the parts mentally to form a recognizable pattern. In the external representation condition, the subjects were to externalize their construction efforts by drawing or doodling with a pencil on a sheet of paper during the assembly period. After each practice trial, the subjects were shown two patterns and their descriptions which could have been created from the parts for that trial. The sample patterns were chosen to illustrate the various ways of combining the parts. For example, the parts for the first practice trial were the capital letter L, a circle, and a square; the sample patterns were a TV and a flag. The subjects were reminded that all parts had to be used in the pattern; they were also told that because the parts had been randomly selected, any particular triplet might be more or less difficult to assemble into patterns, and hence that they should just try to do their best and enjoy themselves.

After the practice trials, the randomly ordered test booklets were passed out and the subjects found out whether they were using mental imagery alone or drawing support on the first block of five trials. After the first block of five trials, the instructions for the internal

and external representation conditions were briefly summarized and the subjects prepared themselves for the next five trials.

After the experimental trials, the subjects turned to a questionnaire based on the one reported by Finke and Slayton (1988). For each representation condition, the subjects had to choose which strategy they used most often from the following four options: (1) combining the parts by trial and error until a recognizable pattern occurred in the image or drawing; (2) thinking of a pattern and then trying to manipulate the parts mentally or physically so as to create it; (3) not using imagery or drawing at all, but thinking about how to combine the parts; and (4) using some other strategy.

Two safeguards were employed to minimize the potential for experimenter bias and demand characteristics. First, all the subjects were tested in one group and received the same instructions from the experimenter. Under these circumstances, it is difficult to imagine how the experimenter could simultaneously give different cues to different, randomly assigned subjects as to how each should perform on any trial. Second, the response protocols were scored blind with respect to representation condition, as described below.

Scoring conventions. Information about representation condition was eliminated from the response sheets, and the response sheets were arranged so that all patterns produced from the same triplet could be rated at the same time. Then three judges independently rated the correspondence between each description and drawing on a scale from 5 (*easy to identify from its description*) to 1 (*impossible to identify from its description*). A pattern that did not contain the three named parts was classified as a wrong-parts pattern and was not rated. After completing the correspondence ratings, the judges classified each pattern as creative or not.

The transformational complexity of each pattern was determined by counting the number of transformations of each part needed to create each pattern. Transformations included size (one or two parts smaller or larger than the third), rotation (of each part less than or greater than 90° from normal), mirror imaging (of P or J only), and embedding (of each part in one or two other parts). The maximum transformation score was 13.

Results and Discussion

Following the lead of Finke and Slayton (1988), all patterns receiving an average rating from the three judges of at least 4.0 were classified as good patterns and those receiving an average rating of less than 4.0 were classified as poor patterns. As can be seen in Table 1, good patterns were generated on as many internal as external representation trials. In the formal analysis, the number of trials in which each subject produced a good pattern in each condition was submitted to an analysis of vari-

ance for a simple Latin square design; the between-subject factor was group (internal or external representation trials first), while the within-subject factors were representation condition (internal and external) and test block (the first five and the last five trials). Only the effect of test block was significant [$F(1,19) = 5.77, MS_e = 1.05, p < .05$]; a good pattern was produced on more of the first five trials than of the last five trials (2.05 vs. 1.29, respectively), regardless of representation condition.

In fact, as can be seen in Table 1, the only difference between the two representation conditions was that subjects were more likely to fail to report a pattern on internal than on external representation trials [$F(1,19) = 5.30, MS_e = .29, p < .05$]. The meaning of this difference, however, was clouded by the fact that wrong-parts patterns were created on more drawing-support trials than mental-imagery-alone trials. It is possible that the subjects were especially sensitive to mental errors (e.g., forgetting a part) when using only mental imagery, and hence failed to report patterns whenever they lost track of the parts for that trial. Because the subjects could see their productions on the drawing support trials, they may have been less sensitive to mental lapses and have consequently committed wrong-parts errors instead of no-pattern errors.

To minimize potential problems associated with the somewhat arbitrary classification of a pattern as good or poor, the mean correspondence ratings for all patterns produced by each subject in each condition were subjected to an analysis of variance. The mean correspondence ratings of patterns produced under internal and external representation conditions did not differ significantly. Test block, however, was significant; mean correspondence ratings were higher for patterns produced in the first test block (3.69) than for those produced in the last test block (3.43); [$F(1,19) = 5.99, MS_e = .12, p < .05$]. The meaning of the significant effect of test block in this and in the previous analysis is unclear; performance may have been higher in the first than in the last test block, either because the triplets in the first block were easier to combine than those in the last block, or because the subjects worked harder on the first block of trials.

Analysis of the transformational complexity data revealed no differences in the complexity of the patterns

Table 1
Mean Performance Measures for Experiment 1

Measure	Representation Condition	
	Internal	External
Trials with pattern (max = 5)	4.29	4.52
Good pattern trials	1.62	1.71
Wrong-parts trials	.05	.19
*No-pattern trials	.67	.28
Correspondence rating (max = 5)	3.55	3.57
Transformation complexity (max = 13)	2.91	2.72
Creative pattern trials (max = 5)	.62	.67
Total subjects producing at least one creative pattern (max = 21)	10	12

*A significant difference at the .05 level or beyond.

produced under internal and external representation conditions. Nor did any other factor approach significance.

As in Finke and Slayton (1988), a creative pattern was defined as one that had been classified as creative by at least two of the three judges. In contrast with Finke and Slayton, all patterns were classified as creative or not. In fact, poor patterns were almost as likely to be classified as creative (16/135) as were good patterns (11/70). With more time and effort, those creative poor patterns might become more recognizable. Figure 1 shows a sample of creative and noncreative good and poor patterns produced under the two representation conditions from the stimulus triplet, PLC.

As can be seen in Table 1, representation condition did not affect either the number of creative pattern trials or the number of subjects producing a creative pattern. The subjects' performance was significantly associated in the two conditions; 9 of the 21 subjects produced at least one creative pattern in both representation conditions, while 8 failed to produce any creative patterns in either representation condition [$\chi^2(1) = 6.05, p < .05$].

Responses to the questionnaire did not differ across conditions; of the 21 subjects, 20 and 19 in the internal and external representation conditions, respectively, claimed to use the trial-and-error strategy for combining parts to synthesize patterns. The remaining subjects claimed that they tried to think of a shape first and then attempted to create it.

These results replicate Finke and Slayton's (1988) report that untrained subjects can make discoveries by using mental imagery alone. Furthermore, these untrained subjects were as likely to produce good or creative patterns when using mental imagery alone as they were when using pencil-and-paper support. Representation condition had

no effect on the quality (e.g., mean correspondence ratings, transformational complexity, judged creativity) of the patterns. Nor did the reported assembly strategies differ. The only difference was that the subjects were more likely to fail to produce any pattern when using only mental imagery than when being able to draw; the cause of this difference, however, was unclear.

EXPERIMENT 2

The focus of Experiment 1, and the experiments reported by Finke and Slayton (1988), was on the production of a single pattern per trial. The data from Experiment 1 show that subjects are equally likely to produce a recognizable or creative product per trial whether mental imagery is augmented by external support or not. But what about quantity? Will mental imagery be as effective for creating multiple patterns when it is used alone as it will when it is augmented by drawing support? Finke, Pinker, and Farah (1989) found that people were able to detect more patterns corresponding to symbolic forms when they inspected a drawing of the target pattern than they were when they imaged it. If, as suggested previously, subjects can commit more cognitive resources to planning, constructing, and interpreting patterns when using external support than when they must perform all aspects of the task mentally, drawing support may facilitate the production of multiple patterns. In Experiment 2, the subjects had to produce and describe as many patterns as they could during a 3-min assembly period with or without external drawing support.

We also examined performance in the first and second halves of Experiment 2 to determine whether the greater success on the first than on the last test block in Experi-
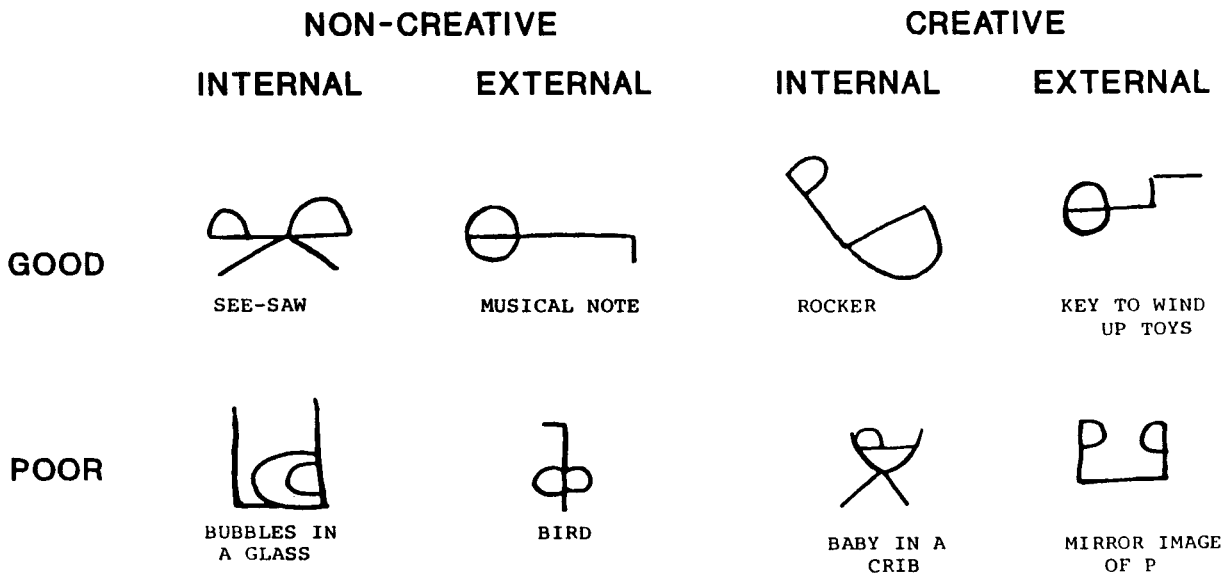

Figure 1. Examples of good and poor patterns produced in each representation condition in Experiment 1 from the stimulus triplet PLC that were classified as creative or not creative.

ment 1 reflected stimulus differences (easy vs. hard triplets) or differential effort across test blocks. Evidence summarized by Finke (1990) suggested that no practice effects would be found.

Method

Subjects. The subjects were 20 advanced psychology students at the University of Bergen (Bergen, Norway) who were naive with respect to the purposes of the experiment. The 16 females and 4 males ranged from 21 to 35 years of age.

Design. Representation condition (internal and external) and session half (first and second) were within-subject factors.

Stimuli. To eliminate the ambiguity in Experiment 1 caused by the instruction that parts could vary in orientation and the specification of a line as horizontal or vertical, "line" was retained as one part; the capital letter E was added to keep the total number of parts at 15. Forty triplets, selected according to the previously described procedures, were randomly assigned to five stimulus sets of eight triplets, with the following restrictions. Within each stimulus set, all parts were sampled at least once but no more than three times, and any repeated parts had to be equally likely to occur as the first, second, or third part of different triplets.

Four presentation orders were created for each stimulus set. The first order was generated by randomly assigning four triplets to the internal and four to the external representation conditions, with the restriction that the frequency of complex parts within the triplets was balanced across representation conditions. The triplets were then semirandomly ordered, so that two internal and two external representation trials occurred in each session half. Reversing the condition assignments and the order of triplet presentation led to the other three presentation orders for each stimulus set.

A response booklet containing two practice and eight experimental response sheets was customized for each of the 20 presentation orders. The upper right corner of each sheet contained the word "mental" or "physical" to cue the subject regarding the representation condition for that trial. The three parts for that trial were drawn at the upper left of the sheet, using a set of templates to ensure that each part was always drawn in the same way. The remainder of the page was divided into 10 numbered sections, each containing a blank line for writing the description and a space below for drawing the pattern.

Procedure. The procedure for Experiment 2 was similar to that for Experiment 1, with the following exceptions. To provide sufficient time to assemble and write out the descriptions of one or more patterns, the assembly period was increased from 2 to 3 min; after each assembly period, the subjects were given time to draw all the patterns they had described. Because each trial took longer to complete in Experiment 2 than in Experiment 1, the number of experimental trials was reduced from 10 to 8. Auditory presentation of the parts in Experiment 1 meant that the subjects had to remember the parts in the mental imagery alone trials, whereas they could see them in the drawing support trials. To eliminate faulty memory for the parts as the cause of differences in the number of wrong-parts patterns and no-pattern trials between the representation conditions, the three parts for each trial were presented visually and auditorily.

To minimize unintentional bias during testing, the subjects were tested in pairs whenever possible, so that on each trial, while one subject was using mental imagery alone, the other was drawing. The subjects were first shown an example of three parts and three different descriptions and patterns that had been created from those parts. After the experimenter briefly described the task and the pool of 15 parts and their names, the subjects read a page of instructions. These instructions were similar to those in Experiment 1, except that the subjects were instructed to try to produce as many patterns as they could during the 3-min assembly period of each

trial. In particular, the subjects were instructed to generate patterns, using mental imagery, with or without external support, until they came up with a recognizable pattern. At that time, they were to write a brief description of it on the response sheet; then they were to try to create another new pattern. At the end of the 3-min assembly period, the subjects were given time to draw the pattern below each description.

After they had read the instructions, the subjects were given two short (90-sec) trials in which to practice using mental imagery alone and pencil-and-paper support, respectively. After each practice trial, questions about the procedure were answered and the subjects were shown some patterns and descriptions created from the parts for that trial. Finally, the subjects were told that during each assembly period, while one of them would be using mental imagery alone, the other would be drawing.

Scoring. All rating decisions were made without knowledge of the production variables of subject, trial, and representation condition. Information about representation condition was eliminated from the response sheets. Then the response sheets from the 4 subjects who had received the same triplet were grouped so that all patterns produced from that triplet could be rated at the same time. Three judges independently rated the patterns from the 40 triplets on a 5-point scale (from *very easy* to *impossible to identify the pattern from its description*). Then all 152 variations on the same pattern description across triplets and subjects (e.g., the 23 different signs, 24 faces, or 13 umbrellas) were traced onto a single page and compared. Two judges ranked the patterns; the mean pattern rating was adjusted when necessary, to force a consistent ranking. Ratings for only three patterns (5%) were adjusted by more than 1 point on the 5-point scale, testifying to the reliability of the original ratings. Judges then classified each pattern as creative or not. Finally, the transformational complexity of each pattern was scored.

Results and Discussion

Inspection of the doodling sheets from the drawing support trials revealed that subjects produced over 500 complete or partial patterns. They described 172 of them on their response sheets; 83 were later rated as good patterns, and 26 were classified as creative. It is, of course, impossible to know how many patterns were generated during the mental-imagery-alone trials. However, the subjects described a total of 125 patterns in the mental imagery trials; 53 were rated as good patterns and 13 were classified as creative patterns. Across trials and conditions, the subjects described between 0 and 6 patterns per trial, with a maximum of 4 good patterns per trial. In all the analyses, the criterion for a good pattern was an average correspondence rating of at least 4.0.

The number of trials in which each subject produced at least one pattern, one good pattern, or no patterns was determined. Table 2 shows that the subjects were more likely to produce at least one good pattern on external than on internal representation trials [$t(19) = 2.63, p < .05$]. The subjects were also more likely to fail to produce any patterns on mental imagery than on drawing support trials [$t(19) = 2.52, p < .05$]. Although these data are not strictly comparable to those of Experiment 1 (because of task differences), the data replicate the apparent greater difficulty of producing any patterns by using only mental imagery that was found in Experiment 1, and they suggest that the provision of external support increases productivity.

Table 2
Mean Performance Measures for Experiment 2

Measure	Representation Condition	
	Internal	External
Trials with any pattern (max = 4)	3.15	3.65
*At least one good pattern	1.95	2.75
*No patterns	.85	.35
*Patterns/trial	1.56	2.15
*Good patterns/trial	.66	1.04
p(Good pattern|all patterns)	.44	.52
Correspondence rating (max =5)	3.56	3.71
Transformation complexity (max = 13)	2.98	3.33
*Creative patterns/trial	.16	.33
p(Creative pattern/all patterns)	.09	.16
No. subjects producing at least one creative pattern (max = 20)	9	15

*A significant difference at the .05 level or beyond.

In the analysis of productivity, the total number of patterns produced by each subject on the internal and external representation trials in each session half were subjected to an analysis of variance with representation condition and session half as within-subject variables. As is shown in Table 2, subjects produced significantly more patterns per trial under external than under internal representation conditions [$F(1,19) = 22.68, MS_e = 1.22, p < .01$]. Although slightly more patterns were produced in the second than in the first half, session half was not significant as a main effect or in interaction with representation condition.

In the analysis of the number of good patterns, the subjects produced more good patterns per trial with than without external drawing support [$F(1,19) = 9.83, MS_e = 1.14, p < .01$]; again, although slightly more good patterns were produced in the second than in the first half, neither session half nor the interaction with representation condition was significant.

The nonsignificant increases in performance across session halves in the preceding analyses suggest that the effect of test block in Experiment 1 reflected stimulus differences, not differences in effort. The fact that session half did not interact with representation condition in Experiment 2 also suggests that the drawing support and mental imagery alone tasks were of comparable difficulty, complexity, or novelty.

To compare the overall quality of the products of mental imagery with and without external support, additional analyses were performed on the conditional probability of producing a good pattern, the mean correspondence rating, and the mean transformational complexity of the patterns. Because session half was not significant in the previous analyses, these analyses were based on performance collapsed across all trials. In the first analysis, the number of good patterns, given the total number of patterns produced per session, was calculated for each subject. Although the probability of producing a good pattern was slightly higher for the external than for the internal representation condition, the difference failed to

achieve significance [$t(19) = 1.09, p > .10$]. In the second analysis, the mean correspondence ratings for all patterns produced under the two representation conditions were calculated for each subject. As in Experiment 1, the mean correspondence ratings for the drawing support and mental imagery alone conditions did not differ significantly [$t(19) = 1.79, p < .10$]. In the third analysis, the mean transformational complexity for all patterns produced by each subject under the two representation conditions was calculated. Again, there was no significant difference [$t(19) = 1.70, p > .10$].

Figure 2 shows a sample of the good and poor patterns that were identified as creative by at least two of the judges. The total numbers of creative patterns produced by each subject under each representation condition were compared, as were the conditional probabilities of creative patterns (e.g., the numbers of creative patterns, given the total numbers of patterns produced per session by each subject). Table 2 shows that twice as many creative patterns were produced with external support as without [$t(19) = 2.29, p < .05$]. When conditionalized by the total number of patterns, however, the difference between external and internal representation conditions was not significant [$t(19) = 1.64, p > .10$]. In contrast with the findings of Experiment 1, creative performance in one representation condition was not associated with creative performance in the other condition: 8 of the 20 subjects produced at least one creative pattern in each representation condition, while 4 failed to produce any creative patterns in either condition [$\chi^2(1) = .61$].

In summary, being able to draw led to greater productivity than did using mental imagery alone. The subjects produced more patterns, more good patterns, and more creative patterns when they were able to draw than when they had to perform all operations mentally. But being able to draw did not lead to higher quality patterns than did using mental imagery alone. Neither the probability of producing a good pattern nor a creative pattern, conditionalized on total productivity, differed significantly across repre-

EXPERIMENT 2

| INTERNAL | EXTERNAL |

EXPERIMENT 3

| INTERNAL | EXTERNAL |

GOOD

SLINGSHOT
VCJ

NUTCRACKER
OXV

"AND" SIGN
CJ\

GOLF
OP△

POOR

SLEEPING IN BED
□VO

INSECT
8X\

TELESCOPE
SURVEYING MOON
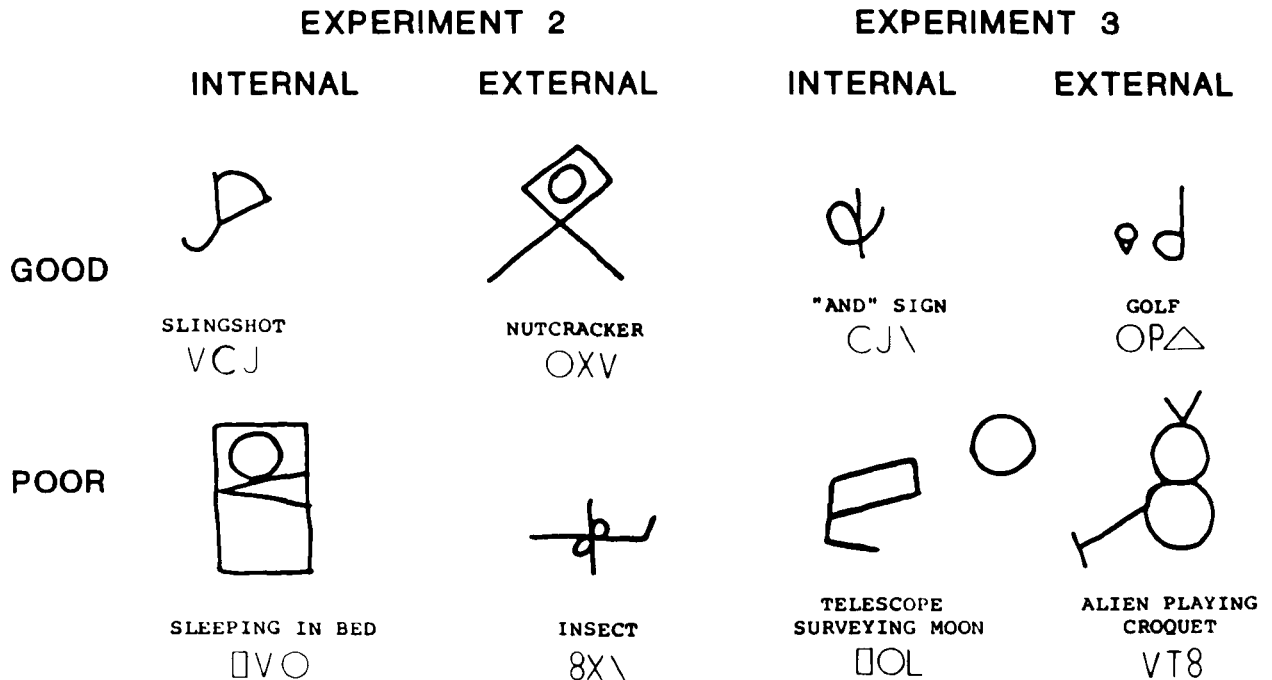□OL

ALIEN PLAYING
CROQUET
VT8

Figure 2. Examples of good and poor patterns from Experiments 2 and 3 that were classified as creative; the parts used to create each pattern are drawn below the description.

sentation conditions. Nor was there a difference between other measures of pattern quality, such as mean correspondence ratings or transformational complexity.

Taken together, the results of Experiments 1 and 2 suggest that mental imagery is remarkably effective for constructing and interpreting new, recognizable, and, sometimes, creative patterns from randomly selected sets of three simple shapes. These results were contrary to the initial belief, shared by most experimenters and subjects alike, that the use of pencil and paper to construct patterns should facilitate performance.

It could be argued that a belief in the efficacy of pencil-and-paper support might have influenced the performance of the subjects in these experiments, thereby leading to an underestimate of the effectiveness of the use of mental imagery alone. That is, because subjects were given complete information about both representation conditions prior to all experimental trials in Experiments 1 and 2, information about one representation condition could have influenced performance in the other condition in any of several ways. In particular, anticipation of difficulty or failure could have led to reduced effort on the mental imagery alone trials relative to drawing support trials. However, the results of two experiments patterned after Experiment 1, where representation condition was a between-subject factor (Helstrup & Anderson, 1991), were remarkably similar to those reported for Experiment 1 and suggest that expectations did not influence overt performance in Experiment 1. Nonetheless, it seemed reasonable to pursue the issue of expectations directly.

EXPERIMENT 3

In Experiment 3, we examined the effects of the subjects' expectations on overt performance in the visual discovery task and on their judgments of their own task performance. As in Experiment 2, the subjects were instructed to produce as many patterns as they could during a 3-min assembly period; as in Experiment 1, the experimental trials were blocked by representation condition. In contrast with the subjects in both Experiments 1 and 2, the subjects in Experiment 3 were not told about the alternate representation condition until they had completed the first block of trials. If expectations affect overt performance, significant differences in task performance should be observed between the first and second test blocks. To examine the effect of the subjects' expectations on their judgments of their own performance, we had them rate the correspondence between the descriptions and patterns that they had produced after completing all experimental trials. If subjects believe that drawing facilitates performance, their ratings should be higher for patterns produced under external rather than internal representation conditions.

Method

Subjects. The 20 subjects (10 males and 10 females) were summer session students at Memorial University who were paid for their participation.

Design, Stimuli, and Procedure. Representation condition (internal and external) was a within-subject variable. The six experimental trials were blocked by condition and counterbalanced so that 10 randomly selected subjects received the three internal

representation trials first, while 10 other subjects received the three external representation trials first. To enhance productivity and, hence, the probability of observing differences between representation conditions, the 30 experimental triplets were those from which subjects in Experiment 2 had successfully created patterns. As in Experiment 2 but not in Experiment 1, four presentation orders were created for each of the five different stimulus sets consisting of six randomly selected triplets to yield 20 different presentation orders. The response sheets and booklets were identical to those in Experiment 2, except that the second practice trial was inserted after the first block of three experimental trials.

The task was to produce as many recognizable patterns as possible during the 3-min assembly period. The main procedural difference between Experiments 2 and 3 was that in Experiment 3, the subjects were initially told only about one representation condition and were then given one practice trial and three experimental trials in that condition. After the first block of trials, the subjects were told about the alternate representation condition and were given one practice and three experimental trials in the alternate condition. The subjects were tested in groups of 1–5; all subjects in each test group were in the same representation condition on the same trials.

After the last experimental trial, the subjects were given a rating sheet and were asked to rate the correspondence between their verbal descriptions and their drawings on the 5-point scale (*very easy* to *impossible to identify*) used before. To illustrate each point on the scale, one to three patterns and their descriptions taken from Experiments 1 and 2 were provided on the rating sheet.

Three judges scored the patterns independently as in Experiment 2, except that the judges also used the rating sheet described above to rate the correspondence between each pattern and its description.

Results and Discussion

The subjects in the internal representation first group produced a total of 80 patterns (8 were rated as good patterns and 6 were classified as creative) when using mental imagery alone, and 240 doodles for a total of 93 patterns (18 good, 11 creative) when using pencil-and-paper support. The subjects in the external representation first group produced a total of 69 patterns (16 good, 4 creative) when using mental imagery alone, and 248 doodles for a total of 68 patterns (17 good, 9 creative) when working with

pencil and paper. Although the subjects in the internal representation first group produced more patterns in both representation conditions than did the subjects in the other group, the group factor did not approach significance in any analysis reported below.

The number of trials in which at least one pattern and at least one good pattern were produced was determined for each subject. These values were subjected to analyses of variance for a simple Latin square design. The between-subject factor was group (internal or external representation trials first) and the within-subject factors were representation condition (internal and external) and test block (the first and second sets of three experimental trials). Table 3 shows that the provision of external support led to slightly more trials in which any pattern or at least one good pattern was produced than when subjects had to do all the work mentally; however, no effects approached significance in either analysis.

Scores on each of the performance measures per trial listed in Table 3 were determined for each subject and were subjected to separate analyses of variance for the Latin square design with the addition of the within-subject factor of the test trial. Neither the group nor the test block factor approached significance as a main effect in any analysis. Test block was marginally significant in interaction with trials in one analysis; the number of patterns produced per trial increased across trials in the first test block and decreased across trials in the second test block [$F(2,72) = 3.43$, $MS_e = 1.55$, $p < .05$]. The subjects' expectations did not produce any obvious effects on overt task performance in this experiment.

As can be seen in Table 3, none of the measures of overt task performance were affected by representation condition; there were no significant differences in the mean number of patterns per trial, the mean number of good patterns per trial, the conditional probability of a good pattern, the mean correspondence rating, or the mean transformational complexity of the patterns. Although the selec-

Table 3
Mean Performance Measures for Experiment 3

Measure	Representation Condition	
	Internal	External
Trials with any pattern (max = 3)	2.75	3.00
At least one good pattern	.95	1.35
No patterns	.25	.00
Patterns/trial	2.48	2.68
Good patterns/trial	.40	.58
p(Good pattern|all patterns)	.17	.25
Correspondence rating (max = 5)	3.14	3.37
Transformation complexity (max = 13)	2.90	2.90
*Subject-correspondence rating (max = 5)	2.83	3.35
Subject-good patterns/trial	.80	1.12
*p(Subject-good pattern|all patterns)	.30	.46
Creative patterns/trial	.17	.33
p(Creative pattern|all patterns)	.06	.10
No. subjects producing at least one creative pattern (max = 20)	6	13

*A significant difference at the .05 level or beyond.

tion of "easy" triplets in Experiment 3 was intended to enhance potential differences in quantity or quality between the two representation conditions, hindsight suggests that it may have been a misguided effort. Easy triplets probably demanded less of the available cognitive resources and, hence, may have minimized performance differences between the two representation conditions.

Although representation condition did not affect overt performance or objective ratings of performance, significant differences were observed in the subjects' correspondence ratings of their own productions. Patterns produced with the aid of pencil and paper received significantly higher subject correspondence ratings than did those produced with mental imagery alone $[F(1,18) = 9.65, MS_e = .83, p < .01]$. Although slightly more subject-defined good patterns were produced with external support than without, the difference only approached significance $[F(1,18) = 3.94, MS_e = .76, p = .06]$; however, the probability of a subject-defined good pattern conditional on the number of patterns produced by each subject was significantly higher in the external than in the internal representation condition $[F(1,18) = 9.26, MS_e = .09, p < .01]$.

The agreement between the judges and the subjects in their ratings of each pattern as good or poor for patterns produced on internal and external representation trials was 70% and 67%, respectively. As might be expected, most failures to agree (78%) occurred when the judges rated the pattern as poor and the subject rated it as good.

Figure 2 shows some creative patterns produced by subjects in Experiment 3. Although slightly more creative patterns were produced with than without external support, representation condition did not approach significance in any analysis. More subjects produced at least one creative pattern with external pencil-and-paper support than without. Creative performance in one representation condition was not associated with creative performance in the other condition; 4 of the 20 subjects produced at least one creative pattern in each representation condition, and 5 failed to produce any creative patterns in either condition, $[\chi^2(1) = .17]$.

The subjects were less likely to produce a creative pattern on Trial 1 than on Trial 2 or 3, as is revealed by a significant effect of trials in the analysis of the number of creative patterns $[F(1,72) = 3.86, MS_e = .25, p < .05]$ and in the analysis of the conditional probability of producing a creative pattern $[F(1,72) = 3.94, MS_e = .03, p < .05]$. The conditional probability of producing creative patterns was .02, .12, and .10 for Trials 1, 2, and 3, respectively.

The results of Experiment 3 are consistent with those found in Experiments 1 and 2, suggesting that in a visual synthesis task, the use of mental imagery alone is quite effective for the production of new patterns that are recognizable and sometimes creative. In addition, they show that subjects believe that being able to draw facilitates performance. Interestingly, this belief does not seem to trans-

late into measurable performance differences. Neither the overall productivity nor the quality (correspondence and creativity) of the patterns as rated by external judges differed as a function of representation condition.

GENERAL DISCUSSION

Our initial motivation for the present experiments stemmed from observations that perceptual assistance enhances performance on some mental imagery tasks. For instance, subjects are much more successful at finding hidden parts in complex figures and at reinterpreting ambiguous figures when the figures are presented as pictures than when the figures are presented as mental images (Chambers & Reisberg, 1985; Reed & Johnsen, 1975). Consequently, we expected that the provision of external drawing support would enhance performance in the visual discovery task, by increasing either productivity or the quality of the patterns.

The results of Experiments 1-3 show that beginning and advanced university students from North America and Norway were remarkably adept at using mental imagery alone in a visual discovery task. When they were required to produce one pattern per trial, mental imagery was as productive when used alone as when augmented by drawing support (Experiment 1); when they were required to produce as many patterns as possible per trial, the provision of pencil-and-paper support sometimes led to greater productivity than did the use of mental imagery alone, but not to higher quality patterns (Experiments 2 and 3). Neither the mean correspondence rating nor the transformational complexity of patterns varied as a function of representation condition in any experiment.

Nor did the provision of drawing support enhance the production of creative patterns. In Finke and Slayton's (1988) experiments, 16% of the good patterns were classified as creative. In the present experiments, where all patterns were classified as creative or not, 47% of the creative patterns were poor patterns. Across all three experiments, with mental imagery alone, 16% of the good patterns and 10% of all patterns were classified as creative by at least two external judges; with drawing support, 20% of the good patterns and 14% of all patterns were classified as creative.

Before interpreting the results of these experiments, we should point out that although subjects clearly believed that patterns produced with external support were "better" than those produced by the use of mental imagery alone, overt measures of performance showed no effect of representation condition (Experiment 3). The failure of subjects' expectations to affect overt performance in Experiment 3 is consistent with the results of experiments requiring the production of only one pattern per trial reported by Helstrup and Anderson (1991), where representation condition was a between-subject factor, and the one reported by Anderson and Helstrup (in press), where representation condition and instructions were similar to those

of Experiment 3. Hence, it is unlikely that the tacit knowledge of the subjects or demand characteristics (cf. Intons-Peterson, 1983) can be used to explain the present results.

We hypothesized that drawing support would provide perceptual assistance to reduce the demands on limited cognitive resources. Consequently, subjects could allocate those freed resources to plan, construct, and interpret new patterns in the visual synthesis task. The fact that the provision of drawing support did not affect either productivity or quality in Experiment 1 suggests that the construction and maintenance processes involved in producing only one pattern per trial did not place undue demands on cognitive resources. However, when productivity was emphasized in the visual discovery task (Experiments 2 and 3), perceptually assisted discovery led to the production of more patterns, but not to better quality patterns, than did working with mental imagery alone, especially when difficult stimulus triplets were used (Experiment 2). Because subjects had to describe each new pattern before beginning to construct another one, the emphasis on productivity should not have led to an increase in the simple maintenance demands of the task (e.g., holding the parts and patterns in mind) beyond those required in Experiment 1. However, it is possible that one needs more resources to maintain an internal representation of a newly created pattern than an external representation of it while simultaneously trying to decompose it and recompose another pattern. Alternatively, the difference in productivity might suggest that the production of the first pattern somehow interferes with the production of subsequent patterns more when all work must be performed mentally than when external support is available. In fact, many subjects independently suggested that it was more difficult to decompose and recompose patterns when they relied solely on mental imagery than when they used pencil-and-paper support.

This perspective returns us to the research showing that it is more difficult to reconstrue or transform some figures in imagery than it is in perception (Chambers & Reisberg, 1985; Reed & Johnsen, 1975; Reisberg & Chambers, 1991; Rock, Wheeler, & Tudor, 1989). In fact, subjects can readily combine simple parts or lines mentally and "see" emergent patterns and components in their images (Finke, Pinker, & Farah, 1989; Intons-Peterson, 1989; Roskos-Ewoldsen, 1989). The subjects in those studies, and in the present experiments, created their patterns from familiar, well-defined parts prior to interpreting the patterns. In contrast, reconstrual difficulties seem to arise primarily when subjects are first presented with whole coherent patterns. This comparison suggests that successful reconstrual depends on facile identification of pattern components or knowledge of how to decompose the pattern (cf. Roskos-Ewoldsen, 1989). If the parts are familiar or readily identifiable within the whole, subjects will be able to compose and decompose patterns. If the parts are unfamiliar or if it is difficult to identify the component parts of the figure, demands on limited resources may exceed the supply and lead to a failure to decom-

pose or parse the pattern. This suggestion receives some support from Peterson et al. (1992). Subjects who formed their images from good parsings of the ambiguous snail-elephant stimulus were more likely to provide valid reversals of the ambiguous figure than were subjects who formed their images from poor parts.

Although differences in productivity may reflect additional demands on cognitive resources for the maintenance of newly generated images or the greater difficulty of reparsing newly generated images as opposed to percepts, the rate of pattern production need not affect the quality of the pattern. And, as summarized above, representation condition did not affect correspondence rating, transformational complexity, or the proportion of creative patterns. However, the provision of drawing support did seem to be associated with increased numbers of subjects producing at least one creative pattern in Experiments 2 and 3, where productivity was emphasized. In addition, although most subjects in Experiment 1 produced either at least one creative pattern or none in each of the representation conditions, this result was not found in Experiments 2 and 3. Whether these differences between Experiment 1 and Experiments 2 and 3 are a consequence of the production of high numbers of patterns, some of which are creative "by fluke," or whether they are due to some benefit of the use of external representations, cannot be determined at this time.

Should we be surprised at the effectiveness of mental imagery alone in the production of new, recognizable, and, sometimes, creative patterns? Yes—cognitive theory has been remarkably reluctant to grant many powers to the nonverbal mind. In retrospect, however, if we consider the possible functions of imagery, these results should not be surprising. A variety of mental processes and structures have been proposed to allow people to model current reality and to anticipate future realities (e.g., anticipatory images, Neisser, 1976; mental models, Johnson-Laird, 1983). If people can mentally rearrange items to survive in the real world, they should certainly be able to rearrange simple shapes mentally to create novel patterns.

In closing, we suggest that creative visual discovery most often involves an interaction between the use of internal and external representations and procedures. Although our subjects were often initially quite resistant to using mental imagery alone, they were also initially hesitant about doodling during the drawing support trials. Even after practice and encouragement, they often paused, perhaps to create a plan or a pattern in their minds, before beginning to draw. Mental imagery seemed to be the initial source of potential ideas and to be critically involved in the synthesis and discovery process. In contrast, drawing seemed to be especially useful for the subsequent production and refinement of patterns. Once a complete pattern had been drawn during the assembly period, our subjects sometimes resketched part(s) lightly in different locations or sizes on the original pattern or physically rotated the paper in order to view the pattern in different orientations. Some subjects commented that the patterns that they created by using men-

tal imagery alone did not always look as good as they had expected when they drew them on the response sheets.

Taken together, the results of these experiments suggest that mental imagery can be used effectively by many people to generate an initial idea, perhaps via top-down construction processes (cf. Roth & Kosslyn, 1988). This idea can then be modified and refined most effectively in drawing, in order to turn the task into a lower level perceptual task (cf. Finke et al., 1989). When pure mental imagery fails, perhaps because more cognitive resources are needed than are available, drawing can be used to reduce these demands to workable levels and to help generate additional potential ideas.

REFERENCES

Anderson, R. E., & Helstrup, T. (in press). Multiple perspectives on discovery and creativity in mind and on paper. In B. Roskos-Ewoldsen, M. J. Intons-Peterson, & R. E. Anderson (Eds.), *Imagery, creativity, and discovery: A cognitive approach* (pp. 223-253). Amsterdam: Elsevier.

Chambers, D., & Reisberg, D. (1985). Can mental images be ambiguous? *Journal of Experimental Psychology: Human Perception & Performance*, 11, 317-328.

Farah, M. J. (1984). The neurological basis of mental imagery: A componential analysis. *Cognition*, 18, 245-272.

Finke, R. A. (1989). *Principles of mental imagery*. Cambridge, MA: MIT Press.

Finke, R. [A.] (1990). *Creative imagery: Discoveries and inventions in visualization*. Hillsdale, NJ: Erlbaum.

Finke, R. A., Pinker, S., & Farah, M. J. (1989). Reinterpreting visual patterns in mental imagery. *Cognitive Science*, 13, 51-78.

Finke, R. A., & Slayton, K. (1988). Explorations of creative visual synthesis in mental imagery. *Memory & Cognition*, 16, 252-257.

Helstrup, T., & Anderson, R. E. (1991). Imagery in mental construction and decomposition tasks. In R.H. Logie & M. Denis (Eds.), *Mental images in human cognition* (pp. 229-240). Amsterdam: Elsevier.

Hyman, I. E. (in press). Imagery, reconstructive memory, and discovery. In B. Roskos-Ewoldsen, M. J. Intons-Peterson, & R. E. Anderson (Eds.), *Imagery, creativity, and discovery: A cognitive approach*. Amsterdam: Elsevier.

Intons-Peterson, M. (1983). Imagery paradigms: How vulnerable are they to experimenters' expectations? *Journal of Experimental Psychology: Human Perception & Performance*, 9, 394-412.

Intons-Peterson, M. (1989, November). *Appropriate and inappropriate priming before, during, and after image generation*. Paper presented at the meeting of the Psychonomic Society, Atlanta.

Johnson-Laird, P. N. (1983). *Mental models*. Cambridge, MA: Harvard University Press.

Kaufmann, G., & Helstrup, T. (in press). Mental imagery: Fixed or multiple meanings. In B. Roskos-Ewoldsen, M. J. Intons-Peterson, & R. E. Anderson (Eds.), *Imagery, creativity, and discovery: A cognitive approach*. Amsterdam: Elsevier.

Kosslyn, S. (1980). *Image and mind*. Cambridge, MA: Harvard University Press.

Logie, R. H., & Marchetti, C. (1991). Visuo-spatial working memory: Visual, spatial or central executive? In R. H. Logie & M. Denis (Eds.), *Mental images in human cognition* (pp. 105-115). Amsterdam: Elsevier.

Neisser, U. (1976). *Cognition and reality*. San Francisco: W. H. Freeman.

Peterson, M. A. (in press). The ambiguity of mental images: Insights regarding the structure of shape memory and its function in creativity. In B. Roskos-Ewoldsen, M. J. Intons-Peterson, & R. E. Anderson (Eds.), *Imagery, creativity, and discovery: A cognitive approach*. Amsterdam: Elsevier.

Peterson, M. A., Kihlstrom, J. F., Rose, P. M., & Glisky, M. L. (1992). Mental images can be ambiguous: Reconstruals and reference-frame reversals. *Memory & Cognition*, 20, 107-123.

Reed, S. K., & Johnsen, J. A. (1975). Detection of parts in patterns and images. *Memory & Cognition*, 3, 569-575.

Reisberg, D. (1987). External representations and the advantages of externalizing one's thoughts. *Proceedings of the Ninth Annual Conference of the Cognitive Science Society* (pp. 281-293). Hillsdale, NJ: Erlbaum.

Reisberg, D., & Chambers, D. (1991). Neither pictures nor propositions: What can we learn from a mental image? *Canadian Journal of Psychology*, 45, 288-302.

Rock, I., Wheeler, D., & Tudor, L. (1989). Can we imagine how objects look from other viewpoints? *Cognitive Psychology*, 21, 185-210.

Roskos-Ewoldsen, B. (1989). *Detecting emergent structures of imaginal patterns: The influence of imaginal and perceptual organization*. Unpublished doctoral dissertation, Indiana University.

Roth, J. D., & Kosslyn, S. M. (1988). Construction of the third dimension in mental imagery. *Cognitive Psychology*, 20, 344-361.

Shepard, R. N. (1978). The mental image. *American Psychologist*, 33, 125-137.

Thompson, A. L., & Klatzky, R. L. (1978). Studies of visual synthesis: Integration of fragments into forms. *Journal of Experimental Psychology: Human Perception & Performance*, 4, 244-263.

Weber, R. J., & Harnish, R. (1974). Visual imagery for words: The Hebb test. *Journal of Experimental Psychology*, 102, 409-414.