Incubation benefits only after people have been misdirected

EDWARD VUL AND HAROLD PASHLER
University of California, San Diego, La Jolla, California

Incubation refers to the popular idea that stopping work on a problem may, at times, be a more efficient means by which to reach a solution than continuing to work. Empirical studies of incubation have used few participants and have provided ambiguous and discrepant results. We investigated three potential accounts of incubation in retrieval and search problems (subconscious work, spreading activation, and fixation forgetting) with the help of a large Internet-based participant pool. The amount of time allotted for explicit work on each of 12 problems was controlled, while the distribution of that time was manipulated in several incubation conditions. When problems were presented by themselves, none of the incubation conditions aided in the solution of either anagram or remote associate test problems. However, incubation benefits arose when participants were given misdirecting clues (probably because time delays facilitated forgetting of these clues).

"You're out for a jog, you're completely relaxed, your mind a pleasant blank. Then all of a sudden into your head pops the solution to a problem you have been mulling over for days or weeks. . . ." So opens *The Creative Spirit* (Goleman, Kaufman, & Ray, 1993, p. 15), one of many how-to guides to creativity and problem solving. Folk psychology is rife with similar evocations of what has come to be referred to as *incubation*: the idea that a solution is more easily reached if work on the problem is interrupted, rather than continued without interruption.[1] The concept of incubation has been popularized not only by anecdotal endorsements of well-known mathematicians (e.g., Poincaré; cited by Olton, 1979) and scientists (e.g., Feynman; cited by Seifer, Meyer, Davidson, Patalano, & Yaniv, 1995), but also by the rather appealing notion that creative problem-solving ability may be improved by doing *less* work on the problem. Since Wallas (1926) listed incubation as one of four primary stages of creative thought, a period of effortless progress has been incorporated into many other models of creativity (e.g., Barron, 1988; Osborn, 1953) and numerous self-help guides (e.g., Goleman et al., 1993; Higgins, 1994; Ray, 1989).

Despite the popularity of the concept, consistent empirical support for incubation has been scarce. Olton (1979) found eight studies of incubation that he judged to be without serious methodological flaws; of those, he found four studies confirming the phenomenon and four questioning it. He also reported that three of the four studies yielding support were not replicated in follow-up research (Olton & Johnson, 1976).

Such empirical inconsistencies may be caused, in part, by the fact that the literature on incubation consists of many studies with small sample sizes. The resultant low

power may cause Type II error in failures to replicate incubation effects. However, even when significant, results are less precise with small sample sizes and may muddy the literature. The great bulk of early incubation experiments consisted of just one difficult problem, and many of those studies used few participants (e.g., Gavurin, 1965, 20/condition; Murray & Denny, 1969, 18/condition; Olton & Johnson, 1976, 10/condition). Although later studies assigned between 10 and 48 problems to each participant, the total number of participants was still low: only 8 to 24 per condition (Goldman, Wolters, & Winograd, 1992; Smith & Blankenship, 1989, 1991). The small subject counts of previous studies, therefore, render null results perhaps somewhat suspect and create ambiguity in significant results; thus, it is still unclear to what degree previous studies of incubation effects should be trusted.

In this article, we seek to improve upon the designs of past research by exploring incubation effects in problems involving memory retrieval and search, using much larger numbers of participants than have ever been used in previous studies. In doing so, we attempt to clarify some of the mixed results in the literature and discuss the implications of our findings for theories of incubation effects.

Theories of Incubation Effects

Perhaps the most popular and intriguing account of incubation would be the idea of subconscious work (Wallas, 1926). According to this view, while conscious work on the problem is suspended and other tasks are undertaken, the subconscious mind continues working toward a solution. In a constrained solution space, this can be viewed as a process of trial-and-error search, but it could also involve more robust or elaborate sorts of problem-solving opera-

E. Vul, evul@mit.edu

tions. The fruits of this work become available when the solution spontaneously pops into the individual's mind. Peterson (1974) found evidence in support of such an account, reporting that the total time required for the solution of an anagram was decreased when the time allotted for solution was provided in the form of six 20-sec blocks separated by interruptions during which the participants worked on other anagram problems, rather than one uninterrupted 120-sec period. Assuming that the amount of work needed to solve an anagram problem is unchanged, the fact that less time was needed for conscious work in the six 20-sec blocks (incubation) condition suggests that some of the necessary work was undertaken subconsciously while the participants directed conscious work to other problems.

Unfortunately, however, Peterson (1974) reported only the solution times for each condition but did not report the number of problems solved. If solution rates decreased with average solution times, the finding may have been the result of participants' working consciously on problems that were supposedly incubating, at the cost of ignoring those that were immediately presented. Such a result would not be unexpected, since a study in which a similar paradigm was used showed that participants were more likely to solve a problem if the solution time was massed (Gavurin, 1965). Together, these studies only further muddy any conclusions one might have drawn about the reality and efficacy of subconscious work during a break from problem solving.

A slight variation of the subconscious work account is that of spreading activation (Yaniv & Meyer, 1987). This account characterizes the work taking place without the guidance of consciousness as activation spreading in an associative network (e.g., *dog* activates *bone*, which activates *dinosaur*, etc.), which suggests a process more akin to memory retrieval than to algorithmic computations or a simple trial-and-error search. Thus, if any benefits of an interruption are caused by spreading activation, one would expect such benefits to be most evident during the solution of semantic problems involving retrieval of associations.

Another popular and plausible account of incubation is sometimes termed *fixation forgetting* (Smith, 1995; Smith & Blankenship, 1989, 1991). *Fixation* refers to the idea that problem solvers often wander down an incorrect solution path and lose the ability to recover from this error (becoming *fixated* on the erroneous path). When this occurs, more work on the problem will involve futile efforts to proceed further down the erroneous path. Thus, a period of incubation may be helpful because the interruption allows fixation to dissipate. Although this account is undeniably plausible, it raises two additional questions. Is there evidence of self-induced fixation during problem solving? If so, does a period of incubation help relieve fixation?

Smith and Blankenship (1989, 1991) reported that an interruption during a period of problem solving resulted in more problem solutions only under conditions of artificially induced fixation in which participants were purposely led down the incorrect solution path. These findings imply that an interruption can, indeed, ameliorate

fixation if it has been artificially induced but that people may not become fixated on their own. However, Smith and Blankship's studies showed a slight (albeit not significant) improvement in solution rates for incubation without artificially induced fixation and involved relatively few participants (8–11 per condition), resulting in a power for detecting that effect of only .24. Due to these factors, no strong conclusions can be drawn about the prevalence and dissipation of fixation when it is not artificially induced.

The first two of these accounts (subconscious work and spreading activation) ascribe fairly sophisticated problem-solving, or memory-searching, capabilities to the subconscious mind. Although the faculties of the subconscious may not be well known, the capacities of conscious processing are generally believed to be limited by a central bottleneck (Pashler & Johnston, 1998; Welford, 1967). This bottleneck forces central-processing stages to operate sequentially whenever people attempt to perform two, even fairly simple, tasks involving response selection. Chronometric studies indicate that cued memory retrieval is also subject to this central bottleneck (Carrier & Pashler, 1995). If the subconscious mind can solve difficult problems while the conscious mind is otherwise occupied, it would seem surprising that so little parallel central processing would be observed with dual-task experiments using mere choice reaction time tasks (which often involve such seemingly trivial cognitive operations as recalling that the left button should be pressed in response to a high-pitched tone).

The Present Study

The present study is the first to explore multiple accounts of incubation effects, using what we believe to be an improved design that not only addresses the unanswered questions mentioned in the previous section, but also tests a sufficient number of participants to attain reasonable statistical power. In Experiment 1, we investigated the effects of interruptions on problem solving by replicating the basic design of Peterson (1974), which seemed to offer the most positive support for the *subconscious work* account. In this design, work on anagram problems took place in either six 20-sec interrupted blocks, or one 120-sec uninterrupted block. In addition, the original design was supplemented with more detailed measurements of response accuracies and latencies, to disambiguate the original findings. Furthermore, two new incubation conditions were added: two 60-sec blocks interrupted by more anagram problems and two 60-sec blocks interrupted also by an unrelated video game task. The latter was added to address the potential concern that a concurrent verbal task of solving other anagrams might interfere with subconscious work on incubated problems. If the subconscious may be employed to aid in the solution of anagrams, we would expect to find the greatest incubation effect in the condition with the longest interruption: two blocks plus a video game. Similarly, we would expect this condition to be optimal for forgetting fixation if participants become fixated during algorithmic search. The spreading activation account, however, does not guarantee benefits of an interruption in this experi-

ment, because it may not be possible to solve anagrams by activation in an associative network.

Since it is possible that spreading activation and fixation occur selectively for semantic problems, in Experiment 2 we focused on remote associate test problems to address these two accounts of incubation. The problems would seem to be an especially good candidate for a task that might profit from putative spreading of activation in an associative network. To thoroughly explore the fixation-forgetting account, we investigated the rate at which fixation occurs and dissipates by comparing improvement resulting from the interruption period with two more conditions involving artificially induced fixation. This manipulation was designed to increase the tendency of participants to follow erroneous solution paths. We could then compare improvement due to interruption with and without artificial fixation, to check whether there was any trend toward fixation forgetting when it was not artificially induced. Here, the fixation-forgetting account predicts great improvement from an interruption when fixation is artificially induced and, if self-fixation occurs, similar trends in improvement without artificially induced fixation. The spreading activation account, however, predicts that an interruption is more likely to be useful if no misleading semantic cues are present, since those will lead to activation's spreading in an inappropriate part of the associative network.

EXPERIMENT 1

Because so much of the literature on incubation has been riddled with low power studies, the primary goal of Experiment 1 was to try several classical manipulations of incubation on a large group of participants. To do this, we employed the increasingly popular methodology of Internet sampling, which allows for the collection of samples larger than would be practical in the laboratory. Although Web-based data collection is still relatively unusual in cognitive psychology, in our own lab we have found good agreement between lab- and Web-based memory studies, and numerous other groups have also reported excellent correspondence between the results of studies in which Internet and laboratory samples were used (e.g., Birnbaum, 1999; Krantz & Dalal, 2000; McGraw, Tew, & Williams, 2000; Reips, 2002).

Method

Participants. Participants ($N = 233$) were recruited from a diverse University of California at San Diego Attention and Perception Lab Internet research panel (for details, see Pashler, Cepeda, Wixted, & Rohrer, 2005). The participants enrolled in this panel participated in various online behavioral science studies in return for incentives such as enrollment in drawings for prizes.

Materials and Design. The 12 five-letter anagrams originally used by Peterson (1974) were employed here. These anagrams were based on words chosen from Thorndike and Lorge (1944) according to their usage frequency. Six words (*sugar*, *train*, *judge*, *admit*, *lower*, and *given*) were categorized as frequent because they occurred more than 100 times per million. The other six words (*liken*, *caste*, *usher*, *covet*, *farce*, and *cower*), which occurred between 1 and 10 times per million, were categorized as infrequent. Two levels of difficulty, based on the number of perturbations from the original

letter order, were used. There were six 2-move perturbation patterns (24351, 32415, 41523, 51324, 14532, and 31254) and six 3-move perturbation patterns (43521, 53421, 35241, 15432, 25314, and 45213). One anagram problem was created with a 1-move perturbation pattern out of the word *house* (the anagram was *houes*) to serve as a test for basic anagram competency and cooperation on the task. An unrelated video game served as the interpolated task. In this game, the participants used a mouse to click on moving yellow circles to progress through the game's levels; with each level, the number of targets and their velocity increased.

This experiment was a $4 \times 2 \times 2$ mixed design with incubation condition varying randomly between subjects and two within-subjects factors: word frequency (high vs. low) and perturbation pattern (two vs. three moves). The participants were randomly assigned to one of the four incubation conditions, which differed in the order and the relative timing of anagram solution and the video game, as described below. For each participant, anagrams were generated by randomly assigning words to perturbation order conditions such that three of the six low-frequency words were in the three-move perturbation condition and three in the two-move condition; the remaining perturbation patterns were assigned randomly to high-frequency words.

Procedure. After reading a description of the study, the participants clicked on the experiment link and completed a consent form and demographic questions. Every participant began the experiment with a 1-min training session to become familiarized with the video game and finished the experiment with a 60-sec presentation of the anagram competency question (solving the anagram *houes*).

For every condition, anagram instructions were presented only once, before the anagrams. The instructions briefly described the anagram problems, gave one example (*veryus = survey*), and explicitly stated that the participants could provide solutions only for the anagram presented on-screen at any given point. During presentation, anagrams were presented in random order in bold, capital letters, with a text box for the participants' answers or guesses immediately below the anagram. After an incorrect guess, the participants were shown the message "Incorrect, please keep trying" for 1.5 sec. When the correct answer was entered, the feedback "Correct" was displayed for 1 sec, and then the next anagram was presented.

Regardless of condition, every participant was exposed to each anagram problem for 120 sec total or until the problem had been solved, and every participant played the video game for a total of 6 min (including the 1-min training session). The order of the various tasks between the video game training phase and the anagram competency question was as follows (see Figure 1).

1. No interruption (NI) condition: After the video game had been played for 5 min, all of the anagrams were presented in a random order for an uninterrupted time interval of 120 sec each or until a solution was reached.

2. Interrupted 20-sec blocks (I20–5) condition: After the video game had been played for 5 min, the anagram problems were presented in a random order for 20 sec each or until a solution was reached; then unsolved anagram problems were presented again for 20 sec each, and so on, for a total of six 20-sec trials for any unsolved problem.

3. Interrupted 60-sec blocks (I60–5) condition: After the video game had been played for 5 min, the anagram problems were presented in a random order for 60 sec each or until a solution was reached; then unsolved problems were presented again for 60 sec each.

4. Interrupted 60-sec blocks plus 5 min of unrelated task (I60+5) condition: The anagrams were presented in a random order for 60 sec each or until solution was reached; then the participants played the video game for 5 min. After that, they were presented with previously unsolved anagrams for another 60 sec or until a solution was reached.

Data were collected through a Web site running on the free and open-source LAMP (Linux, Apache, MySQL, and PHP) framework. The Web site was tested on and proved to be accessible through all major Web browsers. The video game used in the experiment was programmed in Macromedia Flash. All response time data were col-
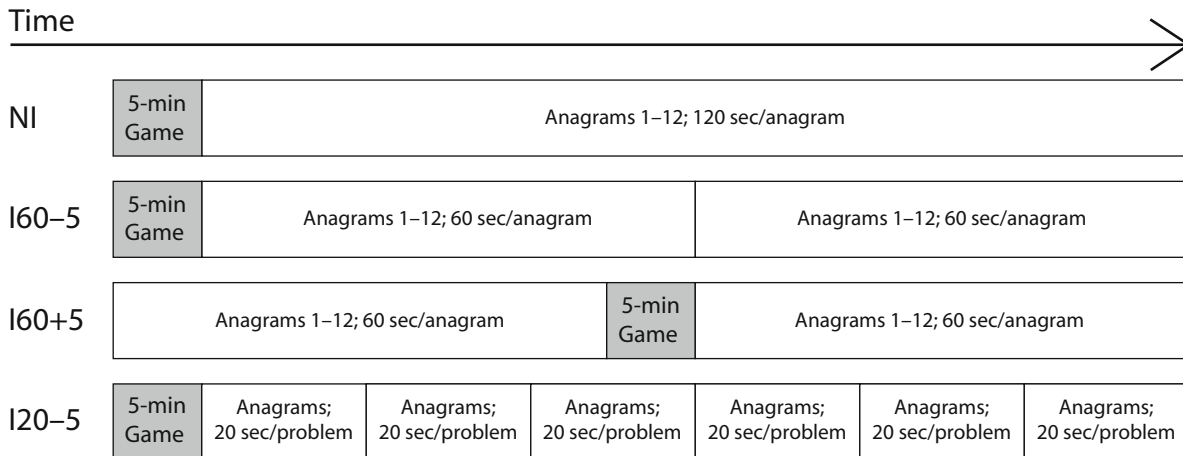
Time

Figure 1. Sequence of anagram and interpolated tasks for the four incubation conditions in Experiment 1.

lected using client-side JavaScript, to allow for more accurate timing data than would be allowed with server-side measures (thus, Internet communication lags were not erroneously entered into our response latency measurements). In independent tests, we determined that our client-side JavaScript latency measurements were accurate within approximately 75 msec. Such temporal resolution is more than sufficient to detect the differences (on the order of many seconds) expected in verbal problem solving, as the results discussed below will confirm.

Results and Discussion

A total of 263 participants completed the experiment. Of those, 30 were dropped from the analysis because they did not answer the anagram competency question correctly (producing *house* from *houes*). Of the remaining 233 participants, 69% participated in the experiment while alone at home, and the other 31% were fairly evenly distributed among other environments. Interestingly, the environment did not have any detectable effects on performance [$F(5,232) = 0.573$, $MS_e = 0.033$, $p = .721$].

To assess performance, total proportion correct and proportion correct conditionalized on word frequency and perturbation difficulty were calculated for each participant. These numbers were then averaged across all participants within each condition (see Table 1). A $4 \times 2 \times 2$ mixed ANOVA with incubation condition as a between-subjects factor and word frequency and perturbation difficulty as within-subjects factors revealed that problems with high-frequency, as opposed to low-frequency, words as their solutions were more often solved [$F(1,229) = 274.98$, $MS_e = 0.054$, $p < .001$]. Surprisingly, three-move perturbation orders were more often solved than two-move perturbation orders [$F(1,229) = 4.39$, $MS_e = 0.051$, $p < .05$]. The interaction between these within-subjects factors was also significant [$F(3,229) = 8.75$, $MS_e = 0.051$, $p < .005$], reflecting the fact that the benefit of a three-move perturbation order was apparent only for high-frequency words. Incubation condition, as a between-subjects factor, had no significant main effect [$F(3,229) = 0.602$, $MS_e = 0.13$, $p = .644$], nor did it interact significantly with any of the within-subjects factors (all $ps > .18$).

The time of solution was recorded for every item solved by each participant (see Table 2). Anagrams of high-frequency words were solved more quickly than anagrams of low-frequency words [$F(1,228) = 156.47$, $MS_e = 875.14$, $p < .001$]. Neither perturbation order nor incubation condition significantly affected solution times (for all main effects and interactions, $p > .15$). Contrary to the effect described by Peterson (1974), in which interrupted 20 sec produced faster average solution times than did 120 sec of continuous work on the anagrams, there was a trend in the opposite direction: 120 sec of continuous work produced solutions about 1.4 sec faster, on average, than interrupted 20-sec blocks. Our power to detect the effect observed by Peterson (effect size = .7) was greater than .99; if the effect was there, we would have detected it. It is possible that we failed to replicate Peterson's findings (while using his exact manipulation and materials) because the problems he had used in 1974 were too easy for our participants in 2004. This account is lent some credibility because the solution times we observed were shorter than those he reported: We saw mean solution times across conditions, ranging from 22 to 44 sec, as compared with the 29–79 sec that Peterson reported. Similarly, solution rates for our problems were on the order of 75%, whereas Goldman et al. (1992) reported between 17% and 50% correct solutions; it could be the case that problems must be more challenging and time consuming for an interruption to prove beneficial.

Table 1
Proportion of Anagrams Solved As a Function of Item and Participant Condition

| Incubation Condition | Frequent Words | | | | Infrequent Words | | | | | |
| | Two Moves | | Three Moves | | Two Moves | | Three Moves | | Total | |
	$\bar{X}$	SD	$\bar{X}$	SD	$\bar{X}$	SD	$\bar{X}$	SD	$\bar{X}$	SD
NI	.83	.23	.89	.17	.58	.32	.61	.32	.73	.18
I20−5	.85	.22	.90	.20	.65	.29	.60	.36	.75	.20
I60−5	.79	.25	.93	.13	.56	.31	.57	.33	.72	.17
I60+5	.83	.25	.87	.20	.67	.29	.63	.30	.75	.18
Total	.83	.24	.90	.18	.62	.31	.60	.32	.74	.18

Table 2
Solution Time (in Seconds) of Anagrams As a Function of
Item and Participant Condition

| | Frequent Words | | | | Infrequent Words | | | | | |
| | Two Moves | | Three Moves | | Two Moves | | Three Moves | | Total | |
Condition	$\bar{X}$	SD	$\bar{X}$	SD	$\bar{X}$	SD	$\bar{X}$	SD	$\bar{X}$	SD
NI	23.0	24.3	24.3	27.5	36.0	32.2	35.7	31.6	28.6	29.1
I20−5	24.5	27.6	23.4	26.2	41.9	33.9	39.8	33.3	31.0	31.0
I60−5	25.6	30.1	25.5	27.9	39.4	36.6	41.7	32.8	31.5	32.1
I60+5	24.2	25.8	21.8	25.0	41.2	32.4	43.9	33.9	31.4	30.5
Total	24.2	26.7	23.9	26.8	39.4	33.6	39.8	32.8	30.4	30.5

Figure 2 shows the cumulative proportion of anagrams solved during the 120 sec that the participants worked on these anagrams (plotted at intervals of 2 sec). Note that the 120 sec refers to a continuous 120-sec interval in the no-interruption (NI) condition, whereas in the other conditions, the 120 sec is broken at various points, depending upon the condition. For example, in the interrupted 20-sec blocks (I20–5) condition, there was a break after every 20 sec. One would expect a subconscious work mechanism to be manifest in increased solution rates for conditions involving interruptions, relative to no interruptions; however, we found no support of this. Nonetheless, such work may still have occurred at a rate so slow that it failed to produce significant effects, even with 233 participants.

To address concerns about possibly unjustified reliance upon acceptance of the null hypothesis, we calculated the minimum effect that we could have detected with a power of .8 and an alpha level of .05, given our subject count and observed variance. The minimum effect size would have

to be .37. This means that a period of interruption would have to result in the solution of one extra problem for only 3 out of 4 participants and Peterson's (1974) manipulation would have to speed up solution by only 5.8 sec (about one quarter of the difference he reported). So, although our results cannot unequivocally refute the presence of *any* benefit of interruptions in anagram solution, we can say that such benefits are minimal (if present at all).

EXPERIMENT 2

Although the results of Experiment 1 suggest that no (or only a very trivial amount of) subconscious progress was made in the solution of anagram puzzles, one might suggest that anagrams are less likely than certain other tasks to trigger, or benefit from, subconscious processes. Anagrams can, at least potentially, be solved by sequentially checking each possible permutation of the constituent letters. By contrast, incubation effects might be restricted to more unconstrained search problems that depend less
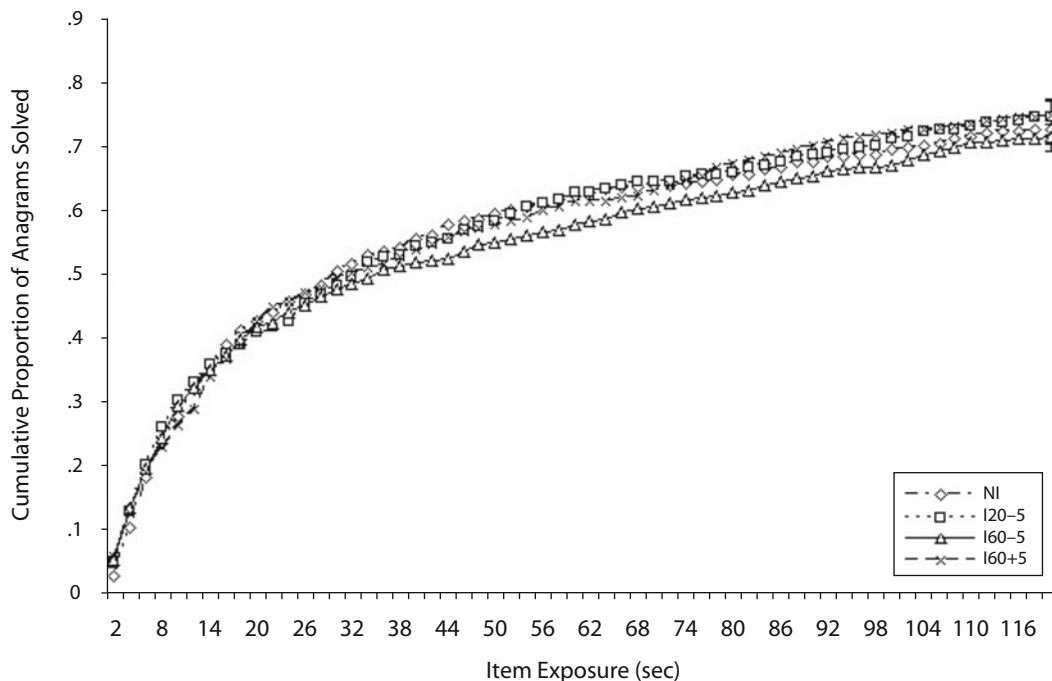

Figure 2. Cumulative proportion of anagrams solved in Experiment 1 as a function of explicit item exposure time (in seconds) and incubation condition.

upon explicit sequential hypothesis testing and search. For example, problems conducive to being solved by spreading activation in an associative network (Yaniv & Meyer, 1987) might seem especially likely to generate incubation phenomena. Experiment 2 addressed this possibility through the use of the Remote Associate Test (RAT). This task, unlike anagrams, is much less likely to be solved through a strategy of trial-and-error hypothesis testing but would, instead, appear to rely upon activation of multiple elements in semantic memory.

Experiment 2 addressed the question of whether incubation effects are more likely for this form of search task by replicating the basic design of Experiment 1, but using the RAT. In addition, we investigated the fixation-forgetting account of incubation effects (Smith & Blankenship, 1989, 1991) by including two new conditions involving artificially induced fixation. This manipulation was designed to increase the tendency of the participants to follow erroneous solution paths.

Method

Participants. A total of 300 new participants were recruited from the same online research panel as that in Experiment 1; no one participated in both experiments.

Materials and Design. Twelve RAT problems were chosen from a list of 144 on the basis of their ease of solution, as reported by Bowden and Jung-Beeman (2003). Of the 12 problems, 6 were "easy," with an average solution percentage after 30 sec of between 55% and 65% (mean: 60%), and 6 were difficult (average solution percentage after 30 sec of between 35% and 45%; mean: 40%). To serve as a test of basic RAT competence and cooperation with the task instructions, one RAT problem (*cottage*, *swiss*, and *cake*; answer: *cheese*) was selected on the basis of 96% accuracy after 15 sec, as reported by Bowden and Jung-Beeman.

A misleading associate was generated for each of the three words in every RAT problem. This word was selected to have an association with its paired RAT word in the target problem, but not with the answer or the other stimulus words (e.g., for the problem *tank*, *hill*, and *secret*, *water* would be paired with *tank*, *ant* with *hill*, and *hideout* with *secret*; each of these words would prime a meaning distant from the correct answer, *top*: *tank-top*, *hill-top*, and *top-secret*). The video game used as the interpolated task in Experiment 1 was used in Experiment 2 as well.

This experiment was a 2 × 6 mixed design, with RAT problem difficulty (easy vs. difficult) as a within-subjects factor and six levels of a between-subjects incubation condition. The four incubation conditions from Experiment 1 were included (NI, I20–5, I60–5, and I60+5), along with two new conditions involving artificially induced fixation, designed to replicate the findings in Smith and Blankenship (1991). The two new conditions—fixation with no interruption (FNI) and fixation with interrupted 60-sec blocks plus 5 min of unrelated task (FI60+5)—were created by adding an artificial manipulation of fixation to the NI and the I60+5 conditions, respectively. To induce fixation in these conditions, misleading associates were presented for 60 sec of the 120-sec total presentation time; the exact timing will be described below.

Procedure. The participants were first given instructions, along with two examples of RAT problems. The RAT problems in Experiment 2 were administered in the same way as the anagram problems in Experiment 1 for the NI, I20–5, I60–5, and I60+5 conditions. For all the presentations, the three RAT words were presented in a vertical column in bold uppercase letters. For the new conditions, FNI and FI60+5, misleading associates were shown in lowercase italics to the right of each of the three RAT words. Figure 3 illustrates the timing in the fixation conditions as described below.

1. Fixation, no interruption (FNI): After the video game had been played for 5 min, all of the RAT problems were presented in a random order for 120 sec each or until a solution was reached. Misleading associates were shown for the first 60 sec of each 120-sec presentation.
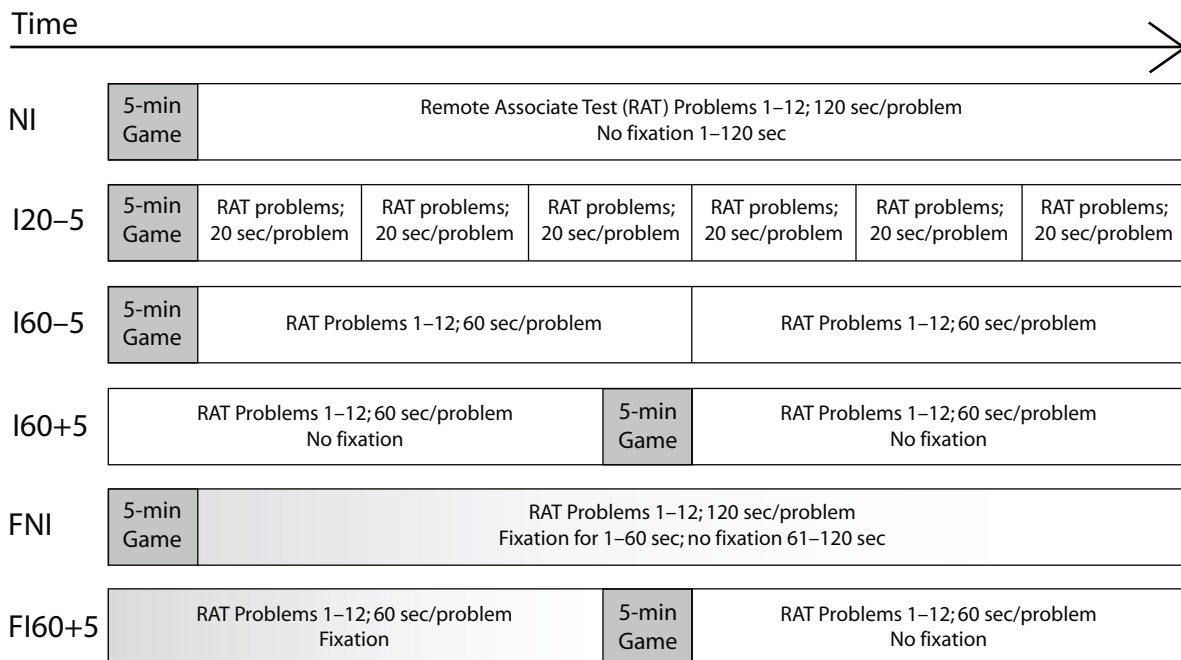
Time

Figure 3. Sequence of Remote Associate Test (RAT) and interpolated tasks for the six incubation and fixation conditions in Experiment 2.
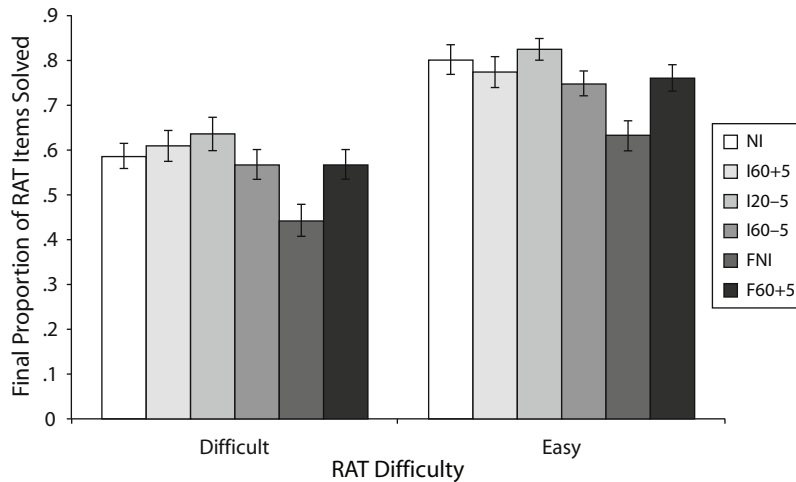
Figure 4. Final proportion of Remote Associate Test (RAT) items correctly solved in Experiment 2 as a function of incubation/fixation condition and RAT item difficulty.

2. Fixation, interrupted 60-sec blocks plus 5 min of unrelated task (FI60+5): All of the RAT problems were presented in a random order for 60 sec each (or until a solution was reached) with misleading associates shown. Then the participants played the video game for 5 min. Following that, the participants were given an additional 60 sec to solve each unsolved problem without misleading associates.

RAT task instructions for the two fixation conditions showed one of the example problems with misleading associates and explained that the words were presented "to make the task more challenging." This was explicitly stated, to avoid confounds in conditions in which the participants may have figured out that the associates were misleading and proceeded to ignore them.

Results and Discussion

A total of 338 participants completed the entire experiment. Of those, 38 did not answer the RAT competency question correctly and thus were dropped from further analysis. Of the 300 remaining participants, 68% participated alone from home; as in Experiment 1, environment did not have any detectable effect on performance [$F(5,299) = 1.274, MS_e = 0.042, p = .275$].

Total proportion correct was analyzed using a 2×6 mixed ANOVA with RAT item difficulty as a within-subjects factor and six levels of the incubation/fixation condition as a between-subjects factor. Figure 4 shows final proportion correct as a function of RAT item diffi-

culty and incubation condition. Difficult RAT items were solved less often than easy RAT items [$F(1,294) = 188.25, MS_e = 0.028, p < .001$]. Incubation/fixation condition had a significant main effect [$F(5,294) = 5.52, MS_e = 0.078, p < .001$], which was driven by significantly lower performance in the FNI condition, relative to all the other conditions (all Tukey's pairwise contrasts with FNI, $p < .05$; all other pairwise contrasts, $p > .4$).

To explore the effects of fixation and interruption, all other analyses were done on fixation with no interruption (FNI), fixation with interrupted 60-sec blocks plus 5 min unrelated task (FI60+5), and their no fixation counterparts: NI and I60+5. A $2 \times 2 \times 2$ mixed ANOVA was calculated, with total proportion solved at 60-sec (at the end of the fixation period for fixation conditions) and total proportion solved at 120 sec as two levels of a within-subjects factor (time) and with fixation and incubation as two between-subjects factors. Table 3 shows mean proportions solved at 60 and 120 sec and mean improvement scores for the two fixation and two incubation conditions. Time had a significant effect, with more problems solved after 120 sec than after 60 sec [$F(1,190) = 252.56, MS_e = 0.004, p < .001$]. The time × incubation, time × fixation, and time × incubation × fixation interactions were all highly significant [$F(1,190) = 7.23, MS_e = 0.004$,

Table 3
Improvement Scores From 60 to 120 sec As a Function of Fixation and Incubation Condition

| | Proportion Solved | | | | | |
| | At 60 sec | | At 120 sec | | Improvement | |
Condition	$\bar{X}$	SD	$\bar{X}$	SD	$\bar{X}$	SD
Fixation						
No incubation	.45	.22	.54	.22	.09	.07
Incubation (60+5)	.49	.20	.67	.21	.17	.11
No fixation						
No incubation	.60	.19	.69	.16	.09	.10
Incubation (60+5)	.61	.22	.69	.22	.08	.07

$p < .01$; $F(1,190) = 11.0$, $MS_e = 0.004$, $p < .005$; and $F(1,190) = 11.79$, $MS_e = 0.004$, $p < .005$, respectively]. This reflected the fact that the incubation conditions performed better at the 120-sec mark, fixation was most detrimental at the 60-sec mark, and the incubation condition with fixation showed the greatest improvement from 60 to 120 sec, relative to the other conditions.

A more thorough analysis was performed on improvement scores (performance at 120 sec − performance at 60 sec) between these four conditions. Figure 5 shows the cumulative improvement from the 60-sec mark as a function of time for each incubation and fixation condition. Tukey's contrasts between FI60+5 and all the other conditions showed that improvement from 60 to 120 sec was far greater in that condition than in any other condition (all $ps < .001$); all other pairwise contrasts were far from significant (all $ps > .9$). This indicates that an interruption increased the rate of improvement when the participants were provided with misleading clues, suggesting that artificially induced fixation was indeed forgotten during an interruption. However, improvement was identical in all the other conditions, meaning that we found no evidence confirming the hint of an incubation effect in conditions without artificial fixation seen by Smith and Blankenship (1991). Our power to detect the small effect they showed (0.39) was 0.8. From this, one of two conclusions must be drawn: Either (1) while working on the sorts of problems investigated here, participants do not become fixated independently, and thus there is no fixation to be forgotten during an interrup-

tion, or (2) such self-induced fixation cannot be forgotten during interruptions.

GENERAL DISCUSSION

Experiment 1 showed no improvement in average solution times when the time allotted for anagram problem solution was interrupted and opportunities for incubation were provided. This failure to replicate Peterson's (1974) finding was corroborated by a lack of an observable difference in the average number of problems solved in various incubation and nonincubation conditions. These null effects suggest that subconscious work was not taking place during periods of incubation. To assess how seriously these null results should be taken, an estimate was computed for a likely upper bound on the beneficial effect of an interruption, given our data. The conclusions suggested that it was most unlikely that anything more than a rather minor amount of subconscious progress could have occurred.

These conclusions were further supported in Experiment 2, using RAT problems, where, again, incubation showed no benefit under normal conditions (i.e., in the absence of misleading clues). The use of RAT problems in the second experiment was designed to make spreading activation in an associative network more helpful to problem solution. Nonetheless, we found no evidence supporting this account of incubation.

Experiment 2 also showed that periods of incubation are quite helpful in allowing an artificially induced fixation to dissipate. Whereas Smith and Blankenship (1991)
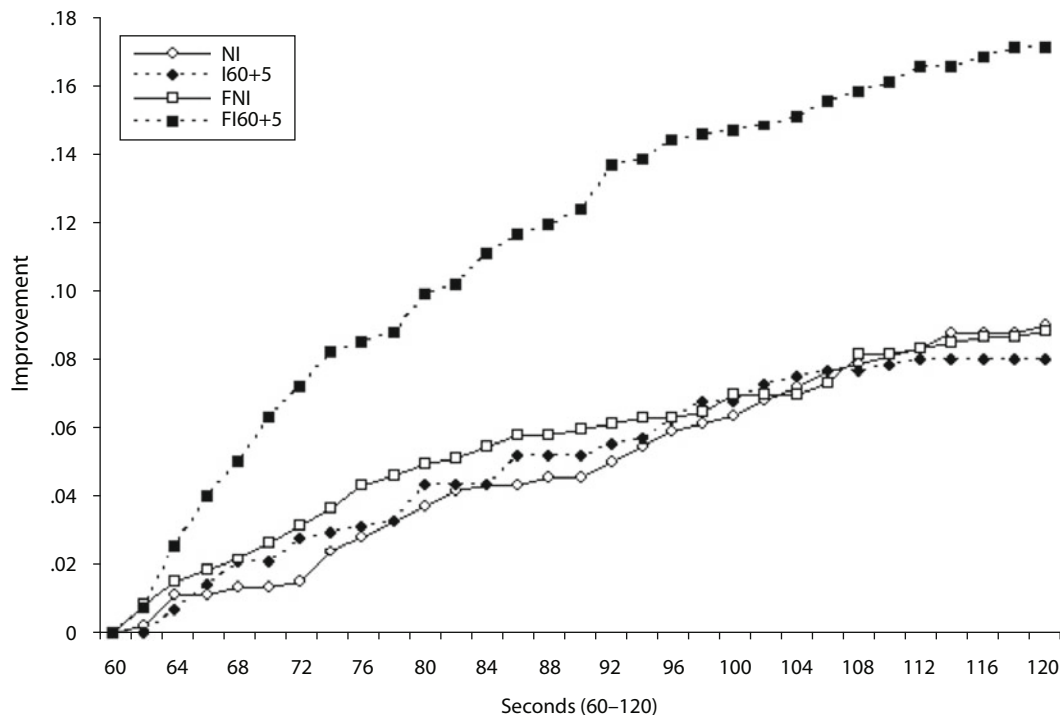

Figure 5. Cumulative improvement (proportion solved) after 60 sec as a function of time (in seconds) and incubation/fixation condition (Experiment 2).

found a slight, albeit nonsignificant, benefit of incubation when participants were not misled, our study showed no such trend, even with sample sizes large enough for the effect to reach significance. A detailed analysis of improvement rates during the time when incubation would have taken place (and fixation presumably dissipated) showed that both incubation and artificially induced fixation were necessary to produce an increased improvement rate. This might be explained in one of two ways: Either the participants did not become fixated when left to their own devices, or if they did, self-induced fixations were not overcome by a period of incubation. It could be, for example, that a self-induced fixation will tend to recur even after incubation (perhaps because whatever caused it in the first place remains present), with the incorrect solution path that someone was drawn to prior to the interruption period proving just as alluring after this period. Wiley (1998) found evidence of just such a phenomenon when fixation induced by prior knowledge was not relieved by interruptions, whereas fixation induced by the experimenter was relieved.

In sum, the results are generally consistent with the possibility that problem-solving cognition is subject to the same central-processing bottleneck as that revealed in dual-task studies involving comparatively more trivial choice reaction time tasks and memory retrieval processes. Furthermore, no evidence was found of any processes that could be reliably employed to the benefit of problem solvers: An interruption in work had a positive effect only when the participants were purposefully misdirected and was ineffective under unobstructed problem-solving conditions.

Conclusions and Limitations

These results suggest that at least with the kinds of problems and ranges of time intervals investigated here, incubation probably does not occur to any substantial degree. What is one to make, then, of the ever-popular anecdotes claiming the reality and power of incubation? On the one hand, these anecdotes could simply be erroneous. It is possible that individuals who think that some problem solution occurred to them while they were not thinking about the problem had, in fact, resumed conscious work. It is also possible that people mistake other phenomena for incubation. For example, interruption may relieve fatigue from long periods of concerted effort, making subsequent efforts more productive. Alternatively, people may sometimes encounter clues (or even the actual problem solution) during such uncontrolled interruptions. Given the design used in the experiments reported here, it is not likely that recovery from fatigue would be mistaken for incubation, whereas in ordinary life such a confusion might be relatively common.

Although either of these scenarios might very well be responsible for anecdotal cases of incubation, neither interpretation would provide any warrant for advising people to cease work on important problems as a way to achieve a solution more rapidly. As was noted in the introduction, precisely this advice is commonly offered in many books purporting to help people become more effective problem solvers.

On the other hand, it would be tendentious and perhaps naive to claim with any certainty that these experiments overturn the observations of thinkers such as Poincaré and Feynman. Like other experimental research on the topic, the experiments presented here have looked at simple problems that have occupied the attention of participants for only a matter of minutes, rather than weeks, months, or years, as would have been the case for major creative achievements. It is quite possible that prolonged and intensive engagement in a particular problem may unleash cognitive processes different than those evoked by the more mundane and casual involvement examined in experimental situations such as ours. It is also possible that performance of problems involving *insight*, rather than mere search or triangulation in semantic memory, as in the present experiments, might possibly benefit more from some form of subconscious work. An important challenge for future research in this area will be to make more prolonged and intensive immersion in a problem amenable to scientific study. For obvious practical reasons, this promises to be a difficult challenge.

REFERENCES

Barron, F. (1988). Putting creativity to work. In R. J. Sternberg (Ed.), *The nature of creativity* (pp. 76-98). Cambridge: Cambridge University Press.

Birnbaum, M. H. (1999). Testing critical properties of decision making on the Internet. *Psychological Science*, **10**, 399-407.

Bowden, E. M., & Jung-Beeman, M. (2003). Normative data for 144 compound remote associate problems. *Behavior Research Methods, Instruments, & Computers*, **35**, 634-639.

Carrier, L. M., & Pashler, H. (1995). Attentional limits in memory retrieval. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **21**, 1339-1348.

Gavurin, E. I. (1965). Effect of distributed practice upon initial solution to a jigsawlike puzzle. *Perceptual & Motor Skills*, **21**, 892.

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

Goleman, D., Kaufman, P., & Ray, M. (1993). *The creative spirit.* New York: Plume.

Higgins, J. M. (1994). *101 creative problem solving techniques: The handbook of new ideas for business.* Winter Park, FL: New Management.

Krantz, J. H., & Dalal, R. (2000). Validity of Web-based psychological research. In M. H. Birnbaum (Ed.), *Psychological experiments on the Internet* (pp. 35-60). San Diego: Academic Press.

McGraw, K. O., Tew, M. D., & Williams, J. E. (2000). The integrity of Web-delivered experiments: Can you trust the data? *Psychological Science*, **11**, 502-506.

Murray, H. G., & Denny, J. P. (1969). Interaction of ability level and interpolated activity (opportunity for incubation) in human problem solving. *Psychological Reports*, **24**, 271-276.

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.

OSBORN, A. F. (1953). *Applied imagination: Principles and procedures of creative thinking.* New York: Scribner.

PASHLER, H., CEPEDA, N., WIXTED, J., & ROHRER, D. (2005). When does feedback facilitate learning of words? *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **31**, 3-8.

PASHLER, H., & JOHNSTON, J. C. (1998). Attentional limitations in dual-task performance. In H. Pashler (Ed.), *Attention* (pp. 155-189). Hove, U.K.: Psychology Press.

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

RAY, M., & MYERS, R. (1989). *Creativity in business.* New York: Doubleday.

REIPS, U.-D. (2002). Standards for Internet-based experimenting. *Experimental Psychology*, **49**, 243-256.

SEIFER, 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). Getting into and out of mental ruts: A theory of fixation, incubation, and insight. In R. J. Sternberg & J. E. Davidson (Eds.), *The nature of insight* (pp. 229-251). Cambridge, MA: MIT Press.

SMITH, S. M., & BLANKENSHIP, S. E. (1989). Incubation effects. *Bulletin of the Psychonomic Society*, **27**, 311-314.

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

THORNDIKE, E. L., & LORGE, I. (1944). *The teacher's word book of 30,000 words.* New York: Columbia University, Teachers College.

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

WELFORD, A. T. (1967). Single-channel operation in the brain. *Acta Psychologica*, **27**, 5-22.

WILEY, J. (1998). Expertise as mental set: The effects of domain knowledge in creative problem solving. *Memory & Cognition*, **26**, 716-730.

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

NOTE

1. Some authors use the term *incubation* to refer to the idea of subconscious work; for various reasons, we prefer to use the term more generically and refer to subconscious work as one possible *explanation* for whatever incubation might occur, if it does in fact occur.