A Review of Experimental Research on Incubation in Problem Solving and Creativity

Rebecca A. Dodds, Thomas B. Ward, Steven M. Smith

Texas A&M University

Nothing is so frustrating as a mental impasse. For example, tip-of-the-tongue experiences, those irritating failures to recall names of people one knows perfectly well, were described by William James as similar to the feeling of needing to sneeze. "Writer's block," another commonly described impasse, is universally feared by authors. Getting stuck when working on a problem or creative project is all too familiar. Which is the better remedy for resolving these mental impasses: proceeding with additional concentrated work, or taking a break?

Many authors have noted that after a period of time away from an unresolved problem, the answer, which was unseen at first, may become readily apparent. Even when brilliant insights do not leap fully developed into one's conscious mind, a key concept or an understanding of the process for achieving the solution may emerge. Wallas (1926) popularized the term "incubation" to refer to the time away from conscious problem solving. Since then, a number of experiments have explored the processes connected with incubation and the variables that may affect those processes. The only review of this body of empirical work prior to the present chapter was by Olton (1979), and was limited to a handful of studies.

Incubation is an often reported experience, but, like reports of ESP or alien abductions, the reality of the phenomenon can be assessed only after critical scientific scrutiny. In the present chapter, we provide a thorough review of the experimental literature on incubation in problem solving and creativity and identify some of the key aspects of the incubation situation that need to be examined. We first describe two general experimental paradigms that have been used in incubation studies, and then assess those studies in which the authors simply attempted to demonstrate incubation effects. In the remainder of the chapter, we divide the studies according to the use of variables that might logically be viewed as important determinants of incubation effects. These determinants include the length of the incubation period and preparation activities manipulated, the types of problems to be solved, the nature of the activities during the incubation period, the clues were presented during incubation, and individual differences among the problem solvers. A summary of the variables examined is provided in Table 1.

Experimental Methods

Incubation experiments are generally conducted using one of two methods. In the first, called interpolated activity, subjects in an incubation group work on a given problem for a predetermined period of time, are given an incubation period away from it, and then return to finish work on the problem. Their performance is generally contrasted to that of a control group that works continuously on the same problem. The contrasts typically include such dependent variables as whether or not the problem was solved, the amount of time required to solve it, the originality of the solution, and how many solutions are given. The second method, the multiple trial/multiple item approach, generally includes several trials in which multiple problems are presented. Experimenters using this method may or may not insert an incubation period between trials. If an incubation period is not inserted, it is assumed that the time between the first and second encounter with a test item allows incubation to occur. Dependent variables for this method include the number of items solved by subjects in incubation and control groups or an increase in number of problems solved if a within-subjects design is used.

Starting with the first attempt to study incubation empirically (Patrick, 1938), only 39 experiments have been reported. Across those experiments, there have been 26 successful demonstrations of incubation effects, 10 reports of failures to find effects, and 3 accounts of finding incubation only for certain groups of subjects. With this roughly 75% overall success rate, the global impression from the experimental literature is that incubation effects represent a real phenomenon worthy of considerable additional study.

As shown in Table 1, the interpolated activity method is somewhat more common, having been used in 23 of the 39 experiments, with 16 reports of successfully demonstrating incubation. The remaining 16 experiments used the multiple trial/multiple item method, with 13 of those studies reporting incubation effects.

Demonstrating Incubation

In this section, experiments conducted to demonstrate incubation effects using certain experimental paradigms will be discussed. These are the most basic types of studies, in which an incubation period was used, but no other variables were manipulated.

In the first incubation experiment on record, Patrick (1938) asked subjects to propose scientific methods to investigate effects of heredity and the environment on humans. The control group worked on the problem immediately after it had been presented. They talked through possible solutions aloud until making a final proposal. The incubation group was given a diary and asked to return within two to three weeks with their final proposals. Incubation subjects were asked to write down any thoughts or ideas they wanted in the diary. Patrick (1938) operationalized incubation as an idea occurring early in the work process, reappearing several times and being the chief topic of the final solution. This definition departs from that used by most experimenters, but the study is included here because it purports to investigate incubation, and it is the historical starting point of the experimental literature on incubation.

Ninety-six percent of subjects in the study showed incubation. Additionally, more modification of ideas occurred for the diary group. Final proposals were rated for merit (feasibility and excellence) by a group of scientists. The proposals of the incubation group received higher scores overall than those of the non-incubation group. Although Patrick interpreted the effects as being due to incubation, the source of the advantage shown by the incubation group is uncertain. For instance, subjects in the diary group were never directed not to work on the problem during the incubation period. Thus, the higher ratings of their proposals may be due to the extra time that they were given to work rather than to taking time to lay the problem aside periodically. It is unknown how long the continuous subjects actually worked before providing solutions, and incubation subjects had between two and three weeks, an exceptionally long time frame for experimental research, during which the participants might have engaged in a wide range of related or unrelated activities. Thus, even though the results are suggestive, it is important to bolster them with findings from more tightly controlled procedures.

In a more recent study, Olton (1979) presented a challenging chess problem to subjects who were chess players between beginner and expert status in chess. Olton used a much shorter time frame than that employed by Patrick (1938). He theorized that, because the problem was of interest to them, subjects would be motivated to solve it, approximating real life conditions. The subjects worked individually at their own pace. Control subjects worked continuously whereas subjects in the incubation group were instructed to take a two-hour break sometime during their work and were allowed to do as they wished during the break. Incubation subjects often returned early and many subjects remained after the experiment ended to discuss the problem and solution. No evidence of incubation was detected, as half of the incubation subjects and half of the control subjects solved the problem. Olton concluded that there was no evidence of incubation. Although the time frame was shorter than in the Patrick (1938) study, the lack of control over when the participants stopped and returned also raises some concerns about the reliability of conclusions drawn from this study.

Kaplan (1989) performed several experiments in an effort to demonstrate the presence of incubation effects using his experimental method. Subjects in the control conditions worked continuously for four minutes on "Consequences" items. These items require subjects to generate consequences to questions such as, "What would happen if everyone suddenly lost the ability to read and write?" Subjects in incubation conditions worked for two minutes and had a 30-minute incubation period before returning to work for another two minutes. In the first experiment, the incubation activity consisted of completing a battery of psychological tests, whereas subjects in the second and third experiments listened to a psychology lecture. Fluency (the total number of ideas generated by subjects) was measured in each experiment.

Results of the first experiment showed that incubation subjects produced more solutions than control subjects during the second two minutes of work. The difference in number of ideas generated in the first and second two minutes of work was computed for the second experiment. Those in the incubation group produced the same number of ideas in each session, whereas those in the control group showed a decrease in number of ideas in the second session. The control subjects in the third experiment also generated significantly fewer solutions during the second two minutes than the incubation subjects. The overall impression from Kaplan's research is that, even in tightly controlled situations, incubation effects can be observed, thus adding credence to the early findings from less controlled experiments.

Preparation Activities

Beyond a simple demonstration of incubation effects, it is essential to begin to understand some of the factors that determine the likelihood of incubation effects occurring. One such factor is the preparation activities before the incubation period. Does it matter, for example, how long people work on a problem, or what they do prior to taking a break?

Silveira (1971, Exp. 1) was concerned with the length of the preparation time before incubation. Silveira used the chain problem, in which a man with four chains of three links each wants the chains joined into a single, closed length. Having a link opened costs $.02 and having a link closed costs $.03. Subjects are asked to tell how the man had the chains joined in a closed circle for only $.15. In the first experiment, subjects worked for a total of 35 minutes. They talked through their solutions aloud while an experimenter unobtrusively took notes. Sessions were also tape recorded. The start of the working session was not designated as being when the subjects received the problem. Instead it was when the subjects demonstrated an understanding that joining the chains in a circle would normally cost $.20 and that putting the chains in a straight line does not constitute the closed chain required by the problem. Because of individual differences, the amount of time it took subjects to reach these realizations varied, so the total amount of time spent working on the problem also varied.

Experimental subjects were interrupted either 3 or 13 minutes after starting. After the incubation period of either 30 minutes or 3.5 hours, subjects returned to work for either 32 or 22 minutes. Control subjects worked continuously for 35 minutes. In the first experiment, a comparison of subjects who solved in the incubation (61%) and control (47%) conditions revealed no significant difference, despite a slight increase in performance due to the incubation period. There was a significant effect of the timing of the interruption such that those interrupted later in the session were more likely to solve than those interrupted early. (See the section "Length of Incubation" for other results from Silveira [1971].)

In her second experiment, Silveira attempted to maximize the results of the first. Subjects worked for a total of 35 minutes on the chain problem. All incubation subjects were interrupted 13 minutes after the start of the work session for a 3.5 hour incubation period, and returned to work for 22 minutes. In this experiment, 38% of the control subjects and 81% of the incubation subjects solved. Clearly the incubation subjects were more likely to solve. Silveira combined the data of the first and second experiments and found no difference between performance of subjects in the control and early interruption conditions, but elevated performance for subjects interrupted after 13 minutes of work. These results provide a hint that incubation effects may be more apparent when the person solving the problem has had more time to become fixated prior to taking a break from the problem.

Kaplan (1989) examined length of preparation as well. Recall that Kaplan performed three experiments in an effort to demonstrate that incubation effects could be revealed by having subjects work on Consequences items for 2 minutes before breaking for an incubation period and returning for 2 minutes (see "Demonstrating Incubation"). In his fourth experiment, Kaplan used a slightly different method to test the hypothesis that increasing the ratio of time in preparation to time during the incubation period results in better performance. All subjects worked for a total of 50 minutes--10 on the Consequences problem and 10 on each of four distractor items. The distractors were a long division problem and three insight problems. Control subjects worked on the consequences item first and then each of the distractors in turn. Incubation subjects were given two trials. On the first trial, subjects chose to work for between two and eight minutes on each of the five problems. Once all five problems had been addressed for up to eight minutes, the second trial began. Subjects were allowed to choose the order of the problems for the second trial and work until the total time spent on each problem equaled 10 minutes.

Kaplan found a significant difference between the number of solutions generated by incubation and control subjects to the Consequences item. Incubation subjects produced a mean of 20.4 ideas. In contrast, control subjects produced only 15.7 ideas on average. There was also a significant difference in the rate of response such that incubation subjects produced ideas more quickly in the last eight minutes. There was no difference between incubation and control subjects in the number of ideas produced or rate of response in the first two minutes of work. Kaplan interpreted these findings as support for his hypothesis.

So, the relative amount of preparation time may matter, but what about the specific preparation activities? How might particular types of prior activities influence problem solving and incubation effects? Problems solved before a critical item may constitute part of the preparation for that item. Kaplan (1989; Exp. 5) used performance on multiple items to examine the effect of incubation on einstellung behavior. Einstellung behavior occurs when a person uses an approach that has been successful previously with a new set of problems that can be better solved with another method. It can be considered as a form of fixation. Kaplan used water jug problems in which subjects are asked how to measure a certain amount of water when given three jars of predetermined sizes. Subjects first work control problems that can all be solved using the same formula. Then they are exposed to the experimental problems. Although the experimental problems can be solved using the same formula as the control problems, a simpler formula also exists.

Kaplan gave his subjects nine water jug problems to solve. Control subjects worked for a total of 5 minutes on each problem. Continuous work subjects brainstormed for 2 minutes about the best ways to solve the problem, had 30 seconds to pick the best of 15 solution methods presented by the experimenter and then 2 minutes to solve. This pattern was repeated for all of the problems. Finally they completed the Remote Associates Test. Conceived by Mednick (1962), the Remote Associates Test (RAT) provides subjects with three words and requires them to find a fourth that is associated with each of the others. For example, given the words "cake", "blue" and "cottage", the correct answer would be "cheese."

Incubation subjects followed the same procedure as continuous work subjects except that they completed the RAT after the five control problems and before the four experimental problems. The incubation subjects were less likely than those in other conditions to use the formula from the control problems to solve the experimental problems. Interestingly, the continuous work subjects who brainstormed before working on the problems were less likely to demonstrate einstellung behavior than the control subjects. Thus, brainstorming during preparation resulted in less fixation.

Frankel (1990) and Parrish (1981) each explored the use of imagery training before incubation. Frankel (1990) prepared his subjects by having them imagine and describe various scenes ("eidetic imagery training"). Then subjects were asked to work for 20 minutes to provide solutions to life-relevant problems such as how to lower abuse of alcohol by teenagers. Control subjects worked for 20 minutes on the problem and then completed a set of anagrams. Incubation subjects worked for 10 minutes, completed anagrams, and returned to work for a final 10 minutes. Frankel used performance during the last 10 minutes of work to derive two scores for each subject: originality and fluency. To determine the originality scores, the solutions generated during the last 10 minutes of work were placed on a frequency distribution. Most frequent solutions scored no points whereas solutions offered less frequently received up to 3 points. A composite originality score for each subject was used in the analysis. The fluency score was simply the number of solutions during the last 10 minutes. Those in the incubation condition had significantly higher originality scores than those in the control condition, but the groups did not differ in fluency. The imagery training had no significant effect on either fluency or originality of solutions offered and did not interact with any other variables. In this experiment, the type of activity was not critically important.

Parrish (1981) manipulated two variables: imagery training during preparation, and presence of an incubation period consisting of 20 minutes of relaxation. The relaxation was assumed to help foster images. Incubation was operationalized as a period when no conscious or voluntary thought was directed toward the preparation (imagery training). It should be noted that the experimental treatments were initiated before giving subjects the criterion treatment and the test was only administered once, unlike most other studies in which experimental subjects are exposed to a problem, given an incubation period, and then return to work on the problem. Thus, this experiment cannot be classified as using either the interpolated activity or multiple trial/multiple item method.

Parrish used the Onomatopoeia and Images subtest of the Thinking Creatively with Sounds and Words test battery (Khatena & Torrance, 1973) to measure incubation effects. During the test, subjects produce verbal images to onomatopoeic words (i.e., crunch, tingle, etc.). No other incubation researchers have used this instrument. There are no correct answers for this test; instead, the images produced by subjects are scored for frequency of occurrence. Control subjects for each condition were given filler activities so that the amount of time spent in the experimental setting was the same for all subjects. There was no significant main effect of the incubation activity; however, there was an interaction between preparation and incubation such that, when given the preparation activity, subjects in the incubation condition scored lower than subjects in the control condition. With no preparation activity, subjects in the incubation condition wrote less frequent words than those in the control condition.

In summary, experiments examining the effects of various preparation activities on incubation have been helpful. First, Silveira (1971) demonstrated that longer preparation periods are more likely to result in increased performance after incubation. Kaplan (1989) showed that higher ratios of preparation to incubation times are beneficial. Results of his final experiment revealed that one advantage of preparation is to break up mental sets formed during work on previous, similar problems. Imagery training during preparation resulted in decreased performance for Parrish's (1981) subjects and had no effect in Frankel's (1990) experiment. Much work remains to be done to show whether other forms of preparation are more likely to influence incubation, and to provide a clearer picture of the role of prior activities in the incubation process.

Types of Problem

Another important question about incubation is whether the effects depend on the type of problem to be solved or are more general in scope, cutting across many problem types. This question is of theoretical significance for understanding the nature of incubation, and of practical significance for advising problem solvers about the types of problems most likely to be resolved through incubation.

Although many types of problems have been used in the incubation studies examined in this chapter, most individual experiments use only a single type. This situation clouds the issue of the effects of problem type because cross-experiment comparisons are fraught with peril. However, some evidence can be obtained from within experiments in which problem type was explicitly manipulated.

In addition to looking for incubation effects, Brockett's (1985) experiment examined differences in types of experimental problem. Subjects worked either on Remote Associate Test (RAT) items or listed unusual uses for a brick. Brockett's control subjects worked for a total of 20 minutes on the RAT or brick items and then completed a battery of psychological tests. Incubation subjects worked for 10 minutes and stopped to complete the battery before the final 10 minutes of work on the problems.

Peterson (1974) used the solution of anagrams as his experimental task. He was interested in whether the frequency of daily use of the anagram words would have any effect. Twelve sets of anagrams were constructed, each containing 12 anagrams apiece. Control subjects worked for 2 minutes on each of 12 anagrams, whereas incubation subjects worked for 20 seconds on each of the problems for a total of 6 trials. Incubation subjects solved problems more quickly than control subjects.

<u>Clues During Incubation</u>

Historically, many anecdotal accounts of successful problem resolution include a reference to clues in the environment, such as Archimedes' noticing of water sloshing out of his bath. Consequently, it is of some interest to try to experimentally assess the role of clues during the incubation period.

Driestadt (1969) was concerned with whether visual analogies (clues) to problem solutions introduced during incubation would have a significant effect. He used two insight problems to examine the use of visual clues during incubation. The first (farm problem) asked subjects to divide an irregularly shaped parcel of land into four equal pieces. The second required subjects to place ten trees into five rows of four trees apiece (star shape). Subjects were given paper and pencils to sketch their solutions. These problems were selected because they were tricky, not generally known by subjects, and not likely to be solved quickly.

Control subjects worked for a total of 20 minutes on one of the two problems. The incubation group worked for five minutes, moved to another room for eight minutes of undirected activity, and then returned to work in the original room on the either the farm or tree problem for the remaining seven minutes. In addition to incubation, visual clues were manipulated. For control and incubation subjects in the clue condition, there were pictures on the walls of the experiment room that incorporated shapes analogous to those that would solve the problem (see Driestadt, 1969, for illustrations).

How close the subjects' final sketches were to the solution of the problem was measured using a scale from one to five points. Subjects were awarded from five points (sketches of the solution) to one point depending on how close the sketches were to the solution. Separate analyses were done for each problem using these "weighted" solution scores. There was no effect of the incubation period. Strong effects of being exposed to the pictorial analogy were seen in both analyses. There was an interaction of analogy by condition in which more solutions and partial solutions were produced by the incubation group who had access to the visual analogies in the farm problem. Driestadt noted that, despite contact with the analogies, most of the subjects were either unaware of the analogies or only vaguely aware that they had helped.

Olton and Johnson (1976) attempted to replicate Driestadt's (1969) results by using visual analogies in their experiment. They used the farm problem, and incubation groups received 10 minutes of work, followed by 15 minutes of incubation and another 20 minutes of work. Recall that 8 minutes of the 20 minute session for Driestadt's incubation subjects was spent in a filler task whereas control subjects worked on the experimental problem for the full 20 minutes. Olton and Johnson hypothesized that the lack of main effects for incubation in Driestadt (1969) might have been due to the control group having more time to work on the problem. Thus, like the incubation groups, the control groups in their study worked on the farm problem for 30 minutes. One control group worked continuously in the same room whereas another worked for 10 minutes, was moved to another room, and then continued working. One experimental group was exposed to visual analogies during the incubation period (see "Activity During the Incubation Period" for discussion of the other experimental groups). Olton and Johnson used Driestadt's method of weighting the solution scores. They found no significant differences between the two control groups so they were pooled. There was no significant difference between the pooled control groups and the treatment group that saw visual analogies during the incubation period.

Browne and Cruse (1988) also used the farm problem and visual clues. Subjects worked for a total of 20 minutes on the problem. Incubation subjects took a break after 10 minutes of work to draw geometric shapes (clues to the answer) for 5 minutes before finishing the farm problem. Significantly more subjects in the incubation than control condition solved the problem. In their second experiment, control subjects worked for 20 or 25 minutes. Incubation subjects followed the same 10-5-10 time pattern as in the first experiment. (See "Activities During the Incubation Period" for further discussion.) The two control groups did not differ from each other. The incubation group that drew shapes showed better performance than the 20-minute control group. There was no overall analysis of the effects of incubation. Despite this, Browne and Cruse concluded that incubation effects had been demonstrated.

Dominowski and Jenrick (1972) also explored problem solving aided by interpolation and hints. In their first experiment, subjects worked on the hat rack problem first used by Maier (1945). For the hat rack, subjects are given two boards and a clamp and asked to construct a rack that could hold a hat or coat. The solution consists of wedging the two boards between the floor and ceiling and clamping them together. The clamp serves as the hook to hold the hat or coat.

Subjects were given a total of 15 minutes to work on the problem, and control subjects worked continuously. After five minutes, those in the hint condition were informed that the ceiling was part of the solution. After receiving a hint or not, incubation subjects were taken to another room for 10 minutes to work on either an anagram task or free association task, and then returned to work on the hat rack problem for the remaining 10 minutes. Subjects who solved the problem in the first 5 minutes of work were excluded from analysis.

An average of 21% of the subjects in the control condition solved the problem, whereas the mean solution rate was 26% in the incubation condition. Of the subjects in the control/hint condition, 00% solved, whereas 91% of the incubation/hint subjects solved. There was a significant difference in performance between subjects receiving the hint and those who did not. The authors concluded that there was an effect of incubation based on an interaction of incubation and ability level (see discussion in "Level of Ability"). However, there was no direct comparison of the number of subjects who solved the problem or solution times of those in the incubation and continuous work conditions, and no exploration of the possible interaction of incubation and hints. Thus, it is difficult to draw conclusions about the effect of hints on incubation from these findings.

In their second experiment, Dominowski and Jenrick (1972) again used the hat rack problem, allowing a total of 13 minutes of work. After three minutes, the hint was given but it was changed to, "In the correct solution, the coat would hang from the clamp" (p. 336). During the 3 minute incubation period, subjects performed an anagram task. It is important to note that, of 62 subjects, only 4 actually solved the problem. As a result, Dominowski and Jenrick's analysis consisted of examining the experimental protocols for the solutions attempted by subjects. After the incubation period, 73% of subjects returned to the same solution they had been attempting before the interruption. Half of the solutions provided by the control subjects were the same as the ones immediately prior to them during any given time in the solution period. There was no significant difference between the patterns of incubation and control subjects. Using this as a basis, the authors concluded that, "There is obviously no support for the notion that interpolated activity facilitates the adoption of a new approach to the problem" (Dominowski & Jenrick, 1972, p. 337).

Mednick, Mednick and Mednick (1964) used the multiple trial/multiple item method to examine the effects of clues on incubation. Subjects were allowed to work for one minute apiece on 26 difficult and 5 easy Remote Associate Test items. Of the 26 difficult items, the experimenter selected 10 that had been missed by the subject. The answers to five of these were primed by having subjects complete simple analogies. Then the 10 missed items were re-administered for one minute apiece. More of the primed items were solved correctly in the second administration than those that had not been primed. Despite defining incubation in their introduction as a time interval away from the problem, Mednick et al. (1964) refer to the priming as the incubation manipulation in their procedures and conclude that priming induced incubation effects.

In their second experiment, Mednick et al. (1964) used a similar procedure. The priming variable was changed to be between-subjects, so the five missed items were primed only for subjects in the priming condition. Priming occurred either before or after a 24-hour incubation period. Subjects given primes were significantly more likely to solve than those not given primes. There was no difference in scores due to the incubation period.

Dorfman (1990; Exp. 3) gave subjects 39 problems that were similar to those in the RAT to test the effect of clues on incubation. Of these problems, 13 were used in the no-delay condition, 13 were used in the 5-minute delay condition, and 13 were used in the 15-minute delay condition. The problems were presented in blocks over seven trials. Subjects began with three clues for each problem, just as in the Remote Associates Test. In subsequent trials, one additional clue was given for unsolved problems until a total of six clues had been given for each problem (see Table 2 for experimental design).

Results showed no effect of delay. There was a main effect of clue such that increasing the number of clues resulted in an increase in solutions. There was also a significant interaction of delay by clue, such that those in the 5-minute delay condition solved more problems than those in the immediate condition when given the same number of clues. Dorfman interpreted these findings as evidence of incubation effects. However, there was a significant difference in number of problems solved in the first trial, which implies that the problems were not equivalent over conditions. Thus, it is difficult to draw conclusions based on these results.

In her fourth experiment, Dorfman used the same procedure of presenting additional clues over trials, this time in a between-subjects format. There was again a significant difference in the number of problems solved in the first trial showing that subjects in the groups were not equivalent at the beginning of the procedure. There was also no significant main effect of delay. Dorfman concluded that an incubation effect existed based on comparison of one incubation group with the control group. The experimental groups differed in length of time during the incubation period (see section on "Length of Incubation Period"). There was a significant effect of hints, such that increasing the number given resulted in an increase in number of problems solved.

Although effects are not universally obtained, the results of studies manipulating <u>helpful</u> clues seem to show that the presence of clues affects incubation performance. But, what happens when the clues are designed to put the problem solver on the wrong track? Fixation due to the use of <u>misleading</u> clues was induced in two sets of experiments done by Smith and Blankenship

(1989, 1991). Smith and Blankenship (1989) used performance on rebuses as their creative task. A rebus is a puzzle in which words are arranged to form pictures. Subjects are asked to give the common phrase associated with the puzzle. For example, "ankoolger" would be deciphered as "look back in anger." The rebuses were presented with either helpful or misleading clues.

In the first experiment, 15 rebuses were presented for 30 seconds each with helpful clues. Then another five were presented with misleading clues. Finally, the last rebus was presented again without a clue and subjects worked on it for one minute. For incubation subjects, either a music perception task or unfilled free time lasting 5 or 15 minutes was used as the distractor before the final problem was presented again. Subjects in the 15-minute incubation group were more likely to solve the critical rebus than those in either of the other groups, but they did not differ significantly from those in the 5-minute incubation group. There was no effect for the music perception versus free time activities (see "Activity During the Incubation Period" for further discussion of Smith and Blankenship [1989]). The control group subjects who worked continuously showed greater recall of the misleading clue on the critical item than either of the incubation groups.

The second experiment had a similar design, with the exception of the critical item being paired with a helpful clue. The incubation period of 5 or 15 minutes was again filled with a music perception task or undirected free time. The results were essentially the same. There was no difference due to the type of activity used during the incubation period. The 15-minute incubation group was more likely to solve the critical rebus than either of the other groups, which did not differ from each other. Subjects in the control group were again more likely to recall the misleading clue paired with the critical rebus than those in either of the other groups.

In the third experiment, the incubation groups worked on the block of 20 rebuses for 30 seconds each, then performed one of the distractors (10 or 15 minutes on other rebus problems, or 10 minutes of rebuses and 5 minutes of math problems, relaxation or the music perception task) and returned to work for 30 seconds on each, whereas the control group worked through the block with one minute on each rebus. Clues were given for each rebus only for the first 30 seconds of work, with 13 of the 22 clues being misleading. Smith and Blankenship (1989) calculated improvement scores as the number of problems solved during the last 30 seconds of work. Thus the maximum possible improvement score was 20. The control group again showed less improvement than any of the incubation groups, but improvement scores did not differ among the incubation groups.

Finally Smith and Blankenship (1989) replicated their third experiment. The only difference was that the distractor tasks used were work on math problems, reading stories or no task. The incubation period for all distractor tasks was five minutes. There was no difference between groups working on the story and math tasks. However, the story group significantly outperformed the control group.

Smith and Blankenship (1991; Exp. 1) demonstrated that fixation can be experimentally induced using misleading clues and will dissipate during incubation. Twenty Remote Associate Test (RAT) items were presented with or without an incorrect clue. Subjects were allowed 30 seconds per item. Then the items were presented again without the clue either immediately or after a 5-minute incubation period that involved reading a short story. An improvement score (number of items solved in the second presentation that were not solved during the first) was calculated for every subject. There was a benefit of incubation only for subjects in the fixation condition (words presented with incorrect clues). Those who were not fixated showed no more improvement with the incubation period than those without it.

Next, Smith and Blankenship (1991, Exp. 2) presented RAT items on a computer with the misleading clues announced by the computer voice. Subjects were given one minute to complete each of 10 RAT items. Control subjects then worked through the items again. Incubation subjects read a story presented on the computer screen for five minutes before returning to work on the RAT items. For one group of control subjects and one group of incubation subjects, a hint of the first two letters of the correct answer was provided during the first trial. The hint was given to all subjects during the second trial. There were significant main effects both for fixation and hints. Subjects who saw and heard the misleading clues were less likely to solve than those who did not. Subjects who saw the helpful hint during the first trial were more likely to solve than those who did not. The interaction of these variables was not significant.

In their third and fourth experiments, Smith and Blankenship (1991) investigated the mechanics of the fixation effect and how to most effectively induce fixation. Finally, Smith and Blankenship (1991; Exp. 5) returned to the role of fixation in incubation. Each of 12 RAT items was presented after three pairs of associates. Each of the pairs contained one of the clues to the RAT item with either a helpful or unrelated word. The pairs preceding half of the test items were helpful and those preceding the other half were unrelated. The test item was presented for

30 seconds and, if unsolved, retested immediately, or after 30 seconds or 2 minutes of incubation. Four of the items were retested immediately, four after 30 seconds of incubation and four after 2 minutes of incubation. During the incubation period, subjects spent 15 seconds writing six associations to either two or eight words unrelated to the test items.

The design of the experiment was within-subjects. Each subject experienced both helpful and unrelated associate pairs, and all three incubation conditions. When subjects worked in the nonfixation condition (helpful associates), there was no difference in improvement scores over the three incubation periods. When subjects worked in the fixation condition, the 2-minute incubation period yielded the most improvement. Combining the results of the experiments, one can conclude that only subjects of Smith and Blankenship's (1991) who were fixated due to misleading clues benefited from the incubation period.

Of the experiments in this section, one experiment (Olton & Johnson, 1976) failed to find any effect of clues on incubation and another (Dominowski & Jenrick, 1972) did not report any analysis of the interaction of incubation and hints. Driestadt (1969), Browne and Cruse (1988), Mednick et al. (1964) and Dorfman (1990) all reported that hints directing subjects toward correct answers significantly increased the possibility of solution, and that those hints were more useful to subjects during incubation periods. Instead of using helpful clues, Smith and Blankenship (1989, 1991) exposed subjects to misleading clues. Their results show that the misleading clues hindered problem solving and that those effects dissipated during the incubation period. It would appear, then, that hints presented immediately before or during incubation have a strong effect on performance. The nature of the effect is determined by whether or not the clue is helpful in reaching the correct answer to the problem.

Activity During Incubation Period

To understand incubation fully, it is important to know, not just if the passage of time, type of problem or the presence of clues matters, but also if the particular activities engaged in during the interval matters. Variations in performance due to differences in incubation period activities could potentially provide some insights into the underlying processes involved.

Gall and Mendelsohn (1967) examined the effects of various activities during the incubation period using the multiple trial/multiple item method. Their subjects worked on 30 Remote Associate Test items for two minutes apiece. The control subjects then returned to work for five minutes on each of five items that they had failed to solve. Incubation subjects participated in a distractor task in which they made judgments about the physical weight of items for 25 minutes. This task was chosen because it did not require verbal solutions and provided no cues to the RAT items. A second incubation group gave free associations on the three words in each of the five missed items for 25 minutes. After the incubation interval, both groups of incubation subjects returned to work on five items that they missed for two minutes apiece.

Gall and Mendelsohn conducted two separate experiments using identical procedures. Results based on the combined data of the two experiments reveal that more RAT items were solved by those in the continuous work group than either of the incubation groups, although subjects in the association group performed slightly better than those in the weight judgment group. Gall and Mendelsohn were careful to note that their experiment did not allow for conclusion of whether time away from the RAT items or the time during the second work session was responsible for solutions in the incubation groups.

Olton and Johnson (1976) aggressively manipulated activity during the incubation period in an experiment that used Driestadt's (1969) farm problem. There were eight experimental groups, each with a different activity during the incubation period: a) undirected incubation time; b) demanding cognitive tasks consisting of the Stroop test and counting backward by threes; c) active review of the farm problem and solutions tried; d) a lecture on mental sets and breaking mental sets; e) relaxation; f) prominent display of visual analogies; g) unobtrusive display of visual analogies; and h) unobtrusive display of visual analogies plus lecture on set breaking. Recall from the "Clues During Incubation" section of the present paper that the two control groups in the experiment did not differ significantly from each other and the results were pooled. Comparisons revealed that none of the treatment conditions differed from the pooled control. Despite no comparison of the pooled control groups versus the pooled treatment groups, Olton and Johnson concluded that there was no effect of the incubation period.

Browne and Cruse (1988, Exp. 2) also examined whether differences in performance resulted from various activities during the incubation period (see also "Clues During Incubation"). Their subjects also worked on the farm problem. Subjects in incubation groups worked for 10 minutes, then either drew shapes, relaxed with music or memorized text, and returned to the farm problem for the last 10 minutes. Subjects relaxing during the incubation period were more successful than other incubation subjects and the 20-minute control group, but did not differ from the 25-minute control group. The two control groups did not differ from each other. In this experiment, relaxation during incubation resulted in the best performance when returning to the problem.

Beck (1979) was concerned with the previous knowledge or experience his subjects might have had with any objects or situations in problems used in typical experiments, but he did not manipulate problem type in his experiment. In an effort to control for previous knowledge, he used a completely different type of task. Beck introduced subjects to a fictional new product called "luminium." After reading a brief description of the product, subjects were asked to list uses and qualities of luminium. Subjects in the incubation condition worked for 12 minutes, then relaxed or worked on an essay for 20 or 30 minutes. Finally, they returned to the luminium task for 12 minutes. Control subjects worked continuously for 24 minutes. When more than 45 of the same answer was given by the 240 subjects, that answer was eliminated from all protocols. Remaining original answers were given one point apiece and a composite score was figured for each subject. Results showed a significant difference between the groups such that the 30-minute incubation group had higher scores than the 20-minute and control groups, which did not differ from each other. The type of activity during incubation apparently had no effect.

Bennett (1975) also examined the effects of engaging in various activities during incubation. She initially gave incubation subjects one minute to work on each of 20 Remote Associate Test items. Those in the incubation condition were then given 10 minutes either to work on math problems (structured) or listen to music (unstructured). Finally, subjects were given 20 minutes to complete any unsolved RAT items. By participating in three trials, each of the three groups of 30 subjects apiece was exposed to every condition: structured incubation period, unstructured incubation period, and no incubation period. There was no difference in the mean number of test items solved by condition.

It is difficult to draw conclusions based on Bennett (1975) given the methodological problems in the study. There was an unexpected group effect showing that the three test groups were not initially equivalent. This seems to indicate a lack of random assignment to conditions. There was also an effect of trial such that all groups scored better during the second trial regardless of condition. Bennett mentioned fatigue as a possible explanation of this result. Each of the trials was 50 minutes and the three trials were run consecutively resulting in three hours in the experimental setting. Better performance on the second trial might be expected as subjects had one trial to warm up and had not yet grown overly tired. Cautious interpretation of the results would seem warranted.

A. S. Patrick (1986) used an unusual methodology to explore the effects of different activities on incubation. Performance on Remote Associate Test items served as the dependent measure in this study. Subjects were presented with 30 test items for two minutes each. Five missed items were given again for one minute apiece. If any item was solved, it was replaced with another missed item from the first trial. Finally subjects worked on these five missed problems for a total of eight minutes. They either worked continuously, or in an interpolated situation in which problems were attempted for two minutes apiece, and another activity was undertaken for five minutes, and then problems were again attempted for two minutes apiece. The interpolated activities included unstructured conversation with the experimenter or work on a mental rotation task. The two-five-two pattern was repeated until subjects spent a total of eight minutes on each of the missed test items. Analysis showed significant effects of incubation condition. Incubation subjects solved more problems in the last six minutes than control subjects. Among the various conditions, only the mental rotation and control groups differed significantly from each other. (See discussion in "Level of Ability".)

Smith and Blankenship (1989) used a variety of activities during the incubation periods in their experiments. Recall that their experiments were concerned with the effects of misleading clues on the solution of rebus word puzzles (see "Clues During Incubation"). In Experiments 1 and 2, a music perception task and unfilled free time were compared. In Experiment 3, time in the incubation period was spent either solving rebuses unrelated to the test items, or solving rebuses and either doing math problems, relaxing or completing a music perception task. Finally, in Experiment 4, incubation subjects were asked either to do math problems or read a story. Smith and Blankenship (1989) failed to find effects of activity on performance in any of their experiments.

A multitude of different activities have been used during the incubation periods in the experiments reported here. Gall and Mendelsohn (1967) found, unsurprisingly, that subjects who developed associations to RAT items during incubation had better subsequent performance than those who made judgments about the physical weight of items. Browne and Cruse (1988) reported that their subjects who relaxed with music did better than those who drew shapes or memorized text. Patrick (1986) found that subjects who conversed with the experimenter during incubation performed more poorly than those who completed a mental rotation task. However, Olton and Johnson (1976), Beck (1979), Bennett (1975), and Smith and Blankenship (1989) used some of the same tasks, including relaxation, music perception, and demanding mental tasks like the Stroop test, and found no differences among groups. It would appear that there is no clear pattern with respect to the effects of activity during the incubation period.

Level of Ability

In addition to any questions about normative aspects of incubation, one might legitimately ask whether the effects are likely to appear in all individuals, or only in certain subgroups of the larger population. Specifically, might expertise in a domain or problem solving ability be influential?

Ability has been measured in several studies to attempt to explain differences in incubation. In Murray and Denny (1969), the ball problem proposed by Saugstad and Raaheim (1957) was used. To solve the problem, subjects are required to move several steel balls from a drinking glass to a metal cylinder using a nail, pair of pliers, length of string, pulley, rubber bands and newspaper. Subjects are not allowed to move close to where the drinking glass and cylinder sit. The glass sits on a movable frame. The solution consists of using the nail and string to hook the frame holding the glass and drag it to the subject. Then the balls are removed and dropped into the cylinder using a tube constructed of newspapers and rubber bands. Murray and Denny allowed subjects to work for a total of 20 minutes. Those in the incubation condition were interrupted after five minutes, taken to another room and given five minutes to work on a paper and pencil measure of "ability." They were then allowed to return to work for 15 minutes. Those in the control condition worked steadily for 20 minutes. After completing the 20-minute experimental session, both groups were given the Gestalt Transformation Test to measure problem solving ability.

Because many subjects who had the correct plan were unable to implement it mechanically, verbal solutions were used by the experimenters for analysis. When the time it took for subjects to report the solution to the problem was measured, there was no main effect for incubation. However, an interaction between ability and incubation was observed. Subjects of low ability were faster to provide solutions in the incubation condition, whereas those of high ability did better in the control condition. Murray and Denny hypothesized that this effect was due to high ability subjects engaging in an orderly search process that was disrupted by the incubation period, and low ability subjects being fixated on inappropriate approaches that were interrupted by incubation, allowing more original responses to surface.

Dominowski and Jenrick (1972) also measured problem solving ability with the Gestalt Transformation Test. The test was administered after subjects finished solving the hat rack problem. Individuals with a tendency to become easily fixated on the use of stereotyped responses were classified as being of low ability. Recall that Dominowski and Jenrick (1972) also manipulated whether subjects received hints about how to solve the insight problem (see "Clues During Incubation").

Results for the first experiment show that, of those receiving the hint, a few in the low ability/incubation group failed to solve whereas all subjects in the other hint conditions solved the problem. There was an interaction between ability level and the hint, such that those of high ability performed slightly better in the incubation than control condition. Those of low ability in the incubation condition required significantly longer to solve than low ability, control subjects.

In the second experiment, Dominowski and Jenrick's (1972) subjects again completed the Gestalt Transformation Test as a measure of ability after finishing work on the hat rack problem. The interaction between hint and ability was observed. In the hint condition, high ability subjects used the hint more quickly after incubation than during continuous work. Low ability subjects showed the opposite pattern, using the hint more quickly in the continuous work condition than the incubation condition, but the interaction was not significant.

It is interesting to note that the results of Murray and Denny (1969) and Dominowski and Jenrick (1972) are opposed with regard to ability level. Murray and Denny found that high ability subjects performed better in the control condition and low ability subjects performed best in the incubation condition, whereas Dominowski and Jenrick found that high ability subjects performed best with incubation and low ability subjects required continuous work. There is an experimental factor that might have contributed to this difference in result. The problem used by Murray and Denny was easier for subjects to solve than the one used by Dominowski and Jenrick. Smith (1995) asserted that, if effects are to be observed, subjects must be "stuck" on a problem before the incubation period. It more likely that the subjects of Dominowski and Jenrick were fixated than those of Murray and Denny. Thus the interaction of ability with the incubation period observed in the studies was likely to be different.

In fact, Murray and Denny suggested that incubation interrupted their high ability subjects during an "orderly search process" (p. 275) causing their performance to be decreased. If their subjects were still working productively, the incubation period would not be expected to help. The highly-able subjects of Dominowski and Jenrick who were fixated in work on a difficult problem would be expected to benefit from the incubation period. They should display better performance than low ability subjects after incubation simply because they were better at problem solving.

Using solution of Remote Associate Test items with intermittent work on a variety of distractors (mental rotation, conversation or other RAT items) as the experimental manipulation, Patrick (1986) demonstrated an interaction of ability level and activity, as well as an incubation effect (see "Clues During Incubation"). The median number of problems solved in the initial presentation of the 30 test items was used to split subjects into high and low ability groups. The interaction revealed that, for subjects of high ability, performance in the mental rotation condition was significantly better than in the continuous condition, whereas the other conditions did not differ. Patrick (1986) concluded that incubation aids subjects more under conditions that completely remove the subject from the problem to be solved, and that this is especially true for high ability subjects.

The increased performance of incubation subjects of high ability is consistent with the results of Dominowski and Jenrick (1972), who found that an incubation period aided subjects of high ability and hindered those of low ability. Patrick's (1986) experimental subjects had more than one trial with items and were unable to solve them before the incubation period was given. This situation is similar to that of Dominowski and Jenrick in which subjects who were fixated in working on an extremely difficult problem. In both studies, fixated subjects of high ability benefited more from an incubation period than those of low ability.

Mednick, et al. (1964, Exp. 2) measured ability using performance on the Remote

Associates Test, splitting their subjects into high and low creative groups based on the scores. During the incubation experiment, solution of RAT items served as the dependent variable (see "Clues During Incubation"). That being the case, it is certainly not surprising that those of high ability solved more items than those of low ability. A more interesting finding would have been an interaction of ability and the other variables in the experiment (incubation, hints via priming), but none of the interactions in the study were significant.

Browne and Cruse (1988, Exp. 1), who used the farm problem and had subjects draw shapes as clues during incubation (see "Clues During Incubation"), added IQ scores to their analysis as a covariate of the incubation effects to control for ability level. Analysis revealed no significant difference of the scores for subjects in the incubation and control groups. Thus, Browne and Cruse ruled out IQ as a possible reason for differences in performance between the groups.

Brockett (1985) observed significant correlations between measures of ability and the fluency and flexibility demonstrated by subjects on the Remote Associates Test and unusual uses for a brick test. Recall that subjects completed either the RAT or the unusual uses test and scores were calculated for the number of responses, and flexibility defined as giving the correct answer for the RAT items or number of categories used in the unusual uses test (see "Types of Problem" for discussion). In addition, subjects completed the vocabulary section of the WAIS and Preconscious Activity Scale (concerned with associations made by "creative" individuals). Significant positive correlations were observed between the vocabulary section of the WAIS and the fluency scores on the unusual uses test, and vocabulary ability and giving correct answers for the RAT. Subjects designated as creative according to the Preconscious Activity Scale tended to score high on both fluency and flexibility in the unusual uses test.

Ability sometimes has an effect on incubation, but no consistent pattern relating ability to incubation can be seen across the reported experimental findings. For Murray and Denny (1969), subjects of low ability performed better when given incubation. However, in the experiments of Dominowski and Jenrick (1972), Patrick (1986) and Brockett (1985), subjects of high ability performed best with an incubation period. Mednick et al. (1964) did not report an interaction effect, and neither did Browne and Cruse (1988). As mentioned previously, the effects of ability on incubation may be tempered when the subjects are not fixated before being given an incubation period.

Gender

A few researchers have examined the effects of gender in incubation experiments. The first were Gall and Mendelsohn's (1967). Recall that their results showed no effects of incubation using the multiple trial/multiple item method with Remote Associate Test items (see "Activity During Incubation Period"). Instead the control group that worked on the RAT items continuously solved more than the incubation groups who either associated to the words in the RAT items or made judgments about the weight of items. Gall and Mendelsohn also analyzed the effects of both subject and experimenter gender on performance on the RAT items. The main effect of experimenter gender was significant as were several interactions involving subject and experimenter gender. More correct responses were made when subjects were tested by male experimenters. Post-hoc analysis revealed that this was due to an increase in performance by female subjects when tested by male experimenters.

Of more interest, there was an interaction of subject gender and condition. Males performed best in the control group. Their performance in the incubation groups that did physical weight measurement and associations with the RAT items did not differ significantly from each other. Subjects in these groups had less success than those in the control group. Females performed equally well in the control and association conditions, and better in both of those than in the weight judgment group. Post-hoc analysis revealed that this interaction was due to better performance by females in the association group than males.

Browne and Cruse (1988, Exp. 1) also examined the interaction of gender and incubation. Recall that they used the farm problem with hints during the incubation period to demonstrate effects (see "Clues During Incubation"). Analysis revealed that male subjects were more likely to solve the problem before the incubation period than females. This result is not in conflict with that of Gall and Mendelsohn (1967) who showed that men tend to perform better when allowed to work continuously.

Brockett's (1985) experiment comparing incubation performance on the unusual uses test and RAT items also included an exploration of the effects of gender (see "Types of Problem"). Brockett's subjects completed the Bem Sex Role Inventory. A significant positive correlation showed that females were more likely than males to have high fluency scores on the unusual uses test. Scores on the Bem were not significantly correlated with fluency on the RAT, or flexibility scores on either instrument.

Frankel (1990) was concerned with the possible effects of subject gender, as well as imagery training. In addition to analyzing the effects of presence of an incubation period and that of imagery training presented before subjects worked on the life-relevant problems (see "Preparation Activities"), Frankel examined whether gender affected the originality and fluency of subjects during the last 10 minutes of problem solving. Neither the main effect of gender nor the interaction of gender with the other two variables reached significance.

Females benefited more from incubation in the experiments by Gall and Mendelsohn (1967) and Browne and Cruse (1988). There was no effect in the study by Frankel (1990). The results of Brockett (1985) seem inconclusive and interpretation is hampered by the measurement difficulties mentioned earlier. The relevance of gender to incubation effects is not at all clear, when viewed across all of these studies.

Length of Incubation Period

A final variable of interest in incubation situations is the length of the incubation period. To what extent does observing incubation effects depend on the length of time away from the problem?

Among the first to be concerned with length of the incubation period were Fulgosi and Guilford. In a set of two experiments (1968, 1972), they attempted to determine the incubation time period that would provide the most powerful effects. A total of 4 minutes was allowed for subjects to work on a Consequences problem. Subjects in the incubation condition worked for 2 minutes, and then were interrupted to work on a number series completion task before returning to the problem for 2 minutes whereas those in the control condition worked continuously for 4 minutes. In the 1968 experiment, incubation subjects worked on the filler task for 10 or 20 minutes. Incubation subjects in the 1972 experiment worked on the filler for 30 or 60 minutes.

In the 1968 experiment, Fulgosi and Guilford measured fluency and originality (how remote or obvious subjects' answers were). A significant incubation effect was observed only when using originality. The 20-minute incubation group received the highest score, followed by the 10-minute group and finally the control group. Fulgosi and Guilford (1968) commented that the 20-minute interval seemed to give the subjects an entirely new start in the problem solving process and suggested that, as at the start of the task, there seemed to be a need to get obvious responses out of the way when returning to work on the problem after an extended break. Results of the 1972 experiment using the fluency scores showed that the 30-minute incubation group made the greatest number of responses whereas performance of the control and 60-minute groups did not differ. Combining the results of the two experiments, Fulgosi and Guilford (1972) suggested that the incubation effect reached its maximum potential between 20 and 30 minutes.

Recall that Beck (1979) also compared the effects of 20- and 30-minute breaks on performance in the "luminium" task (see "Activity During Incubation Period"). He found that the 30-minute group outperformed the 20-minute and control groups. The 20-minute group did not differ from the control group in terms of performance. Thus the results of Fulgosi and Guilford (1968; 1972) and Beck appear to converge on an optimum length of time for the incubation period.

In her first experiment, Silveira (1971) also manipulated length of the incubation period. Recall that Silveira's subjects worked on the chain problem for a total of 35 minutes (see

"Preparation Activities"). Experimental subjects were given an incubation period lasting either 30 minutes or 3.5 hours. During the short incubation period, subjects were allowed to read their own books or one provided by the experimenter. During the long incubation period, subjects were allowed to leave the experimental room, but were asked not to think about the problem or work on it until returning. No incubation effects were revealed, and Silveira made no comparison of the responses of the two incubation groups. Of the 25 subjects receiving the short incubation period, 14 solved the problem, whereas 16 of the 24 in the long incubation period group solved. Despite these data, Silveira concluded that the longer period was most beneficial and used it in her second experiment.

Silveira (1971; Exp. 3) attempted to replicate the findings of Fulgosi and Guilford (1968) using a different Consequences item--what would happen if people no longer needed food to live. Subjects for this experiment were those who quickly solved the insight problem during her first and second experimental sessions. Control subjects worked on the problem continuously for six minutes, whereas incubation subjects worked three minutes and then took a 20-minute break to read before returning to work for a final three minutes. There was no significant difference between the first and second 3 minute work periods for either fluency or originality. Silveira did not compare the performance of subjects in the incubation and control conditions.

Despite claiming that the experiment was a replication, it appears that Silveira's analysis and procedure were not the same as that of Fulgosi and Guilford (1968). First, she did not manipulate the amount of time experimental subjects spent in incubation. Second, Silveira's results were based on the use of subjects who had quickly finished insight problems for her other studies. Thus, instead of being a random sample, subjects in the third experiment were selected based on their ability to solve the insight problem quickly. Third, the subjects in Silveira's (1971; Exp. 3) experiment had already been primed by solving one problem and were asked to solve to another. It is likely that some transfer existed, although the effect on performance cannot be predicted.

Silveira's task was also difficulty because Fulgosi and Guilford's (1968) study is short on details. They stated that two Consequences items were used but not whether all subjects worked on both items or, instead, if there were separate groups. They reported incubation effects found using a 3-way ANOVA but give no indication if the dependent variable was the difference between the first and second work sessions, or only scores on the second work session for each Consequences item. Fulgosi and Guilford (1968) merely report that the subjects receiving a 20-minute incubation period showed higher fluency on the second trial. Thus, it seems unlikely that Silveira (1971; Exp. 3) was able to complete a faithful replication given this lack of information. Nevertheless, Silveira (1971) concluded that there was a failure to replicate and lack of incubation effects.

Smith and Blankenship (1989; Exps. 1, 2, 3) also manipulated the length of time in the incubation period during their experiments with rebus word puzzles (see "Clues During Incubation" and "Activity During Incubation Period"). In Experiments 1 and 2, the incubation subjects were given either 5 or 15 minutes. There was no difference in performance between the groups in the first experiment. However, the second experiment revealed that subjects in the 15-minute group performed better than those in either the 5-minute or control groups. There was no difference in performance of incubation subjects in the third experiment who were given 10 or 15 minutes during the incubation period.

Goldman, Wolters, and Winograd (1992) used the multiple trial/multiple item method to examine the effects of incubation periods of various lengths. Control subjects were given 48 anagrams to solve and allowed 50 seconds to work on each. Incubation subjects worked on 36 anagrams for 15 seconds each. Then, they either answered general knowledge questions for 20 minutes or were asked to return 24 hours later. After the incubation period, subjects worked on the unsolved anagrams from the first trial in addition to 12 anagrams not previously presented for 35 seconds each. Thus, the total time incubation subjects worked on "old" anagrams was 50 seconds, with 35 seconds on each "new" anagram.

There was no difference between groups in ability to solve problems (36 possible) during the first 15 seconds of work. During the last 35 seconds of work, the subjects in the 24-hour incubation condition solved significantly more of the 36 problems than those in the control condition, whereas those in the 20-minute condition did not differ from either of the others. For the incubation groups, there was a significant difference between old and new items, such that better performance was observed on the old items. The authors concluded that this result demonstrates no general improvement in incubation subjects' ability to solve anagrams. Instead, there was only an improvement in solving those items that had been subjected to incubation.

Dorfman (1990; Exp. 3, Exp. 4) was also interested in the effects of various lengths of time during incubation. In both experiments, she gave subjects items like those from the Remote Associate Test and with additional clues over several trials (see "Clues During Incubation"). In the third experiment, there was no main effect of the delay periods of 5 and 15 minutes. In the fourth experiment instead of designating problems to conditions to form a within-subjects experiment, separate groups of subjects were assigned to control, or 3-minute, 8-minute or 13-minute incubation condition. During the incubation period between the first and second trials, subjects worked on math problems. The subjects in the 8-minute incubation condition solved significantly more experimental problems in the first trial than subjects in other groups. However, there was no main effect of length of incubation period in this experiment either.

Many experiments explored the length of the incubation period and their results are difficult to summarize. Fulgosi and Guilford (1968; 1972) reported that their 20 minute incubation group outperformed the 10 minute group, and was outperformed by the 30 minute group. However, the performance of subjects in the 60 minute group was lower than those in the 30 minute group. Beck (1979) also demonstrated that 30 minutes seemed more effective than 20 minutes. But Silveira (1971, Exp. 3) failed to find any incubation effects using a 20-minute incubation period.

Smith and Blankenship (1989, Exp. 1 & Exp. 3) found no differences between groups having 5 and 15 minutes, and 10 and 15 minutes, but their 15-minute group outperformed the 5-minute group in Experiment 2. Dorfman's (1990, Exp. 3) 5- and 15-minute delay conditions resulted in similar performance, while her 8-minute group outperformed both the 3- and 13-minute groups in the fourth experiment. It should be remembered, however, that her groups were not equivalent at the beginning of the fourth experiment. Of studies using longer experimental periods, Silveira (1971, Exp. 1) found 3.5 hours superior to 35 minutes, and

Goldman et al. (1992) showed that 24 hours was better than 20 minutes.

There do seem to be a few general effects of time during the incubation period. For studies using incubation periods of 15 minutes or less, there seems to be little difference in performance due to length of time. Over 15 minutes, increasing the length of the incubation period seems to increase performance until the period reaches 30 minutes and then declines approaching the one hour mark. For incubation periods of 3.5 hours or more, increasing the length seems to help performance. The longest break period in the experimental literature that still resulted in incubation effects was two weeks but, as discussed, Patrick (1938) used the difficult problem of designing an experiment and the study lacked control (see "Demonstrating Incubation").

Conclusion

Our review reveals several variables that appear to affect the incubation process. Longer periods of preparation increase the benefits of incubation. Clues given either immediately before or during incubation interact with incubation effects. Effects of the length of the incubation period depend on whether incubation is short- or long-term, with maximal effects at 30 minutes and longer periods of 3.5 and 24 hours still resulting in increased incubation effects.

Less influential factors are problem type, activities during the incubation period, ability, and gender. Problem type may well affect incubation, but the experimental evidence is inconclusive at this point. Various activities during incubation do not seem to make any substantial difference. Ability level appears to interact with subjects' fixation. For high-ability subjects who are stuck, incubation is more beneficial than for low-ability subjects. Females may be able to make better use of incubation periods, but more research on gender differences is needed.

The experimental literature of incubation in problem solving and creativity has demonstrated that several variables may interact with incubation periods. While time away from conscious problem solving is important, its passage is not the only factor that results in better problem solving upon the return. As this review has shown, it may be time to extend the definition to include the other components that interact with the incubation period to improve problem solving and creativity. More than anything else, this review reveals that much work remains to be done before a clear picture of the true nature of incubation will emerge. We urge researchers not to incubate on incubation, but rather to pursue a direct and deliberate attack on the problem.

Acknowledgements

The present chapter was supported in part by NSF Grant DMI-0115447 to Steven M. Smith and Jami Shah.

References

Beck, J. (1979). The effect of variations in incubation activity and incubation time on creative response production. Unpublished doctoral dissertation, New York University.

Bennett, S. (1975). Effect of incubation on the associative process of creativity. Unpublished doctoral dissertation, University of Notre Dame.

Brockett, C. (1985). Neuropsychological and cognitive components of creativity and incubation. Unpublished doctoral dissertation, Virginia Commonwealth University.

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.

Frankel, A. (1990). The effect of incubation and eidetic imagery training on tasks of problem solving. Unpublished doctoral dissertation, Fordham University.

Fulgosi, A., & Guilford, J. (1968). Short-term incubation in divergent production. American Journal of Psychology, 81, 241-246.

Fulgosi, A., & Guilford, J. (1972). A further investigation of short-term incubation. Acta Instituti Psychologici, 64-73, 67-70.

Gall, M., & Mendelsohn, G.A. (1967). Effects of facilitating techniques and subject-experimenter interaction on creative problem solving. Journal of Personality and Social Psychology, 5, 211-216.

Goldman, W., Wolters, N., & Winograd, E. (1992). A demonstration of incubation in anagram problem solving. Bulletin of the Psychonomic Society, 30, 36-38.

Kaplan, C. (1989). Hatching a theory of incubation: Does putting a problem aside really help? If so, why? Unpublished doctoral dissertation, Carnegie-Mellon University.

Khatena, J., & Torrance, E. (1973). Thinking creatively with sounds and words: Norms-technical manual. Lexington, MA: Personnel Press.

Maier, N.R.F. (1945). Reasoning in humans: III. The mechanisms of equivalent stimuli and of reasoning. Journal of Experimental Psychology, 35, 349-360.

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

Mednick, S. (1962). The associative basis of the creative process. Psychological Review, 69, 220-232.

Olton, R.M. (1979). Experimental studies of incubation: Searching for the elusive. Journal of Creative Behavior, 13, 9-22.

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

Parrish, J. (1981). The effects of preparation and incubation upon production of original verbal images. Unpublished doctoral dissertation, Mississippi State University.

Patrick, A.S. (1986). The role of ability in creative "incubation". Personality and Individual Differences, 7, 169-174.

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

Peterson, C. (1974). Incubation effects in anagram solution. Bulletin of the Psychonomic Society, 3, 29-30.

Saugstad, P., & Raaheim, K. (1957). Problem solving and availability of functions. Acta Psychology, 13, 263-278.

Silveira, J. (1971). Incubation: The effect of interruption timing and length on problem solution and quality of problem processing. Unpublished doctoral dissertation, University of Oregon.

Smith, S. (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 Publishers.

Smith, S., & Blankenship, S. (1989). Incubation effects. Bulletin of the Psychonomic Society, 27, 311-314.

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

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

Table 1

Variables Manipulated by Experiment

Experiment	Inc. Effect	Length of Inc	Prep. Activity	Problem Type	Clues Given	Activity In Inc.	Ability Level	Gender
Beck	yes: I	x				x		
Bennett	no: M					x		
Brockett	yes: I			x			x	x
Browne & Cruse (exp. 1)	yes: I				x		x	x
Browne & Cruse (exp. 2)	yes: I				x	x		
Dominowski & Jenrick (exp. 1)	yes: I				x		x	
Dominowski & Jenrick (exp. 2)	no: I				x		x	
Dorfman (exp. 3)	yes: M	x			x			
Dorfman (exp. 4)	yes: M	x			x			
Driestadt	no: I				x			
Frankel	yes: I		x					x
Fulgosi & Guilford, 1968	yes: I	x						
Fulgosi & Guilford, 1972	yes: I	x						

Gall & Mendelsohn	no: M					x		x
Goldman, et al.	yes: M	x						
Kaplan (exp. 1)	yes: I							
Kaplan (exp. 2)	yes: I							
Kaplan (exp. 3)	yes: I							
Kaplan (exp. 4)	yes: I		x					
Kaplan (exp. 5)	yes: M		x					
Mednick, et al. (exp. 1)	yes: M				x			
Mednick, et al. (exp. 2)	no: M				x		x	
Murray & Denny	yes: I						x	
Olton	no: I							
Olton & Johnson	no: I					x		
Parrish	no		x					
Patrick, 1938	yes: I							
Patrick, 1986	yes: I					x	x	
Peterson	yes: M			x				

Silveira (exp. 1)	no: I	x	x					
Silveira (exp. 2)	yes : I							
Silveira (exp. 3)	no: I							
Smith & Blankenship (1989, exp. 1)	yes : M	x			x	x		
Smith & Blankenship (1989, exp. 2)	yes : M	x			x	x		
Smith & Blankenship (1989, exp. 3)	yes : M	x			x	x		
Smith & Blankenship (1989, exp. 4)	yes : M				x	x		
Smith & Blankenship (1991, exp. 1)	yes : M				x			
Smith & Blankenship (1991, exp. 2)	yes : M							
Smith & Blankenship (1991, exp. 5)	yes : M							

Note: The letters "M" and "I" in the Incubation Effects column refer to the experimental method used.

Table 2

Experimental Design for Dorfman (1990; Exp. 3 & Exp. 4)

	Condition of Items		
Trial	Immediate	5-Minute Delay	15-Minute Delay
1	3 clues	3 clues	3 clues
2	3 clues	Not presented	Not presented
3	4 clues	3 clues	Not presented
4	5 clues	4 clues	3 clues
5	6 clues	5 clues	4 clues
6	Not presented	6 clues	5 clues
7	Not presented	Not presented	6 clues