Creativity Research Journal
2002, Vol. 14, Nos. 3 & 4, 287–304

The Use of Environmental Clues During Incubation

Rebecca A. Dodds, Steven M. Smith, and Thomas B. Ward
Texas A&M University

ABSTRACT: Three experiments tested the prediction that incubation effects are caused by interactions between activation and environmental clues. Participants worked on 20 experimental problems and then were informed that they would have a second chance to work on the problems. Half were told they might see clues before returning to the problems and were instructed to try to use such clues. Participants then had an incubation period during which they generated words from the letters of test words. The test words were either semantically related to experimental problem answers, the actual answers, or unrelated words. Finally, all participants again tried to solve the experimental problems. Resolution, calculated as the number of items solved during the second trial that were not solved initially, was measured. Participants who saw answers during incubation resolved more items than those who saw related words. In Experiment 3, participants receiving no instructions did not differ across clue conditions, whereas instructed participants who saw answers resolved more problems than those who received related words. Participants in the instructed and unrelated condition performed significantly worse than those in the instructed and answer condition. Incubation effects occurred only when participants who were shown answers were also given instructions. No support was found for the theory that incubation effects are caused solely by environmental clues and activation.

Incubation is one of the least studied aspects of creative problem solving. Since Patrick conducted the first experiment in 1938, only 38 others have been completed. Many of these experimenters have relied on Wallas's (1926) conception of incubation as a stage of creativity. According to Wallas, work on a problem begins during the preparation phase but is discontinued when the solver becomes fixated or reaches an impasse. Wallas asserted that the solver then takes a break from active work and may experience improved ability to solve when returning to the problem later (an incubation effect). This use of the term *incubation* implies that incubation effects are due to an automatic, unconscious process.

Several hypotheses have been offered in an attempt to explain how incubation effects occur. One of the earliest theories, offered by Woodworth and Schlosberg (1954), is that the problem solver becomes mentally exhausted during the initial work on the problem. The incubation period allows the solver to rest, and incubation effects are due to returning to work on the problem when refreshed. Another theory is that during the incubation period, information stored during the preparation period is recombined subconsciously to form the answer (see Seifert, Meyer, Davidson, Patalano, & Yaniv, 1995, for a discussion). Smith (1995) proposed the idea that incubation effects are due to forgetting initially inappropriate responses. According to him, incubation effects result when a problem solver becomes fixated during preparation by focusing on an incorrect response. The incubation period allows time for the incorrect response to be forgotten, thereby making the correct response relatively more available.

Yaniv and Meyer's (1987) memory sensitization hypothesis gave a different explanation of how the incubation process might work. Information necessary to solve problems resides in memory and is partially activated during the initial attempt at problem solving. However, because the activation of the necessary

Correspondence and requests for reprints should be sent to Thomas B. Ward, Center for Creative Media, University of Alabama, 462 B Reese Phifer Bldg., University Blvd. and Colonial Dr., Box 870172, Tuscaloosa, AL 35487–0172. E-mail: ward@ccm.ua.edu

information does not reach threshold, the solver is not consciously aware of it or able to use it while working. The activation is slow to decay during the incubation period, and its presence is thought to sensitize the solver to chance encounters with stimuli in the environment that would be helpful in solving the problem. Such encounters raise the activation over the threshold and allow the problem to be solved.

This activation theory was extended by Seifert et al. (1995), who asserted that the inability to solve a problem initially results in special long-term memory traces called *failure indexes.* These traces are partial representations of the problem that are encoded and stored in terms of the characteristics that useful cues in the environment may have (Patalano, Seifert, & Hammond, 1993). During the incubation period, if helpful external stimuli are encountered and encoded, activation spreading from these stimuli will access the failure indexes, resulting in new attempts to solve by assimilating the new information into the problem representation, using it to restructure the representation, or both (Seifert et al., 1995).

A review of the experimental literature regarding incubation reveals only a handful of studies in which the use of environmental clues was manipulated. One of the earliest was done by Maier (1931). Participants were required to tie two cords together that were hanging from the ceiling. The cords were too far apart to allow participants to reach them simultaneously. The solution involved tying the pair of pliers in the room to one cord to form a pendulum. Then the participant had to swing the weighted cord and catch it while holding the other cord to tie them together. Many of the participants (39.3%) were able to solve the problem without any clues. The remaining participants were given a hint in the form of the experimenter brushing by one cord to set it swinging. With the hint, an additional 32% of participants were able to solve the problem. Despite the success of participants who received the clues during problem solving, this experiment gives no indication of whether or not clues encountered in the environment during incubation will result in incubation effects.

A few researchers have examined the effect of clues present in the environment during the incubation period. Driestadt (1969) had participants work on insight problems. During the incubation period, participants in the experimental group were taken to a room where pictures on the wall provided visual clues to the answers to the insight problems. A significant, positive effect of seeing the clues was observed. However, contrary to predictions based on the activation and failure indexes theory, there was no incubation effect. Olton and Johnson (1976) attempted to replicate Driestadt's (1969) results using the same method but found neither effects of the clues nor receiving an incubation period.

Browne and Cruse (1988) used one of the same insight problems as Driestadt (1969). However, instead of placing experimental participants in a room in which visual clues were provided, participants drew geometric shapes that were analogous to the answer. Browne and Cruse (1988) found that the drawing task aided in solving the problem but made no comparison of experimental and control participants to determine whether an incubation effect occurred. No conclusion can be made about whether the data support the activation and failure indexes theory without such comparison.

Dominowski and Jenrick (1972) examined the effect of clues given during the incubation period. Participants were asked to solve an insight problem that involved making two boards and a clamp into a hat rack. Before the incubation period, half of the experimental participants were given the hint that the solution involved the ceiling. Participants receiving hints were more likely to solve the problem than those who did not. However, only 91% of the participants in the incubation condition who received the hint solved the problem, whereas 100% of participants in the control condition who received the hint solved. The researchers did not test for differences between incubation and control participants. The activation and failure indexes model would predict that an incubation effect should have been observed for participants given both the clues and incubation period, but this was apparently not the case given the percentages reported.

Finally, both Mednick, Mednick, and Mednick (1964) and Dorfman (1990) used problems from Mednick's Remote Associates Test (RAT) for their experiments. Mednick et al. (1964) gave participants a series of RAT problems and then had participants solve simple analogies of the form "A is to B, as X is to _____," in which the answers to the analogies were the same words as the answers to five of the RAT problems that participants had previously attempted and failed to solve. Dorfman (1990) increased the number of clue words given over a series of trials with RAT problems.

In both experiments, a significant increase in performance was associated with receiving clues, but no effect of having an incubation period was observed.

In each of the experiments mentioned, with the exception of Olton and Johnson's (1976), receiving clues caused an increase in level of performance. However, no incubation effects were observed. This pattern is in direct contrast to predictions based on the activation and failure indexes theory. If the theory were correct, the experimenters would have observed significant, positive effects when participants returned to problem solving, due to having received the clues during the incubation period.

The experimental evidence to date leaves some doubt about whether activation and failure indexes are the mechanisms responsible for incubation effects. Although both Wallas's (1926) definition of incubation and the activation and failure indexes theory included the assumption that the problem solver is passive during incubation and the processes occur automatically, requiring no conscious attention or effort, this is not necessarily the case. Smith, Sifonis, and Tindell (1998), hypothesizing based on the set change theory, posited a more active process. Specifically, the set change theory states that incubation effects are due to a change in mental set. Such a change does not occur merely when clues are present in the environment during the incubation period (Smith, 1995). Smith et al. (1998) provided evidence that clues in the environment during the incubation period do not aid in subsequent problem solving unless they are intentionally used by the solver. In a series of three experiments, RAT items were presented as the critical problems to be solved, and participants completed lexical decision tasks during the incubation period. Some of the words during the decision task were strongly associated (related words) with the answers to the critical problems. For example, in the RAT problem *apple—house—family,* the correct answer is *tree* and the word *leaves* might be used as a related word. The presence of the related words seen during the incubation period had no effect on resolution of initially unsolved problems unless participants were instructed to attempt to use them in problem-solving attempts. Contrary to predictions based on the activation and failure indexes theory, the presence of clues alone was insufficient to cause incubation effects. Instructions to use the clues were necessary. The results of these experiments would not be inconsistent with the set change theory because clues in the environment did not cause a change in mental set.

This demonstration of the effectiveness of active processing during incubation parallels the more general problem-solving literature concerned with analogical transfer (e.g., Gick & Holyoak, 1980, 1983). In analogical transfer experiments, participants are presented with two problems that require similar solutions (usually convergence of force from several directions). Results show that, despite solving the first problem, participants do not spontaneously transfer the principle of the solution to the second problem. However, successful problem solving ensues when instructions are given directing participants to use the solution to the first problem as a basis for solving the second (see Gick & Holyoak, 1983). It is unclear whether the participants are successful after the instructions are given because the instructions tell them *how* to solve the problem or because they supply the answer to the new problem. In this instance, the instructions to solve might be likened to either the related words of Smith et al. (1998) or the instructions to use the related words, in that they may act as the point from which problem solving proceeds but are not, in themselves, the answers to the problems.

The previous experimental literature seems to suggest that the activation and failure indexes theory of incubation effects may be incorrect, but no critical experimental examination of the theory has been complete. The present investigation tested whether incubation is, in fact, a passive, automatic process caused solely by activation and the creation of failure indexes or requires the active use of environmental stimuli. To meet this goal, the effects of instructions to use environmental stimuli were separated from the effects of clues in the form of either related words or answers to problems. This allowed the separate effects of instructions and clues on incubation to be examined.

Experiment 1 had a between-subject 2×3 design with instructions and clues as the independent variables. First, participants worked on 20 experimental problems like those found in Mednick's RAT for 30 sec each. All participants were told that they would have a second chance to work on the experimental problems. Participants in the instruction condition were informed that they should pay attention, as hints to the experimental problems might be found in the tasks that followed. Next, all participants completed an insight problem and the Make-a-Word task in which they made 3 words from the letters of each of 20 words. The words participants saw during the Make-a-Word task were either answers to the experimental problems, words semantically related to the answers, or words unrelated to the answers. Finally, all participants received a second trial in which they had 30 sec to attempt to solve the 20 experimental problems. Performance was measured by counting the number of experimental problems subjects solved during the second trial that they failed to solve in the first trial. Experiments 2 and 3 had similar designs.

If the activation and failure indexes theory is correct, participants who are exposed to answers and related words during the incubation period should solve more experimental problems than those who see unrelated problems, because answers and related words provide the best fit with the failure indexes. This pattern of results should be the same despite instruction condition because, according to the theory, active use of environmental clues (whether fostered by instructions or not) is not required to produce incubation effects. If the activation and failure indexes theory is incorrect, the results are expected to replicate those of Smith et al. (1998), with an interaction of the type of clue with instructions to use clues. Specific predictions for Experiment 1 included the following:

1. According to the activation and failure indexes theory, there should be a main effect of clues such that participants receiving answers during the incubation period should resolve more items in the second trial than those who receive related words. Those receiving related words should resolve significantly more problems than those receiving unrelated words because the unrelated words do not fit at all with the failure indexes (see Figure 1).

2. If the activation and failure indexes theory is incorrect, an interaction of clue and instruction is expected. Participants receiving no instructions should resolve the same number of problems in all clue conditions. In the instruction condition, participants receiving answers should outperform those in other clue conditions. Those in the unrelated condition who get instructions are expected to perform more poorly than those in other conditions because cognitive resources are being expended to make use of the irrelevant information instead of solving the problems by other means (see Figure 2).

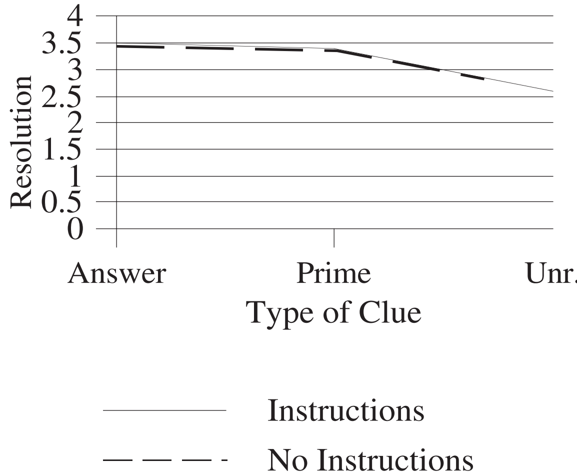

Figure 1. *Theoretical Results: Activation and Failure Indexes Theory Correct*

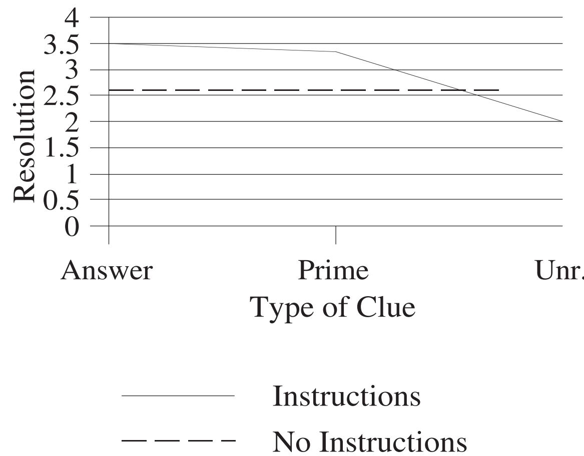

Figure 2. *Theoretical Results: Activation and Failure Indexes Theory Incorrect*

Pretesting

Participants

The participants in the pretesting procedures were 147 undergraduate students from the introductory psychology research pool at Texas A&M University who received credit for completion of class research requirements in exchange for their participation. Sheets for experimental sessions were posted on a bulletin board, and participants signed up for the sessions they wished to attend. The sign-up sheets gave no indication as to the experimental conditions participants would be attending.

Procedure. To compile the experimental problems, a group of 43 participants completed 69 original and revised problems from Mednick's RAT. From these, 20 experimental problems were chosen. The percentage of participants solving the 20 problems that were selected ranged from 31% to 71%. The mean solution rate for the 20 problems was 50% of participants (see Appendix A).

Related words of the problem answers were also determined during pretesting. Related words were words that were closely associated with the problem answers and were likely to make participants think of the answers. A separate group of 41 participants was presented with the 60 words from the experimental problems and their 20 answers. Words were given one at a time on PowerPoint slides, and participants were asked to write the first 15 words that they thought of in response to each. Participants were given 30 sec to write their associations to each word. Responses from participants to each experimental problem word and answer were ranked from most to least frequently given. The related words chosen were the words given most frequently by participants in response to the problem answers. Highly dominant responses that were adjectives, answers to or words from other experimental problems, or those strongly associated (1st to 15th most frequent response) with other answers were not chosen as related words.

To determine whether the related words would elicit answers to the experimental problems, another group of 23 participants was shown the related words on PowerPoint slides. They were asked to list the first 15 words they thought of in response to each of the related words and had 30 sec to write their responses to each. Responses were ranked for each related word from most to least frequently given. Answers to the corresponding experimental problems ranged from first to seventh most frequent responses given as associations for the related words. The average rank for answers given as associations to the related words across the 20 experimental problems was 2.55. Thus, it was likely that the words chosen would act as related words by eliciting the problem answers.

For a fixation condition used in Experiment 2, words chosen from the lists of responses generated as associated to the experimental problem words by the 41 participants during the second pretesting session were used. Association lists to the three words that comprise each problem were examined for a response that was commonly given for two of the words but not for the third or the answer to the problem. Responses for other problems and answers were also examined to make certain that the fixation words chosen did not appear in any other problem, as the answer to another problem, or in the association list for any other problem word or answer. The fixation words were those found only on the association response lists for the problem for which the fixation word was to be used.

To make certain that the fixation words actually had a negative effect on solving the experimental problems, another 40 participants were tested. Twenty were shown the experimental problems with the fixation words appearing below them in parentheses. The instructions for completing the experimental problems (see Appendix B) were shown with the additional sentence: "An example answer will appear in parentheses with each problem." In a separate experimental session, the remaining 20 participants saw the experimental problems with no fixation word or instructions. Next, all participants completed the first trial with the experimental problems, followed by work on an insight problem (see Appendix A) for 5 min. Participants then listed associations to 20 words from RAT problems not used in the experiment for 30 sec each. These words were not contained in any experimental problem; were not the answer to any experimental problem; and did not appear on the association lists for any experimental problem, answer, or related word. These associations were used for the Make-a-Word task (see next paragraph). Finally, participants were given a second trial with the experimental problems. The 20 nonfixation participants completed the same procedure but saw no fixation words or instruction with the experimental problems. The fixation group solved an average of 6.36 problems on the first trial, whereas the group who saw no fixation words solved an average of 9.49 problems. The fixation words and instructions significantly depressed performance, $t(38) = 2.88$, $p < .01$.

Unrelated words for the Make-a-Word task were chosen from the list of RAT problems not used in the experiment. The words did not appear in any experimental problem or as the answer or related word to any experimental problem. Additionally, these words did not appear in the association lists for any experimental problem, answer, or related word. The 40 participants who completed the fixation pretest gave associations for the 20 unrelated words between trials with the experimental problems. The ranked frequency lists of associations for unrelated words were examined to make certain that no experimental problem, answer, related word (ranging from 1st to 5th most popular association), or answer appeared. This was done to ensure that the words given in the unrelated category were actually not related to or associated with any other problems in the experiment.

Experiment 1

Method

Participants. The participants were 164 undergraduate students from the introductory psychology research pool at Texas A&M University who received credit for completion of class research requirements in exchange for their participation. Sheets for experimental sessions were posted on a bulletin board, and participants signed up for the sessions they wished to attend. The sign-up sheets gave no indication as to the experimental conditions participants would be attending.

Materials. Materials included the 20 experimental problems chosen in pretesting from either Mednick's RAT or revised versions of those problems, their answers, and related words (see Appendix A).

Procedure. Participants entered the experimental room and were given an informed consent form to read and sign. All participants signed the form, agreeing to continue the study.

In the first problem-solving trial, participants were given instructions about the experimental problems and how to complete them, along with an example problem (see Appendix B). Then each of the 20 experimental problems was presented for 30 sec. Participants recorded their answers on the sheets provided.

After the first trial, participants were told that, because there were many experimental problems and that they were difficult, the participants would be given another chance to work on the problems later in the session. The following paragraph was read to participants in the instruction condition:

> Some of the things that you will do before going back to the problem set may help you. There may be hints to the problem set answers in the tasks you'll be doing next. You should pay close attention and try to use the hints to solve the problems when you do the set later.

Then participants were given an insight problem and asked to work on it for 5 min (see Appendix A). Next, participants were given 10 min to complete the Make-a-Word task, in which they were asked to make 3 words from letters of each of the 20 words on a list (see Appendix B). They were given 30 sec to record answers for each word. The words in the Make-a-Word task were either the answers to the experimental problems, words semantically related to the answers, or words unrelated to the answers (see Appendix A).

In the second trial, the PowerPoint slides with the experimental problems were presented again for 30 sec each, and participants were asked to write answers for the problems. Participants were instructed to record answers, even if they had written them during the first trial. Finally, participants were given debriefing forms, and any questions the participants had were answered.

All of the instructions and stimuli for the sessions were shown to participants in a PowerPoint slide show in which the presentation times had been preprogrammed. The experimenter stopped the show for each instructional set to allow time for questions to be answered. Participants' responses were made in a packet of answer sheets provided by the experimenter at the beginning of the experimental sessions. Sheets were torn out of the packet and collected by the experimenter after each task was completed.

Results

The overall incubation effect was assessed using a paired t test. The number of problems each participant solved on the first trial was compared to the number solved on the second trial. The mean number of problems solved on the first trial was 9.90, with an average of 12.65 problems solved during the second trial. Participants solved significantly more experimental problems on the second trial, which demonstrates an effect of incubation, $t(164) = 16.72$, $p < .001$. The dependent variable for all subsequent analyses was a resolution score. Resolution was calculated as the number of problems solved during the second trial that were not solved during the first trial. This measure is in keeping with the theoretical definition of incubation in which problems not solved initially are resolved after a rest period. Resolution scores were computed for each participant.

Resolution scores were analyzed using a between-subject analysis of variance (ANOVA) with two levels of instruction and three levels of clues. The ANOVA approached significance, $F(5,158) = 2.17$, $p = .06$. The effects of instruction and the interaction of clues and instruction were not significant, but the effect of clues was, $F(2, 158) = 4.10$, $p < .05$. A follow-up t test showed that those in the answer condition had higher resolution scores than those in the related word condition, $t(110) = 2.92$, $p < .01$. Participants in the unrelated condition also scored significantly higher than those in the related word condition, $t(101) = 2.38$, $p < .05$. Performance in the answer and unrelated conditions did not differ (see Table 1).

To test the predictions made a priori, the effects of clues at each level of instruction were examined. When no instructions were given, there was no difference in performance among participants receiving answers, related words, or unrelated words. However, when participants received instructions, there was a significant effect of clues, $F(2, 80) = 3.24$, $p < .05$. A post hoc Tukey test shows that, when given instructions, participants in the answer condition significantly outperformed those in the related word condition. There was no difference between performance of participants in the related word and unrelated conditions or between those in the answer and unrelated conditions (see Table 2).

Table 1. *Experiment 1: Mean Resolution Scores in Clue Conditions*

Condition	*M*	*SD*	*N*
Answer	3.26^a	2.06	61
Related Word	2.25^b	1.59	51
Unrelated	3.17^a	2.28	52

Note: Means with different superscripts differ significantly at $p < .05$.

Table 2. *Experiment 1: Mean Resolution Scores in Clue × Instruction Conditions*

Condition	*M*	*SD*	*N*
Answer			
No Instruction	2.97	1.83	31
Instruction	3.57	2.25	30
Related Word			
No Instruction	2.38	1.58	24
Instruction	2.15	1.61	27
Unrelated			
No Instruction	3.46	2.21	26
Instruction	2.88	2.36	26

Discussion

Results of this experiment show little support for the passive activation and failure indexes theory of incubation. According to the theory, any helpful environmental clues encountered during incubation should be useful in problem solving. Those clues that establish the "best fit" (i.e., answers) with the stored representation should provide the most assistance. In Experiment 1, when answers to the experimental problems were inserted into the incubation period, participants did not resolve significantly more problems than those in the unrelated condition, in which the participants encountered no relevant information about the experimental items. As predicted based on the activation and failure indexes theory, participants receiving associated information (related words) had significantly lower resolution scores than those in the answer condition. Contrary to expectation based on the activation and failure indexes theory, however, they scored lower than those in the unrelated group. This pattern of results does not support the theory.

Additionally, some evidence contradicting the activation and failure indexes model was observed. If the activation and failure indexes model were correct, participants receiving answers would have outperformed those in the related word and unrelated conditions with or without instructions. The simple main effect of instruction showed that, with instructions, participants in the answer condition performed better than those in the related word condition. When no instructions were given, this pattern was not observed. There were no differences in the clue conditions. Thus, predictions based on the activation and failure indexes theory were not supported.

The results of Experiment 1 failed to provide a critical test of the activation and failure indexes model in one respect. Although performance of participants in the related word condition was less than that of participants in either the answer or unrelated condition, it is impossible, given the design of the experiment, to know whether their performance was depressed below a baseline level of incubation effects by the experimental manipulations or not. Further, there is no way to test whether those in the answer and unrelated conditions actually had elevated performance, because no baseline level had been determined. If the activation and failure indexes theory is incorrect, comparison across conditions to baseline when no instructions are given should show no significant incubation effects and no depression in performance below baseline.

The theoretical definition of incubation provides two options for establishing that incubation effects have occurred. First, a comparison can be made between those receiving an incubation period and those who work continuously on the problem. Alternately, the comparison of those with an incubation period who work on experimental tasks to those who work on unrelated tasks may be done. In the first case, comparison would show any changes in performance due to receiving the incubation period. In the second case, the comparison would show changes in performance due to activities during the incubation period.

If the activation and failure indexes theory is correct, people who work continuously should have the same level of performance as those who receive an incubation period in which they work on an unrelated task because incubation effects are caused by receiving useful information during the incubation period. The comparison of experimental participants who receive related words to a group with no incubation period made it possible to draw conclusions about whether the related words caused a depressed level of performance in Experiment 1. The comparison of participants in the answer and related-word conditions to a group who had an incubation period but did not do the Make-a-Word task allowed conclusions about whether seeing related information results in incubation effects or whether incubation occurs merely with the passage of time. If the activation and failure indexes theory is correct, those receiving answers and related words should have significantly better performance than those who receive time alone. The theory would also lead to the prediction that participants who see unrelated items in the Make-a-Word task should have similar performance as those with the same incubation period who perform other tasks. Two possible conclusions exist to explain the Experiment 1 result of participants who saw unrelated items but performed the same as those who received answers. First, the unrelated words may have been useful in problem solving despite the results of the pretesting. Alternately, the activation and failure indexes theory may be incorrect about the cause of incubation effects.

To make these determinations, two control conditions were added to Experiment 2. An immediate control condition allowed comparison of all experimental participants to those who had no incubation period, and a second, delayed control group received the same amount of incubation time (15 min) but worked on unrelated tasks.

The results of Experiment 1 do not support the activation and failure indexes model, but the interaction of clues and instruction that was predicted if the theory is incorrect was not seen either. Those in the unrelated clue condition did not perform significantly worse when given instructions, and the interaction was not significant. One possible explanation for these results is that only a small effect of instruction was present. In Experiment 1, the effect size of the instruction variable was .015, resulting in power to detect an effect of only .045. In an effort to strengthen this variable, the instructions used in Experiment 2 were altered. Instead of just being told about hints, participants were informed that they might see answers to the experimental problems during the incubation period.

A second explanation for the lack of support for either of the predicted outcomes may be the lack of need for an incubation period. Smith (1995) asserted that, if effects are to be observed, participants must be stuck on a problem before the incubation period. Seifert et al. (1995) discussed the role of fixation or reaching an impasse in problem solving as important because it allows failure indexes to be fully encoded. These indexes have more strength and stability when the thinker initially has failed to solve a problem. Seifert et al. (1995) found that participants who were allowed to reach impasse before incubation, relative to those who were interrupted, demonstrated better recall of the unsolved problems.

In Experiment 1, participants were given 30 sec to solve each problem, followed by an incubation period. Although failure to solve problems on Trial 1 may have been due to fixation, there are alternate explanations such as simply needing more time to work. To examine the effects of instructions and hints in situations that are similar to those outside the laboratory, it is necessary to ensure that participants are fixated before giving them the incubation period. Smith and Blankenship (1991) demonstrated that presenting distractors with RATs can induce that fixation. The most effective distractors are those that are associated with two of the words in the RAT but not the third. For instance, one experimental problem used was "sandwich, Canadian, golf." The answer to the problem is "club," which makes two-word phrases with each of the problem words. The fixation word *bacon* makes the phrases "bacon sandwich" and "Canadian bacon" but is not associated with the word *golf*. Because fixation may be necessary to produce incubation effects and using fixation words with RAT items has been demonstrated to induce fixation, a manipulation of this type was added to Experiment 2.

In Experiment 2, the general design of Experiment 1 was used with the additions discussed previously. This provided the opportunity to replicate the results, as well as to examine the theory more specifically. Predictions for Experiment 2 included the following:

1. The predictions from Experiment 1. Specifically, the main effect of clues was examined for the answer and the related word or unrelated pattern predicted by the activation and failure indexes theory. If the theory is incorrect, an interaction of clues and instruction should be observed.
2. Fixated participants should exhibit better performance than nonfixated participants. According to both Smith (1995) and Seifert et al. (1995), fixated participants should be able to use the incubation period more effectively to resolve problems.
3. Fixated participants should be able to better resolve experimental problems when supplied with related words than nonfixated participants. If Seifert et al. (1995) were correct, fixated participants should have stronger memory traces and thus need less activation than nonfixated participants. According to the activation and failure indexes theory, related words, as compared to answers, should provide less help in problem solving. This effect should be more detrimental to the efforts of nonfixated than fixated participants.

Experiment 2

Method

Participants. The participants were 444 undergraduate students from the introductory psychology research pool at Texas A&M University who received credit for completion of class research requirements in exchange for their participation. Sheets for experimental sessions were posted on a bulletin board, and participants signed up for the sessions they wished to attend. The sign-up sheets gave no indication as to the experimental conditions participants would be attending.

Materials. The materials were the same as those used in Experiment 1, with the addition of the fixation words (see Appendix A). These distractor words were associated with two of the words in the experimental problem but not the third or the answers (see Pretesting).

Procedure. The procedure was identical to that in Experiment 1, except that participants in the fixation condition saw distractors in parentheses with the experimental problems in the first trial. Participants were instructed that these words were "example answers" (see Appendix A). Thus, Experiment 2 had a between-subject 2 (instructions) $\times$ 3 (clue) $\times$ 2 (fixation) design. Additionally, 2 control conditions were included. In the immediate control condition, 81 of the 444 participants completed the first trial with the experimental problems and then had the second trial with no intervening incubation period. The performance of this group demonstrates the absolute baseline level of performance with continuous work on the experimental problems in 2 trials. In the delayed control condition, 50 of the 444 participants completed the first trial and an insight problem. Then they worked on a drawing task (see Appendix B) for 10 min before completing the second trial with the experimental problems. The resolution scores of participants in this group demonstrate baseline performance with a 15-min incubation period when the intervening tasks are unrelated to the experimental problems. Finally, an addition was made to the instructions such that participants in the instruction condition were told:

> Some of the things that you will do before going back to the problem set may help you. There may be hints to the problem set answers, *or even the answers themselves,* in the tasks you'll be doing next. You should pay close attention and try to use the hints to solve the problems when you do the set later.

Results

When using a fixation manipulation, the dependent variable must take into account the number of problems that remain to be solved, as this is not constant across the fixation and nonfixation conditions. If the manipulation is successful, those in the fixation condition will get stuck and solve fewer problems during the first trial, thus artificially inflating the resolution scores. Therefore, the dependent variable for all analyses was the proportion of problems successfully resolved. This score was derived by first figuring the resolution score (the number of problems solved during the second trial that were not solved during the first trial). Next, the proportion of problems resolved was derived by dividing the resolution score by the number of problems not solved after the first trial. The resulting number represents the proportion of problems that were unsolved in the first trial that the participant successfully solved in the second trial. Because the proportional measure adjusts for number of problems left to solve in the second trial, it is an appropriate measure to use with a fixation manipulation.

Recall that in Experiment 1 participants in the related-word condition scored lower than those in either of the other clue conditions, but it was not possible to conclude that related words significantly decreased the level of performance because no control conditions were included in the experiment. In Experiment 2, incubation was assessed in two ways. First, a one-way, between-subject ANOVA was done comparing control participants from the immediate and 15-min delay groups to the experimental participants. The ANOVA was not significant, $F(2, 441) = .82$, $p = .44$. Second, to evaluate the prediction of the activation and failure indexes theory that clues in the environment alone result in incubation effects, participants who received instructions and fixation were excluded from the analysis, and the remaining experimental participants who were exposed only to the variable of clues were compared to those in the control conditions using a one-way, between-subject ANOVA. The result was not significant, $F(2, 199) = .26$, $p = .77$. Post hoc t tests to examine a priori predictions were performed to compare each clue condition to each control condition, but yielded no significant differences. No incubation effects were observed in this experiment (see Table 3).

A between-subject ANOVA with two levels of instruction, three levels of clues, and two levels of fixation was performed, and the results approached significance, $F(11, 301) = 2.17$, $p = .09$. Neither the main effects of instruction or fixation, nor the interaction of instruction and clues were significant. The main effect of clues was significant, $F(2, 301) = 5.73$, $p < .01$. A Tukey post hoc test revealed that participants in the related-word condition had significantly lower proportions of resolution than those in either the answer or unrelated conditions, but the latter two groups did not differ from each other (see Table 4). This replicates the pattern of results observed in Experiment 1.

Because no main effect of fixation was observed in the three-way ANOVA (see Table 5), the ANOVAs were computed again without this variable. Use of proportional scores to account for differing levels of success in the first trial was not necessary because no fixation effect was observed. When resolution scores were used as the dependent variable, the results of the ANOVA approached significance, $F(5, 307) = 1.82$,

Table 3. *Experiment 2: Mean Proportion of Resolution in Control Conditions and Clue Conditions With No Instructions and No Fixation Manipulations*

Condition	*M*	*SD*	*N*
Answer	.34	.20	25
Related Word	.25	.23	25
Unrelated	.30	.19	21
Immediate Control	.28	.20	81
Delayed Control	.28	.18	50

Table 4. *Experiment 2: Mean Proportion of Resolution and Resolution Scores in Clue Conditions*

Condition	Mean Proportion	*SD*	*N*	Mean Resolution
Answer	$.34^{a}$	.22	107	3.40
Related Word	$.25^{b}$	.19	106	2.68
Unrelated	$.33^{a}$	.22	100	3.25

Note: Means with different superscripts differ significantly at $p < .05$.

p = .11. There was a significant main effect of clues [$F(2, 307) = 3.35$, $p < .05$], and Tukey post hoc tests showed that participants in the answer condition scored significantly higher than those in the related-word condition. But the performance of participants in the answer and related-word conditions did not differ from those in the unrelated condition (see Table 6). There was no significant effect of instruction or interaction of instruction and clues.

To test predictions regarding the interaction of the type of clue and instructions, the simple main effects of instruction were examined. The results showed that participants did not differ significantly across clue conditions with or without instructions. There was no evidence that instructions also significantly increased scores within any of the clue conditions (see Table 7).

Discussion

The fixation manipulation did not alter the pattern of means in either the clues or instruction conditions (see Tables 8 and 9). Power to detect differences due to fixation was 1.00. It is apparent that, contrary to the predictions based on previous hypotheses, fixation was not sufficient to cause an incubation effect and did not affect the other variables being used in this experiment.

The results of Experiment 2 almost perfectly replicated those of Experiment 1. Again no support was found for the activation and failure indexes model. The subjects who saw related words had lower proportional resolution scores than those who saw answers. However, participants receiving unrelated information did not differ from those in the answer condition. According to the activation and failure indexes model, unrelated information should not aid in problem solving, and participants receiving such problems should perform at a lower level than those receiving useful information such as related words and answers. More importantly, participants were unable to use any type

Table 5. *Experiment 2: Mean Proportion of Resolution in Fixation Conditions*

Condition	*M*	*SD*	*N*
No Fixation	.30	.21	156
Fixation	.30	.21	157

Table 6. *Experiment 2: Mean Resolution Scores in Clue Conditions*

Condition	*M*	*SD*	*N*
Answer	3.40[a]	2.31	107
Related Word	2.68[b]	2.02	106
Unrelated	3.25[ab]	2.10	100

Note: Means with different superscripts differ significantly at $p < .05$.

Table 7. *Experiment 2: Mean Resolution Scores in Clue × Instruction Conditions*

Condition	*M*	*SD*	*N*
Answer			
No Instruction	3.14	1.97	51
Instruction	3.64	2.57	56
Related Word			
No Instruction	2.47	1.85	49
Instruction	2.86	2.16	57
Unrelated			
No Instruction	3.17	1.86	45
Instruction	3.31	2.29	55

Table 8. *Experiment 2: Mean Proportion of Resolution in Fixation × Clue Conditions*

Condition	*M*	*SD*	*N*
Answer			
No Fixation	.33	.20	56
Fixation	.34	.24	51
Related Word			
No Fixation	.25	.20	51
Fixation	.24	.18	55
Unrelated			
No Fixation	.32	.23	49
Fixation	.33	.22	51

Table 9. *Experiment 2: Mean Proportion of Resolution Fixation × Instruction Conditions*

Condition	*M*	*SD*	*N*
No Instruction			
No Fixation	.30	.21	71
Fixation	.30	.21	74
Instruction			
No Fixation	.31	.22	85
Fixation	.31	.23	83

of clue—including answers—to boost their performance above the baseline level displayed by the control groups. Incubation effects should have been observed in both the answer and related word conditions if the activation and failure indexes model is correct. This result helps to more firmly establish the idea that, contrary to the predictions of the activation and failure indexes model, clues in the environment during incubation are of little help in problem solving.

It was somewhat disappointing that no strong effect of instruction was observed in either Experiment 1 or Experiment 2. The changes in the instructions for Experiment 2 succeeded in increasing the effect size from .015 to .03, with the resulting power increasing from .045 to .08. But without a stronger effect, any interaction of clues and instruction that might exist would not be detected, and a fair test of the activation and failure indexes model is impossible. Therefore, in Experiment 3, the experimental procedure was radically revised in an effort to further increase the effectiveness of the instructions. There are several possible reasons for the lack of a strong effect of instructions. First, participants might have forgotten the instructions by the time they encountered the Make-a-Word tasks in which the clues were embedded. Therefore, the intervening insight problem was not used in Experiment 3. Second, participants may have been confused about which of the tasks in the incubation period (insight problem or Make-a-Word task) contained the clues to the experimental problems. Therefore, in addition to discarding the insight problem, the experimental instructions were given after those for the Make-a-Word task to make it clear to participants exactly where clues might be found.

Third, Yaniv and Meyer (1987) provided evidence that activation could last as long as 30 min, but it is possible that activation could be of shorter duration within the context of this experimental design. Therefore, the activation might have decayed before participants encountered the helpful clues. To eliminate this as a possibility, the time participants were given to respond to experimental and Make-a-Word problems was shortened to 15 sec each. This makes the total time between seeing the experimental problem and answer or related word in the Make-a-Word task approximately 5 min.

Finally, only the delayed control condition was used in Experiment 3. There were two reasons for this change. First, no difference in performance was observed between the immediate and delayed control conditions in Experiment 2, lending support for the idea that incubation effects are caused solely by the passage of time. Second, participants in both the delayed control and experimental conditions receive the same amount of time during the incubation period. Thus, any differences observed between performances in the groups were due to different activities during the incubation period, which is the primary prediction of the activation and failure indexes theory. If the activation and failure indexes theory is correct, participants receiving related words and answers should significantly outperform those in both the unrelated clues and delayed control conditions.

The design of Experiment 3 was the same as that in Experiment 1 with two levels of instruction and three levels of clues. The prediction for Experiment 3 was as follows:

> If the activation and failure indexes theory is correct, a significant main effect of clues should be observed such that those in the answer condition have the best performance, followed by those in the related-word condition. Participants in the unrelated condition should have the worst performance. If the activation and failure indexes theory is incorrect, an interaction of clues and instruction should be observed with those in the answer and instruction condition performing at the highest level and those in the unrelated and instruction condition performing at the lowest level. Participants given no instructions should not differ across clue conditions.

Experiment 3

Method

Participants. The participants were 247 undergraduate students from the introductory psychology research pool at Texas A&M University who received credit for completion of class research requirements in exchange for their participation. Sheets for experimental sessions were posted on a bulletin board, and participants signed up for the sessions they wished to attend. The sign-up sheets gave no indication as to the experimental conditions participants would be attending.

Materials. The materials were the same as those used in Experiment 1.

Procedure. The procedure was identical to that in Experiment 1, with minor exceptions. First, participants received only 15 sec to solve each experimental problem in both trials. Second, participants did not receive the insight problem. Instead, they proceeded directly from Trial 1 with the experimental problems to the Make-a-Word task, in which they received 15 sec for each problem. Task instructions for the experimental problems and Make-a-Word task were altered to reflect the new time period of 15 sec per problem. Because the experimental manipulation containing the clues directly followed the first trial, the instructions were altered slightly. Participants were informed that they would have a second problem-solving task and then return to the experimental problems. Additionally, they were told that hints or answers to the experimental problems would be in the next task they performed. Because of the shorter experimental period and more explicit instructions, it was expected that a main effect of instruction would be observed in Experiment 3. Third, the experimental instructions were given after those for the Make-a-Word task instead of before the incubation period.

There were 37 participants in the delayed-control condition. These participants had 15 sec per problem in both trials, with 5 min spent on the insight problem between trials.

Results

The dependent variable for all analyses was a resolution score—the number of problems solved in the second trial that a participant failed to solve in the first trial. The overall incubation effect was assessed using an independent subjects t test. Participants in the control condition had a mean resolution score of 2.62 problems, whereas the mean of participants in the experimental conditions was 2.69 problems, $t(245) = .25$, $p = .80$. On the whole, the performance of the experimental participants did not exceed baseline levels; no overall incubation effect was observed.

To explore the effects of clues and instruction, a between-subject ANOVA was conducted with three levels of clues and two levels of instruction. The ANOVA was significant, $F(5, 204) = 3.17$, $p < .01$. The main effect of instruction was not significant, $F(1, 204) = .31$, $p = .58$. The instruction variable had an effect size of .08 and power of .21. The main effect of clues was significant, $F(2, 204) = 3.89$, $p < .05$. A post hoc Tukey test showed that participants in the answer condition resolved significantly more problems than those in the unrelated condition. Performance of participants in neither of these differed from those in the related-word condition (see Table 10).

The interaction of clues and instruction was significant, $F(2, 204) = 3.88$, $p < .05$ (see Table 11). One-way between-subject ANOVAs were performed to examine the interaction. For participants receiving no instructions, there was no difference in mean levels of resolution across clue conditions, $F(2, 105) = .07$, $p = .93$. Performance of participants who received instructions varied significantly depending on the type of clue received. A Tukey post hoc test revealed that participants in the answer condition outperformed those in the unrelated condition, but neither of these differed from the related-word condition.

Further analysis of the interaction showed that participants in the answer condition resolved more problems when given instructions than when no instructions were received, $t(73) = 1.99$, $p < .05$. Additionally, those in the answer condition who received instructions demonstrated an incubation effect, as they outperformed participants in the control condition, $t(72) = 2.15$, $p < .05$. In the unrelated clues condition,

Table 10. *Experiment 3: Mean Resolution Scores in Clue Conditions*

Condition	*M*	*SD*	*N*
Answer	3.05^{a}	1.78	75
Related Word	2.58^{ab}	1.44	72
Unrelated	2.37^{b}	1.22	63

Note: Means with different superscripts differ significantly at $p < .05$.

Table 11. *Experiment 3: Mean Resolution Scores in Clue × Instruction Conditions*

Condition	*M*	*SD*	*N*
Answer			
No Instruction	2.66	1.56	38
Instruction	3.46	1.91	37
Related Word			
No Instruction	2.56	1.39	39
Instruction	2.61	1.52	33
Unrelated			
No Instruction	2.68	1.19	31
Instruction	2.06	1.19	32

participants receiving instructions resolved fewer problems than those who received no instructions, $t(61) = 2.04$, $p < .05$. Those in the unrelated condition who received instructions were depressed below baseline, performing more poorly than those in the control condition. However, this result was not significant, $t(67) = 1.79$, $p = .08$. Only participants in the related-word condition did not differ across levels of instruction. Their mean performance at 2.58 was nearly equal to that of participants in the control condition, with a mean resolution score of 2.62.

Discussion

The presence of the significant interaction between clues and instruction in these data allows a complete examination of the predictions made based on the activation and failure indexes theory. According to the activation hypothesis, participants receiving answers perform better than those receiving related words, because the answers are a better fit for the failure indexes created during the initial work on the problem. This was not the case in either of the instruction conditions. Additionally, participants in the unrelated condition who received no instructions performed just as well as those in the answer and related-word conditions who received no instructions. This is counter to the predictions derived from the activation hypothesis, because unrelated material would not fit the failure indexes. Of greater importance, participants in none of the clue conditions exceeded the baseline level of performance when no instructions were given. If the activation hypothesis were correct, participants in the answer and related-word conditions should experience incubation effects because clues that fit the failure indexes were present in the environment during the incubation period.

Further, the significant interaction of instructions and clues is directly counter to the predictions of the activation and failure indexes theory. Proponents of the theory claim that only clues in the environment are necessary to cause incubation effects (Seifert et al., 1995). However, the significant interaction showed that participants told to use clues had increased performance when they encountered answers and decreased performance when attempting to use unrelated material to solve the experimental problems. In fact, the only incubation effects observed were when participants saw answers *and* were told to use them. Therefore, it cannot be environmental clues alone that cause incubation effects, as posited by the activation and failure indexes theory.

Conclusions

Overall, the results of all three experiments lend no support to the predictions of the activation and failure indexes model. No incubation effect was observed when helpful environmental clues were present during the incubation period. With the higher power of the instruction variable in Experiment 3, those in the answer condition showed increased performance, and those in the unrelated condition demonstrated decreased performance. This pattern of results is not consistent with the predictions of the activation and failure indexes hypothesis. Therefore, it must be concluded that, within the confines of the present experiments, incubation effects are not solely due to encountering useful clues in the environment during the incubation period.

The results of the present experiments are consistent with research concerning the effect of clues during the incubation period. Driestadt (1969), Browne and Cruse (1988), Dominowski and Jenrick (1972), Mednick et al. (1964), and Dorfman (1990) all found that receiving clues during incubation had a significant, positive effect on the ability to solve problems on returning to work. In each of these past studies, the clues provided were helpful. In each of the three present experiments, participants who saw answers were better able to resolve problems, although there was not always a significant difference when compared with those in other conditions. In the previous literature, experimenters who tested for the presence of incubation effects when clues were given found none (Dorfman, 1990; Driestadt, 1969; Mednick et al., 1964). Incubation effects were not seen in the present experiments when only clues were given. Both instructions and answers had to be present for incubation effects to be observed. The critical test of the activation and failure indexes theory provided by the present experiments was possible because the effects of clues and instruction were separated, allowing the predictions of the theory to be fully tested.

The activation and failure indexes theory is only one of a number found in the incubation literature (see Seifert et al., 1995), many of which focus on passive, automatic processes as the source of incubation effects. Whereas only the activation and failure indexes theory was systematically tested in the present experiments, these data lend little support to other theories that propose passive processes. For instance, recovery from mental exhaustion does not appear to be responsible for the incubation effects in the present experiments. If this theory were correct, incubation effects should have been constant across experimental groups because the incubation periods were the same. Instead, incubation effects were observed only when participants received clues and attempted to use them. However, it seems unlikely that 30 sec on a problem, or 10 min on a set of experimental problems, would cause debilitating mental fatigue, or that 15 min away from the problems would result in recovery from such fatigue, especially when the break was filled with other mentally demanding activities. Thus, although the fatigue recovery theory cannot be discarded altogether as a possible cause of incubation effects as a thorough test has yet to be made, it is doubtful that recovery would explain the incubation effects of the present experiments.

The response competition theory of incubation was not vigorously tested in the present experiments either. In Experiment 2, the fixation manipulation consisting of additional words associated with two of the three items in the experimental problem was used. This manipulation might be thought of as an intentional effort to induce response competition. The fixation manipulation did not significantly affect the measure of incubation used (proportion of items resolved). It is possible that the manipulation was not successful in inducing competition, that the proportion measure was not sensitive to the effects of forgetting competing responses, or that this mechanism does not cause incubation. Without more direct, controlled experimentation, the only conclusion that can be drawn is that forgetting competing responses does not appear to be responsible for the incubation effects observed in the present experiments.

There are two theories in the incubation literature that focus on more active processes and may explain the results of the present experiments. The conscious work hypothesis suggests that incubation effects are the result of additional conscious efforts to solve the problem, but these efforts are forgotten after the solver achieves the answer (Browne & Cruse, 1988; Seifert et al., 1995). If the instructions to use clues resulted in additional conscious efforts to solve the experimental problems, then the incubation effects may be due to this extra work. However, this explanation is somewhat suspect, as the effect size associated with the instruction variable was very small in each of the present experiments. Also, the conscious work theory cannot account for these results entirely, as it does not explain the interaction of the type of clue and instructions observed in Experiment 3.

In addition, the results of the present experiments are not inconsistent with the predictions of the set change theory, which states that no incubation effects should be observed unless one's mental set changes during the incubation period. Hints or answers in the environment would only be helpful if participants had a change of mental set that allowed the use of the clues. If the assumption is made that instructions to use environmental clues causes a change of mental set, then the set change theory would predict the pattern of results seen in Experiment 3, with the addition that those in the related-word condition should show elevated performance. In Experiment 3, participants in the related-word condition had a slight, nonsignificant increase in performance when instructions were given. This effect of instruction in the related-word condition may be because participants needed more time to work with the related words to solve the experimental problems, as compared to participants in the answer condition who had no need to expend further effort to make the clues they encountered into answers for the problems.

Of the theories of incubation effects considered here, the predictions based on the set change theory most closely match the results observed in the present experiments. However, because these experiments were specifically designed to test the activation and failure indexes theory, other theories cannot be validated or discarded. Further, it is impossible to know whether or not multiple processes described by several theories were responsible for the results of the present experiments. But spreading activation and the creation of failure indexes should no longer be considered a viable alternative when attempting to explain these results.

At this point it is fair to ask if incubation is the passive, automatic process that has been posited for years by previous researchers and theorists. Given the results of these experiments, it would appear not. Instead, something more active must be occurring to cause incubation effects. Just as instructions may serve to either tell the steps necessary to solve a problem or provide the answer, incubation may provide time for the problem-solving process to change or the answer to be discovered.

The research of the processes of incubation is a journey filled with small hesitant steps. There are theoretical paths that lead astray and some that circle endlessly. Experiments that reveal them are valuable because they serve as markers for those who follow. With more careful, systematic study that exposes both the helpful and misleading trails, an accurate map may be created that leads to a complete understanding of the incubation process.

References

Browne, B. A., & Cruse, D. F. (1988). The incubation effect: Illusion or illumination? *Human Performance, 1,* 177–185.

Dominowski, R., & Jenrick, R. (1972). Effects of hints and interpolated activity on solution of an insight problem. *Bulletin of the Psychonomic Society, 26,* 335–338.

Dorfman, J. (1990). *Metacognition and incubation effects in insight problem solving.* Unpublished doctoral dissertation, University of California, San Diego.

Driestadt, R. (1969). The use of analogies and incubation in obtaining insights in creative problem solving. *Journal of Psychology, 71,* 159–175.

Gick, M. L., & Holyoak, K. J. (1980). Analogical problem solving. *Cognitive Psychology, 12,* 306–355.

Gick, M. L., & Holyoak, K. J. (1983). Schema induction and analogical transfer. *Cognitive Psychology, 15,* 1–38.

Maier, N. R. F. (1931). Reasoning in humans II: The solution of a problem and its appearance in consciousness. *Journal of Comparative and Physiological Psychology, 12,* 181–194.

Mednick, M., Mednick, S., & Mednick, E. (1964). Incubation of creative performance and specific associative priming. *Journal of Abnormal and Social Psychology, 69,* 84– 88.

Olton, R. M., & Johnson, D. M. (1976). Mechanisms of incubation in creative problem solving. *American Journal of Psychology, 89,* 617–630.

Patalano, A. J., Seifert, C. M., & Hammond, K. J. (1993). Predictive encoding: Planning for opportunities. In *Proceedings of the Fifteenth Annual Conference of the Cognitive Science Society.* Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.

Patrick, C. (1938). Scientific thought. *The Journal of Psychology, 5,* 55–83.

Seifert, C. M., Meyer, D. E., Davidson, N., Patalano, A. L., & Yaniv, I. (1995). Demystification of cognitive insight: Opportunistic assimilation and the prepared-mind perspective. In R. J. Sternberg & J. E. Davidson (Eds.), *The nature of insight* (pp. 65–124). Cambridge, MA: MIT Press.

Smith, S. M. (1995). Creative cognition: Demystifying creativity. In C. Hedley, P. Antonacci, & M. Rabinowitz (Eds.), *Thinking and literacy: The mind at work* (pp. 31–46). Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.

Smith, S. M., & Blankenship, S. E. (1991). Incubation and the persistence of fixation in problem solving. *American Journal of Psychology, 104,* 61–87.

Smith, S. M., Sifonis, C. M., & Tindell, D. R. (1998). Hints do not evoke solutions via passive spreading activation. In M. Gernsbacher & S. Derry (Eds.), *Proceedings of the Twentieth Annual Conference of the Cognitive Science Society* (pp. 3–6). Mahwah, NJ: Lawrence Erlbaum Associates, Inc.

Wallas, G. (1926). *The art of thought.* New York: Harcourt, Brace.

Woodworth, R., & Schlosberg, H. (1954). *Experimental psychology* (2nd ed.). New York: Holt, Rinehart & Winston.

Yaniv, I., & Meyer, D. (1987). Activation and metacognition of inaccessible stored information: Potential bases for incubation effects in problem solving. *Journal of Experimental Psychology: Learning, Memory and Cognition, 13,* 187–205.

Appendix A: Experimental Materials

Experimental Problems	Answers	Related Words	Distractors	Proportion That Solved
board, magic, death	black	hole	marker	.47
painting, bowl, nail	finger	print	broken	.31
bald, screech, emblem	eagle	America	tires	.71
walker, main, sweeper	street	road	floor	.71
chocolate, fortune, tin	cookie	oatmeal	soldier	.61
widow, bite, monkey	spider	web	love	.42
catcher, license, hot	dog	cat	plate	.36
duster, bed, weight	feather	pillow	room	.63
peach, arm, stop	pit	fruit	signal	.36
ship, outer, crawl	space	star	tunnel	.45
worm, scotch, red	tape	cassette	robin	.72
river, note, blood	bank	money	suicide	.36
hearted, feet, bitter	cold	winter	tender	.40
gravy, show, tug	boat	lake	wishbone	.63
sandwich, Canadian, golf	club	caveman	bacon	.59
motion, poke, down	slow	drive	pin	.45
news, tiger, doll	paper	write	stuffed	.36
made, cuff, left	hand	palm	tailor	.72
stool, powder, ball	foot	toes	milk	.40
wood, liquor, luck	hard	rock	cabinet	.40

Unrelated Words

candle	butterfly
electric	stalk
cabbage	Saturday
there	surprise
slicker	window
staff	friend
shout	stiff
habit	stat
gnome	king
soap	Southern

Insight Problem

Place 10 blocks in 5 rows with 4 blocks in each row.

Appendix B: Task Instructions

Experimental Problems

Each of the puzzles that you will see on the screen will consist of three words. The solution to each puzzle is a word that is linked or associated with each of the three words. For example, if the words were *check, mark,* and *note,* the correct answer would be *book.* The solution word *book* is linked with each of the three words: *checkbook, bookmark, notebook.* For each puzzle, write your answer in the numbered space on the answer sheet. You will have 30 sec to complete each problem, and there will be 20 problems.

Insight Problem

The following is a timed task. Do not turn the page over until I say "go." You will have 5 min to complete the problem. If you finish before time is called, please sit quietly.

Make-a-Word instructions. A set of words will be shown on the screen. For each word, please write three words that can be made from the letters of the word shown. For example, if the word shown was *elephant,* you could make *ant, pane, and late.* Each word you write should contain only letters found on the screen. If a letter is found only once in the word on the screen, you may use it only once in each of the words you create. The words you make can be of any length. Please write only English words. There will be 20 words, and you will have 30 sec for each.

Drawing task. (Note: This task was used in the 15-min control condition in Experiment 2.)

Please try to imagine going to another planet somewhere in the galaxy. Imagine that this planet has many different species of plants and animals. One of the species is highly intelligent, and they use many different types of tools. Think about what their tools would be like and how they would use them. *Specifically, think about tools that do not require electricity, motors, computers, and so on. What we mean here is simple tools.* Please draw one of these tools in enough detail that someone would be able to identify it from your drawing. Depict as much as you can think of that would be relevant for identifying it. Divide the open space in half. On the top half draw a front view and on the bottom half draw a side view. *When you are finished with your front and side drawings,* please answer the questions and follow the instructions at the bottom of the page.

If you need more space to answer any questions, please continue on the back of this sheet. What is this tool? What is it used for? Please go back to your drawing and label the different parts of the tool. Also indicate what function each part serves.