BOOTH, RACHEL A. M.A. Does Mind Wandering Create A Mental Context Change That Facilitates Memory Retrieval? (2022)
Directed by Dr. Michael J. Kane. 61 pp.

This study tested whether task-unrelated thoughts (TUTs) can benefit memory retrieval by inducing mental context changes in the presence of output interference. Specifically, Study 1 examined the effects of an incubation period on recall in a long-duration verbal fluency task (recalling animal names). Study 2 replicated Study 1, but with longer task times. Both studies also assessed whether TUT rates during an incubation period predict post-incubation recall output, as would be expected if TUTs create mental context changes that reduce output interference between recall periods. Study 1 tested 204 participants, and Study 2 tested 211, and both presented a recalling-animals verbal fluency task and an incubation-period reading task with embedded thought probes to assess TUTs. In Study 1, participants recalled animals for 7 min pre-incubation and 2 min post-incubation, whereas in Study 2, participants recalled animal names for 9 minutes pre-incubation, and incubation-task duration increased from 11 to 12.5 min. Neither experiment showed the predicted incubation effects, nor a correlation between TUT rates and post-incubation fluency; only Study 2 showed a short-lived incubation effect in an exploratory analysis.

DOES MIND WANDERING CREATE A MENTAL CONTEXT CHANGE THAT FACILITATES MEMORY RETRIEVAL?

by

Rachel A. Booth

A Thesis
Submitted to
the Faculty of The Graduate School at
The University of North Carolina at Greensboro
in Partial Fulfillment
of the Requirements for the Degree
Master of Arts

Greensboro

2022

Approved by

____

Dr. Michael J. Kane
Committee Chair

APPROVAL PAGE

This thesis written by Rachel A. Booth has been approved by the following committee of the Faculty of The Graduate School at The University of North Carolina at Greensboro.

Committee Chair	Dr. Michael J. Kane
Committee Members	Dr. Dayna Touron
	Dr. Peter Delaney

November 21, 2022
Date of Acceptance by Committee

November 21, 2022
Date of Final Oral Examination

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER I: INTRODUCTION

Sometimes, when trying to remember something, a person may experience the feeling that they know the answer but can't quite name it. This "tip of the tongue" (TOT) experience is often solved by stepping away from the problem and changing one's mental context. Mind wandering may be one way to create such a mental context change, affording people the opportunity to reach remote solutions after they have reached impasse in retrieval.

Retrieval Impasse and Incubation

Impasse is a point at which a person has been searching for a solution and is no longer making progress. When a person reaches impasse during retrieval, they may or may not feel the answer is accessible in memory. This is separate from, but similar to, the TOT phenomenon (Brown and McNeill, 1966) because an impasse need not be accompanied by a feeling of knowing the answer (Hart, 1967). Taking a break from attempted retrieval to reach more or better solutions is called an incubation period. This incubation period can help overcome TOTs and other forms of impasse.

For example, Choi and Smith (2005) gave participants descriptions of items or pictures of people to name that were likely to elicit a retrieval impasse with a TOT sensation. If subjects were unable to name an item, they either had to immediately try again (in the control group) or spend 17 to 25 min working on other descriptions until they could try the missed descriptions again (in the experimental group). Those who tried the missed problems again after a delay resolved 15% more missed problems than did those who tried again immediately. Therefore, Choi and Smith demonstrated that incubation when experiencing a TOT assists in TOT resolution.

Though the present study did not examine TOTs specifically, it examined retrieval after impasse, and the TOT literature indicates that an incubation period facilitates correct retrieval after an individual has exhausted their initial memory search. Incubation periods are most widely studied in the problem-solving literature, and so looking to that field should inform our understanding of retrieval impasse.

Incubation and Impasse in Problem Solving

During problem solving, people form an initial strategy to achieve solution, and will typically attempt that strategy multiple times, even if it does not help them with solution (Smith and Blankenship, 1991). This leads them to impasse, in which there is no apparent progress being made towards solution. Several reviews indicate that an incubation period often helps people overcome impasse and reach solution. Dodds et al. (2004) examined 39 experiments investigating incubation and problem solving, and found that 26 demonstrated incubation effects, with studies that used an incubation period of between 15-30 min being most likely to report significant incubation effects. Likewise, Sio and Ormerod (2009) meta-analyzed 117 empirical studies using incubation during problem solving and found that 85 reported positive incubation effects (overall unweighted mean effect size, $d = .41$). There were a wide range of effects examined, with two particularly relevant to the current study. First, incubation periods benefitted creative problems where participants need to extend their mental search of their knowledge base and widen that search as time continues. Second, problems with verbal materials, which also likely involved memory (e.g., from the Remote Associates Test; see below) particularly benefited from low cognitive demand tasks during the incubation period.

“Fixation” refers to a particular form of impasse in problem solving, a mental block that arises where one is stuck at that initial starting point or on early solution attempts (Beeftink, 2008; Smith & Blankenship, 1991). Koppel and Storm (2014) argue that overcoming fixation can be done through inhibition of inappropriate responses, as measured by retrieval induced forgetting (RIF). RIF happens when inhibition has blocked non-target responses during retrieval practice from interfering with target responses on a final recall test. A common measurement of RIF is presenting a series of category pairs to participants. These pairs will have a category first, and then an exemplar, such as *fruit-strawberry*, then *fruit-grape*. The series also includes pairs other than fruit. Participants will only practice retrieving half of the series, so *fruit-strawberry* but not *fruit-grape*, and not the other categories. On the final test, participants recall all pairs in the series. RIF has occurred when *grape* is not recalled as well on the final test as completely unpracticed pairs, like *animal-dog*.

Koppel and Storm (2014) inserted incubation into RAT problems for half of their participants to see how solution correlated with RIF, assessed separately. Under normal non-incubation conditions, RIF was significantly correlated with RAT performance (r = .32), indicating that participants' ability to inhibit wrong solutions was linked to better RAT performance. For incubation participants, however, RIF no longer correlated with their problem-solving performance (r = -.01). Therefore, Koppel and Storm (2014) drew the conclusion that incubation replaced the need for inhibition because incubated participants became less fixated than non-incubated participants. This study shows how incubation can help alleviate fixation, which builds the argument that incubation will be helpful for impasse during retrieval; it also shows how incubation can be related to useful forgetting, which is expanded upon further below.

Mental Context Change

What might be underlying these incubation effects is a mental context change. The history of context's role in memory research is long and well documented (for reviews, see

Roediger et al., 2017; Smith and Vela, 2001). We use context to retrieve, and so it makes sense that by changing context, one also changes retrieval success. Overall, the literature shows that matching or mismatching contexts at encoding and retrieval either assists or harms retrieval, respectively. Being in the same environmental context at both encoding and retrieval leads to better recall at test than being in, or mentally simulating, separate contexts at encoding and retrieval. These same physical match and mismatch findings can be replicated by state or mood dependency as well. Therefore, it follows that a mental context change should impact retrieval, which Sahakyan and Kelley (2002) demonstrated.

Sahakyan and Kelley (2002) used the notion of mental context change to explain the costs and benefits of directed forgetting. Directed forgetting refers to when people intentionally forget information on command, so they are less able to recall it later; the forgotten information is also less likely to interfere with subsequently learned information. Directed forgetting is typically studied by having participants learn List 1, then asking them to forget or remember that list, then having them learn List 2, and then testing memory for both List 1 and 2. Directed forgetting "costs" refer to worse recall of List 1 for subjects instructed to forget it versus subjects instructed to remember it. Directed forgetting "benefits" refer to better recall of List 2 for subjects instructed to forget List 1 versus subjects instructed to remember List 1.

Sahakyan and Kelley (2002) presented a list of words to participants (List 1), then instructed them to either forget or recall that list. Then, half of participants in both the forget and remember conditions were given a mental context change, by imagining for 45 seconds what they would do if they were invisible; the other half waited for the same span of time for List 2 to be presented. Finally, all participants studied List 2 and then recalled both lists. For participants who were instructed to remember, those who had a context change recalled significantly fewer words from List 1 and significantly more words from List 2 than did participants who didn't experience a context change. These context-change effects were similar in magnitude to those for participants who received only a directed-forgetting instruction, suggesting that a mental context change can reduce the likelihood of interference from previous information during recollection. Indeed, Delaney et al. (2010) found that manipulations producing more dramatic mental context changes (e.g., imagining a foreign trip) created larger directed-forgetting-type effects than did less dramatic context changes (e.g., imagining a domestic trip).

The idea that a mental context change can replicate directed-forgetting effects was built upon by Smith and Beda (2020), who examined the effect of mental context change on fixation using the RAT problem-solving task. They first had participants memorize 24 word-triads that were displayed on the computer, overlaid on a photo (one photo per triad); participants were then tested on their memory for those triads. Finally, the participants were given the RAT, which used different triads than the fixation task, and each problem was given twice, no matter if at the first presentation the response was correct or incorrect; at the second presentation, the problems were either overlaid on the photo from the fixation task, or over a new photo. Also for the second presentation, participants were either immediately shown the problem again, or were shown after a delay. Participants solved more unresolved problems after a delay than after immediate retesting, and this incubation effect was even stronger for problems shown over a new photo (i.e., after a context change) than over the old photo. Smith and Beda interpreted these results as participants needing both an incubation period and a context change to perform the best on retested, unresolved problems. Although incubation alone relieved fixation, incubation was even more useful when it also provided a clear mental context change.

Mind Wandering as a Mental Context Change

In empirical research, mind wandering is frequently defined as task-unrelated thoughts (TUTs) during an ongoing task (Kane & McVay, 2010). TUTs are measured in variety of ways, though primarily via thought probes inserted during tasks that ask participants whether they were just thinking about the task or about something else. That 'thinking about something else' is why mind wandering might act as mental context change. Recall that Delaney et al. (2010) found that the greater the mental context change experienced during an imagination task, the poorer was one's memory for the pre-imagination events. That is, the geographically farther the place participants thought of (so the more unrelated they were to the present context) the less they could recall items from a present geographical place. They argued that daydreaming had effects on memory due to context change.

As discussed earlier, reviews of the incubation literature noted that, generally, a low cognitive demand incubation task increased incubation effects (Sio & Ormerod, 2009). Mind wandering increases in low demand tasks compared to high demand tasks, as well (e.g., Smallwood et al., 2009). Therefore, low-demand incubation periods may facilitate mind wandering, which then may facilitate a mental context change.

In a similar line of thinking, Baird et al. (2012) found that mind wandering during incubation was related to successful subsequent creative thinking. Baird et al. (2012) asked participants to work on the unusual uses task (UUT), which measures how many creative (i.e., unique and apt) uses for a common item, such as a brick, that participants can generate within a brief time period. For example, "to build a wall" is not a creative response, but "to hold over your head to protect from rain" is. After generating ideas for 2 min, participants either had no incubation, a difficult incubation task, or an easy incubation task, meant to facilitate mind wandering. The authors proposed that this easy task should increase mind wandering because it requires less attention and so, with increased mind wandering, people should be better able to reach remote solutions (they did not theorize what underlies this process, but rather encouraged further study of how and why this phenomenon occurs). Participants in the easy condition retrospectively self-reported significantly higher levels of mind wandering during incubation on a 1–5 scale ($M = 2.47$, SD = 0.66), than did participations in the demanding condition ($M = 2.15$, SD = 0.67); they also performed 40% better on the UUT than the demanding condition, and 50% better than the no break condition. These findings thus suggested that mind wandering during incubation bolsters creative problem-solving performance.

However, several conceptual or direct replication attempts of the Baird et al. (2012) study have not found an association between incubation mind wandering and creative cognition. First, Smeekens and Kane (2016) found that TUT rates during an incubation task, measured via thought probes rather than a post-incubation rating, were not correlated to post-incubation creativity scores in the UUT (or pre-to-post increases in creativity). Steindorf et al. (2020) attempted to replicate both Baird et al. (2012) and Smeekens and Kane (2016). Ultimately, they too could not establish a relationship between mind wandering and creativity because TUTs during incubation were not correlated with creativity performance after incubation. The most recent replication effort has come from Murray et al. (2021), but again, they did not find a relationship between TUTs and creativity.

Divergent thinking is accomplished through many different cognitive processes, a prominent one of which is memory retrieval. Before an individual can think of creative uses for an object, they first must retrieve the object and its typical uses from memory. And so, if TUTs create a context change that facilitates memory, then divergent thinking may be too complex a task to use for examining the question. Instead, a somewhat simpler memory task might be sensitive enough to demonstrate any possible mental context change effect on retrieval.

Examining Impasse and Incubation in Fluency Tests of Memory Retrieval

The present study addressed the controversy about mind wandering and creativity by taking a step back to ask whether memory retrieval, reflecting a simpler set of processes that only partially contribute to creative idea generation (e.g., Beaty et al., 2014; Gilhooly et al., 2007), shows an incubation benefit driven by TUTs. As discussed earlier in this introduction, fixating on a strategy that is not helping to retrieve new items leads to can lead to impasse, and so mental context change during an incubation period might combat that impasse.

Impasse, within the memory literature, is most closely linked to output interference. Output interference occurs when, during recalling items from a set, those first recalled from the set block retrieval of to-be-recalled items in the remainder of the set (Smith, 1995). Smith and Vela (1991) showed that an incubation period relieves this output inference compared to continuous recall. Across three experiments, their participants memorized a list of objects for later recall, and then, when they suffered from output interference upon the first recall period, an incubation period inserted between their first and second recall period helped participants name previously unrecalled items in the second period.

A task that has been used repeatedly to elicit output interference has been the naming animals task. The use of naming animals as a measure of verbal fluency was first described by Thurstone (1938) in his book on mental abilities. For this task, participants simply name as many animals as they can in a given time period. The time between retrieved items increases as recall proceeds, even within the first few minutes of the task (Rosen and Engle, 1997). People could not possibly have exhausted their knowledge base of animals so soon, which indicates the effects of output interference. In many ways, verbal fluency is a building block to divergent thinking tasks, like the UUT task used in Baird et al. (2012). Both are cue-based, and both have participants naming as many “category” exemplars as possible. The UUT involves an ad hoc category (e.g., uses for an object) but also adds a step, to creatively alter those items retrieved from memory. Both kinds of task, however, start with item generation from memory.

Smith et al. (2017) asked participants to retrieve as many items as possible from memory for two different categories. In one group, participants named items from one category for three minutes uninterrupted before moving onto the second category. In the incubation group, participants alternated naming items between the two groups every 60 s. The switching from one group to another served as the incubation manipulation, as the authors hypothesized that forgetting one list while recalling in another works as incubation. The switching-categories group named more items in both categories than did the continuous group. This is a demonstration of evidence for the forgetting-fixation/context-change hypothesis, which argues that incubation is helpful because it allows people to forget their fixation and access new ideas, such as new words to recall.

Goals and Hypotheses

The current study examined whether an incubation period can facilitate overcoming an impasse in memory retrieval, especially (or only) when the incubation task promotes mind-wandering experiences (i.e., TUTs) that can create mental context changes. Experiment 1 tested whether an incubation period during retrieval benefits subjects more than completing a retrieval task without incubation (and Experiment 2 replicated Experiment 1 with a few methodological changes). Subjects completed a verbal fluency task wherein they named as many category items as they could in a timed setting (i.e., for 7 min). They either had an 11 min incubation task (reading a passage) inserted, or continued to the final retrieval block with no break. During incubation, TUTs were measured via ten thought probes at randomized durations. Participants' fluency was measured by number of new items recalled in the final 2-min block of the retrieval task.

The reading task used in the incubation condition was also given to the no-incubation group *after* retrieval. Therefore, I probed for TUTs during the reading task in both conditions. Also, strategy questionnaires were given to all participates after both retrieval blocks, asking for the first and second retrieval block strategies. If incubation relieved fixation in between retrieval blocks, then participants may have chosen a new, previously unthought of strategy in the second block.

I sought to determine whether an incubation period is helpful to retrieval because it relieves interference by creating a mental context change; if the reading task is an effective incubation period, then individuals who had an incubation period during retrieval would outperform individuals who did not have an incubation period during retrieval on fluency measures. As well, if incubation produces a mental context change via mind wandering, then within the incubation condition, TUT rates should positively predict post incubation retrieval, whereas within the no incubation condition, rates of mind wandering (after the fluency task) should not predict post incubation retrieval.

CHAPTER II: METHODS

Experiment 1

Participants

Two hundred four participants, aged 18-40 years, from the University of North Carolina at Greensboro and Prolific Academic (payrate $10.82/hour), were tested online. One hundred two participants were randomly assigned to the incubation condition and 102 to the no incubation condition. Sample size was determined by G*power (Faul et al., 2007) using $\alpha = .05$, 2-tailed, for a *t*-test contrasting Block-2 retrieval fluency between the incubation and no incubation groups, with a target effect size of $d = .40$. With 102 subjects in each group, the study design had 80% power to detect an effect size of Cohen's $d = 0.39$, 90% power for $d = .45$, and 95% power for $d = .50$. For a 1-way ANCOVA, which will follow-up the *t*-test with fluency in the first 7 min of retrieval (i.e., Block 1) as a covariate, 102 subjects per group will provide 80% power to detect $f = 0.20$, 90% power for $f = 0.23$, and 95% power for $f = 0.26$. And finally, at 102 subjects per group, the study had 80% power to detect a TUT-by-fluency correlation within each group of $r = 0.27$, 90% power for $r = 0.31$, and 95% power for $r = 0.34$.

Tasks and Measures

All tasks were programmed using Labvanced (Goeke et al., 2017).

Retrieval Fluency

Participants recalled and reported as many animals as they could, without repetition. After typing each animal name into a response box, they pressed a button onscreen to submit their response and the box was cleared. There were two separate retrieval assessments (blocks) with different timing parameters, explained under Procedure. The primary dependent variable for each block was the number of unique animal names retrieved.

After participants received their fluency-task instructions, they took a compensation instruction quiz. It asked what reward they would receive based on their performance, and to continue, they had to correctly select the answer that they would not be rewarded based on performance. The goal of this quiz was to reduce motivation to cheat during the fluency task by looking up animals online, for example. If they answered incorrectly, they were shown a screen saying “Incorrect. Remember we are interested in how people think not how well you can do. Please do not use outside resources.” and were taken back to the quiz to reattempt. The instructions for the fluency task alongside the cheating reduction measures are provided in Appendix A.

The first fluency task block lasted 7 min, and the second for 2 min. An informal review of the literature containing time course intervals for fluency-task retrieval suggested that by roughly 7 min, retrieval has slowed down considerably on average, suggesting some degree of impasse (Bousfield & Sedgewick, 1944; Fitzgerald, 1983; Gronlund & Shiffrin, 1986; Kail & Nippold, 1984). Furthermore, whereas participants in both Gruenewald and Lockhead (1980) and Rosen and Engle (1997) had decreasing recall after the 7-min mark, on average, they were still able to produce some new animal names for the rest of their 10- or 15-min tasks.

The second retrieval block of 2 min should have provided the incubation group with enough time to name new animals they were able to think of post-incubation, while not pushing the no-incubation group into giving up on the study out of frustration, tiredness, or boredom.

Strategy Assessment

I assessed strategy use during the fluency task with a checklist created by Schelbe et al. (2012). The 16-item checklist (in Appendix B) asked participants if they used any of the common strategies listed, such as mentally searching for sub-types of animals, rhyming animal names, or thinking of different animal locations or habitats. Participants responded by clicking a checkbox beside the applicable strategy. Both groups received this checklist after the second retrieval task block only. It was then presented twice, first to ask about the 2-min naming animals block (Block 2) they *just* competed, and then again to retrospectively ask about the 7-min naming animals block (Block 1) they *first* completed.

Incubation Passage Reading

All participants read a modification of a personal essay, *Lazy Thoughts of a Lazy Girl,* by Jenny Wren (1891), presented one sentence at a time on-screen; participants in the incubation condition read it between the two retrieval blocks, and participants in the no-incubation condition read it after both retrieval blocks were completed. The sentences were presented in Lato font, size 24, centered on screen, in black text against a light grey background. The essay focused on the writer's preference for dogs over children and was modified (by removing words or entire sentences) to shorten it while preserving the meaning and not altering any text directly speaking to the narrator's comparison of dogs to children. The 85 presented sentences were 6–20 words long (M = 14.35, SD = 2.12) and had a mean of 6.7 Flesch-Kincaid reading level.

According to Sio and Omerod's (2009) meta-analytic review of incubation effects, this type of reading task reflects an "easy" task, and therefore should be conducive to producing incubation effects. Because the passage refers often to dogs, it should have helped maintain the "animals" context from the animal-recalling fluency task (following the logic of Beda & Smith, 2020). Previous research shows that dogs are one of the first animals named (Kail & Nippold, 1984) and so, despite supporting the prior mental context, the passage should not have given any additional cues to participants for the post-incubation recall period.

Time Estimate

Immediately after reading the incubation passage, all participants were asked to give their estimate, in minutes, for the reading task duration. Sahakyan and Smith (2014) demonstrated that participants who underwent a mental context change significantly overestimated the time their experiment took compared to participants who did not experience a mental context change. Participants typed in their free response estimate numerically.

Comprehension Task

As a check to assure subjects were reading carefully, they were told before the reading task to expect a comprehension test. After they complete the reading task, they answered four multiple-choice questions by pressing the number assigned to their answer choice (see Appendix C). The number of correctly answered questions was counted as the DV.

Thought Probes

To measure TUTs, both groups received ten thought probes on-screen, based on McVay and Kane (2012), during the reading task. Participants had 7 s to answer each probe; even if the probe was answered before 7 s, participants were told that it would remain onscreen for the full duration. This ensured that the total incubation time for the reading task was the same for all participants. The probes' placement was randomized through the passage (with the constraints of a minimum of four sentences in between probed sentences, and no repeats among the three counterbalancing conditions), and each participant saw one of three randomized versions of probe placement. With three probing placement options, and ten probes, thirty out of the eighty-five sentences were followed by a probe across subjects.

Participants received probe instructions onscreen, which highlighted that they needed to report what they had *just* been thinking about:

> It is perfectly normal to sometimes think about things that are not related to the reading task. For instance, you may think about off-task things such as something you did recently or will be doing later, your current emotional or physical state, personal worries, daydreams, or your external environment. Please try your best to honestly assess your thoughts and choose a response that best describes the thoughts you were JUST having in the INSTANT before we asked.

Participants typed the corresponding number to select their answer from seven options presented onscreen. The options, alongside an explanation of each option, are listed below. Participants only saw the full explanations (below in parentheses) during the instructions. During the task, participants only saw the numbers listed (1-7) and matching phrase next to each number (e.g., "5. Daydreams").

1. On-task on the passage (for thoughts about what was written in the passage at that time);
2. How well I'm understanding the passage (for evaluative thoughts about comprehending [or not] what was being presented on-screen);
3. Everyday personal concerns (for thoughts about normal everyday things, life concerns, or worries);
4. Current state of being (for thoughts about one's current physical, psychological, or emotional state [e.g., thinking about being sleepy, hungry, or fascinated]);
5. Daydreams (for fantasies or unrealistic thoughts);
6. External environment (for thoughts about anything present or happening in the room, aside from the reading task).
7. Other (for any thoughts not fitting into the above categories).

To continue with the study beyond the thought-probe instructions, participants had to correctly answer an instructions quiz. They were asked what timeframe across which they should report thoughts when given a probe. They were shown three options and had to correctly select the option to only report thoughts *just* before the probe to continue; if they selected an incorrect response, they were brought back to the probe instruction quiz to retake until they achieved the correct answer.

I coded any response of 3–7 to be a TUT. The proportion of probes on which a subject reported a TUT was the DV.

Inclusion Questionnaire

After completing the experimental tasks, participants filled out a demographics and inclusion questionnaire. For demographics, they responded with their age, gender, race/ethnicity, and education level. For the inclusion questions, they answered four questions about their environment during the study, such as how noisy their environment was, how distracted they were, how often they interacted with other multimedia, and how sleepy they were. The questions had a response scale of four options: not at all, somewhat, very, and extremely. The multimedia-use question was broken into three sub-questions, asking for their levels of interaction with: (a) their phone for texts and calls, (b) video games, and (c) any form of social media, during the experiment. The multimedia-use question had a response scale of four options: never, some of the time, most of the time, and all of the time. Finally, they answered a yes/no question about if they used any outside resources during the either of the naming-animals task Blocks.

Procedure

After participants were told they would be undergoing a study on human memory and that they will not be rewarded based on performance (to disincentivize “cheating” by looking up animals on-line, for example), they were asked to affirm a non-cheating statement, where they selected yes, that they would complete the study alone, or no, that they would complete the study with outside help. If they selected yes, they would use help, they were removed from the online study immediately, and so their data were not collected.

In the first block of the fluency task, subjects typed animals one at a time for 7 min and then pressed the enter key to submit each answer. Their task instructions indicated that they were to name items from a category but did not reveal the category until the task began, so that participants were not able to start formulating responses.

After the first 7 min, the procedures for the incubation and no-incubation groups diverged. The incubation group silently read the incubation passage, with instructions to read for a comprehension test (recall that they also had to correctly answer a thought-probe instructions quiz before proceeding). The passage was presented sentence by sentence, at an automated 7 s pace. Participants saw 10 thought probes on-screen. The reading task took 11.10 min, and then participants completed the comprehension test.

Incubation participants then responded to the verbal fluency prompt once more, for an additional 2 min. They were given instructions to not repeat animals they had named in the first phase of this task, and to name only new animals. They then completed the retrieval strategy questionnaire twice, the first time answering for the retrieval block they had just completed, and the second time answering for the first retrieval block.

The no-incubation group, after completing the first 7 min of the retrieval task, did not have an incubation period. Instead, they were instructed to retrieve animals for an additional 2 min, with no break other than to allow for task instructions (which matched those for the incubation subjects). At the end of the second retrieval block, they then completed the retrieval strategy questionnaires, in the same manner as the incubation group, and then the reading task with thought probes and comprehension test.

To end the study, all participants reported if they used outside help or not.

Results

For all analyses, I report null hypothesis significance tests with an alpha of .05.

Data Analysis Exclusions

For both experiments, subjects' data were excluded from all analyses if three or more TUT probes were unanswered before timing out, or if two or more reading comprehension answers were incorrect. Further, participants' data were dropped from analyses if either: (a) they responded "extremely" to at least two of the three post-experiment inclusion questions regarding their environment, or (b) they responded to any of the three the media-multitasking questions with "Most of the time" or "All of the time." Finally, if they responded to the final honesty question with "Yes, I did use outside help," their data were dropped.

Of the 250 participants tested, data from 46 were dropped for one of the exclusion criteria above, resulting in 204 usable participants (102 in each experimental condition; 90 women; *M* = 31.6 years old). Data from all 46 were dropped for getting 2 or more reading comprehension check questions wrong; out of the remaining subjects, all had abided by the required probe responses, did not report severe distraction, and did not report using outside recourses. This makes sense, because if participants were paying close enough attention to the reading task to answer the comprehension questions correctly, then they would have also seen all probes during the reading task, would likely not have been distracted, and likely were dedicated enough to not cheat on the recall task. Demographics for included participants are reported in Table 1.

Recall-Task Scoring

Recall of unique, non-repeated animals in the fluency task was counted by me. I used the criteria provided by Welhaf et al. (in preparation) to score responses within retrieval Blocks. These scoring criteria included collective and singular versions being counted as repetitions rather than unique responses (cow and cows), as were baby and adult responses (calf and cow); in contrast, superordinate and subordinate subcategory members counted as unique responses (bear and polar bear). The output from the first and last 20% of usable subjects on both retrieval task blocks were also checked by a second scorer, to confirm my scoring and to check that any animal repetitions between a subject's first and second retrieval blocks were caught and not counted toward the fluency score. Repetitions, non-animals, and other non-unique responses were eliminated from data analysis of the primary DV. In Experiment 1, the score correlations between myself and the secondary rater were high for both Blocks 1 and 2, rs(48) = 0.99 and .97, respectively.

Table 1. Subject Demographics Across Experiments 1 and 2.

Demographics	Experiment 1		Experiment 2	
	N	%	N	%
Gender				
Male	89	55.62	94	44.54
Female	68	42.50	112	52.34
Non binary/gender non-conforming	3	1.88	5	2.36
I don't identify as any of these	0	0.00	0	0.00
Race[a]				
Asian: East Asian	21	10.29	11	5.39
Asian: Indian descent	5	2.45	12	5.68
Black or African descent	8	3.92	21	9.95
Hispanic or Latin descent	6	2.94	4	1.89
North American or Alaskan Native	2	0.98	3	1.42
Middle Eastern, Arab, or North Africa[n]	3	1.47	4	1.89
White or European descent	131	64.21	173	81.99
Education (Prolific)				
High school/ A-levels	42	20.58	51	24.17
Technical/community college	22	10.78	23	10.90
Bachelor’s degree	60	29.41	90	42.69
Graduate degree	31	15.19	42	19.90
Doctoral degree	5	2.45	5	2.36
Age				
18-25	27	13.24	44	20.85

26-30	37	18.13	53	25.12
31-35	51	25.00	54	25.59
36-40	45	22.05	60	28.44

Note. Demographic data were only collected for 160 subjects in Experiment 1 due to a programming error.

[a]Subjects allowed to choose more than one answer.

For exploratory purposes, semantic clustering behaviors were also scored for each fluency block. Participants' propensity to shift between clusters of related exemplars, called switching, is an important component to fluency (Troyer et al., 1997). Switching means that an individual is moving past an exhausted semantic cluster to name items from a new cluster. Clusters are semantic sub-categories, usually with discernable temporal gaps between them (reflecting switches). For example, if a participant named *octopus*, *fish*, *whale*, *lion*, *zebra*, *giraffe*, that would reflect two clusters, the first being sea animals and the second safari animals.

Troyer et al. (1997) established three main Animal categories of "living environment, human use, and zoological categories" and several subcategories therein. The full list of scored subcategories is detailed in Troyer et al.'s (1997) appendix. If participants named an animal that fit into two clusters, it was scored twice, once in each cluster. For example, if a participant named several safari animals, then a lion, and then several large cats, lion was scored both in the safari animal cluster and in the large cat cluster. If a cluster contained smaller clusters within itself, only the overarching cluster was counted. Cluster size reflected the number of animal names recalled in each subcategory, and a switch was counted for each transition between subcategories. So, if a participant named twenty animal names, that was their overall fluency score. If they named those animals into four equally sized subcategories in turn, then they would have four clusters with a cluster size of five each, and they would have three switches between the four clusters. A research assistant scored the first 10% of responses for cluster size and switching and then I scored the remaining 90% of responses.

Preliminary Analyses of Pre-Incubation (Block 1) Fluency Scores

Before addressing my theoretical questions about post-incubation (Block 2) retrieval, I first assessed whether the experimental groups differed in fluency scores during Block 1, before the incubation manipulation. The means and standard deviations are given for Block 1 fluency, clustering, and switching in Table 1. Output did not differ between conditions, $t(202) = 0.67$, $p = 0.501$, though the no-incubation condition did produce larger clusters, $t(202) = 2.86$, $p = 0.004$, $d = 0.39$, but again no difference for switches $t(202) = 1.10$, $p = 0.274$.

Table 2. First Block Fluency

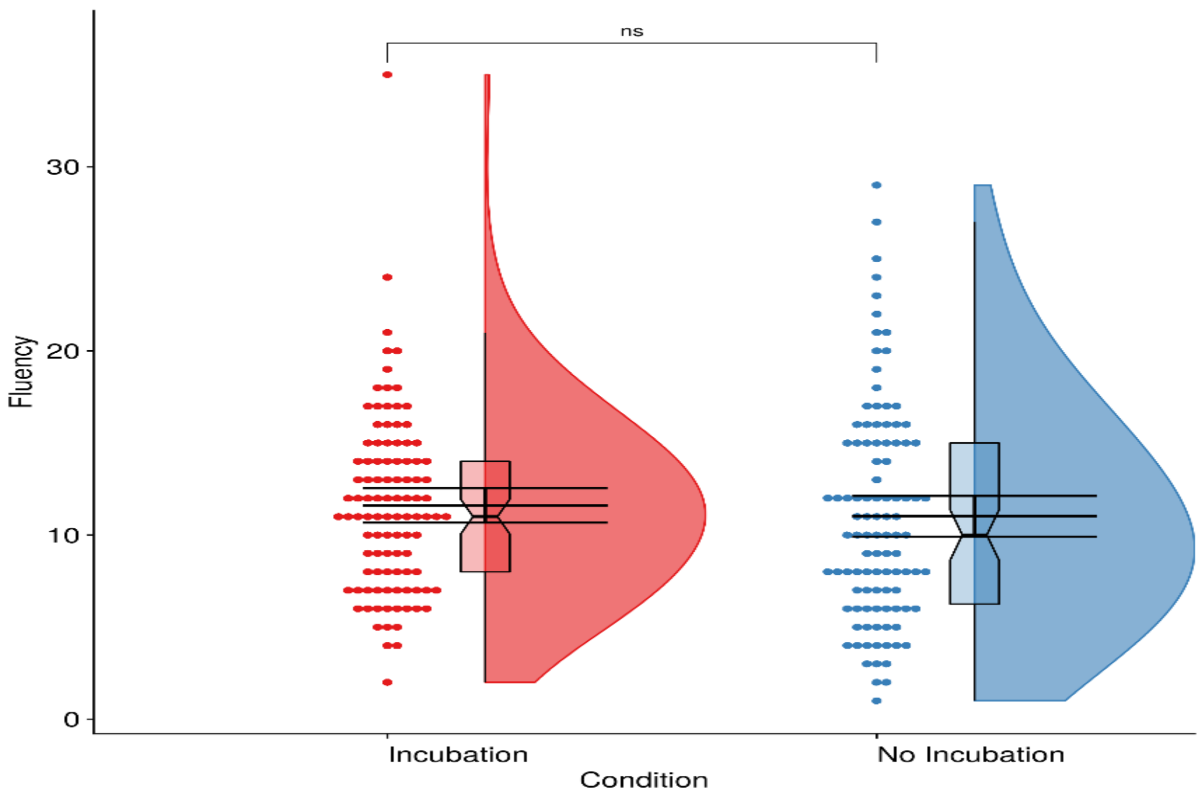
	Inc.			No Inc.		
DV	Mean	SD	Range	Mean	SD	Range
Output	57.0	18.6	19.0-102.0	58.7	17.4	27.0-104.0
Cluster size	2.9	3.1	1.7-8.9	3.9	1.7	1.7-20.5
Switches	28.7	9.0	11.0-51.0	27.4	7.9	11.0-49.0

Primary Analyses

Fluency Scores

Contrary to the prediction that an incubation period would reduce output interference in the second fluency block, Figure 1 illustrates that there was no significant difference in fluency between the incubation condition ($M = 11.6$; $SD = 4.9$, skewness was 1.22 and kurtosis was 4.17) and no incubation condition ($M = 11.0$; $SD = 5.9$, skewness was 0.68, kurtosis was 0.26); $t(203) = -0.76$, $p = 0.446$, $d = 0.11$. To detect a true population effect of this magnitude with 80% power ($\alpha = .05$, 2-tailed) would require 1,299 participants per group. Since the incubation condition possessed a high kurtosis, I re-ran the Student's t-test as a Welch's test, and found similar null-results; $t(201) = -0.39$, $p = 0.697$.

Figure 1. Number of Unique Animal Names Recalled in Second Fluency Task Block by Condition



Note. The lines indicate the mean and 95% CI.

TUT Rates

Although not hypothesized, participants had higher TUT rates in the no incubation condition, (M = 0.290, SD = 0.248), than the incubation condition (M = 0.214, SD =0.224); $t(202) = 2.31$, $p = 0.022$, perhaps because the "incubation" task occurred later in the session for the no incubation condition. Though I predicted that TUT rates and fluency would be positively correlated in the incubation condition, there was no significant correlation between TUT rates and fluency, for either the incubation condition, $r(100) = -.12$ [−.31, .08], $p = 0.184$, or the no incubation condition, $r(100) = -.17$, [−.02, .35], $p = 0.078$; if anything, these correlations were non-significantly *negative*. Scatterplots for these correlations are presented in Figures 2 and 3, respectively.

Figure 2. TUT Rates and Second Block Fluency in the Incubation Condition

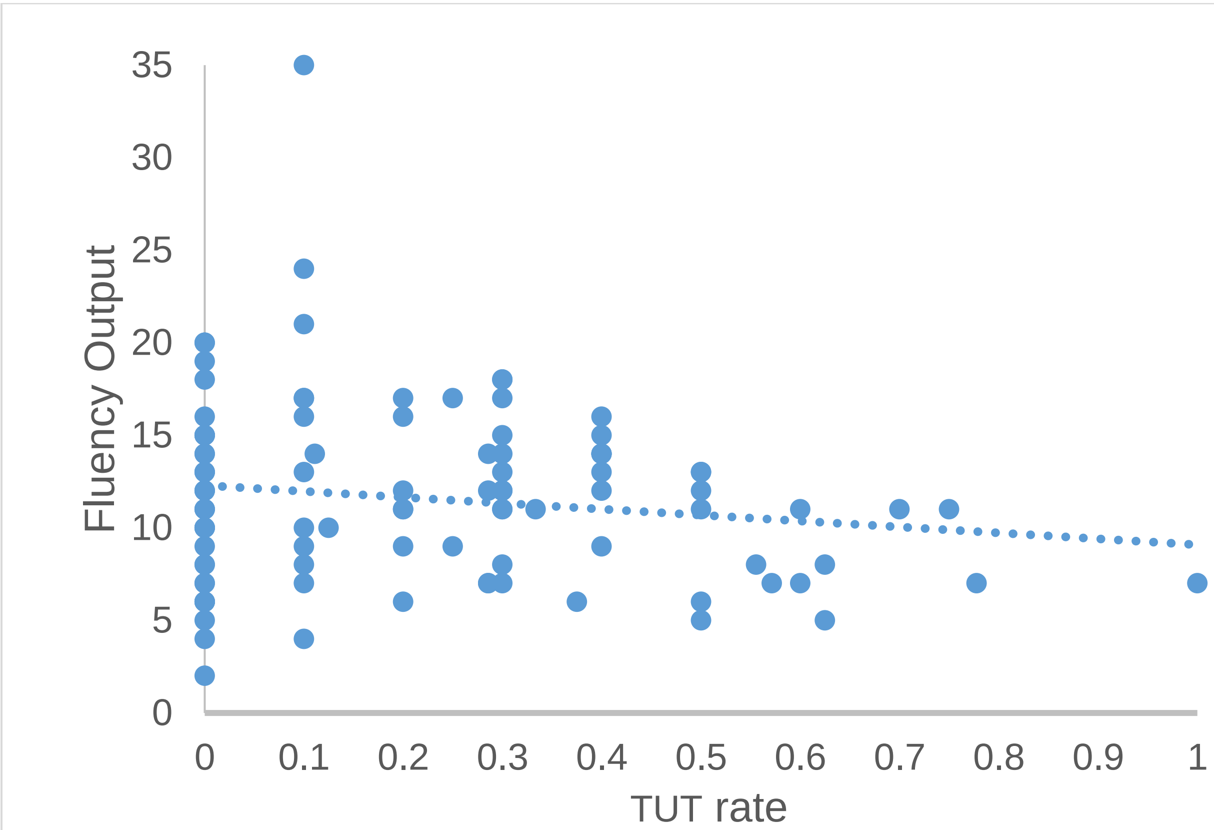
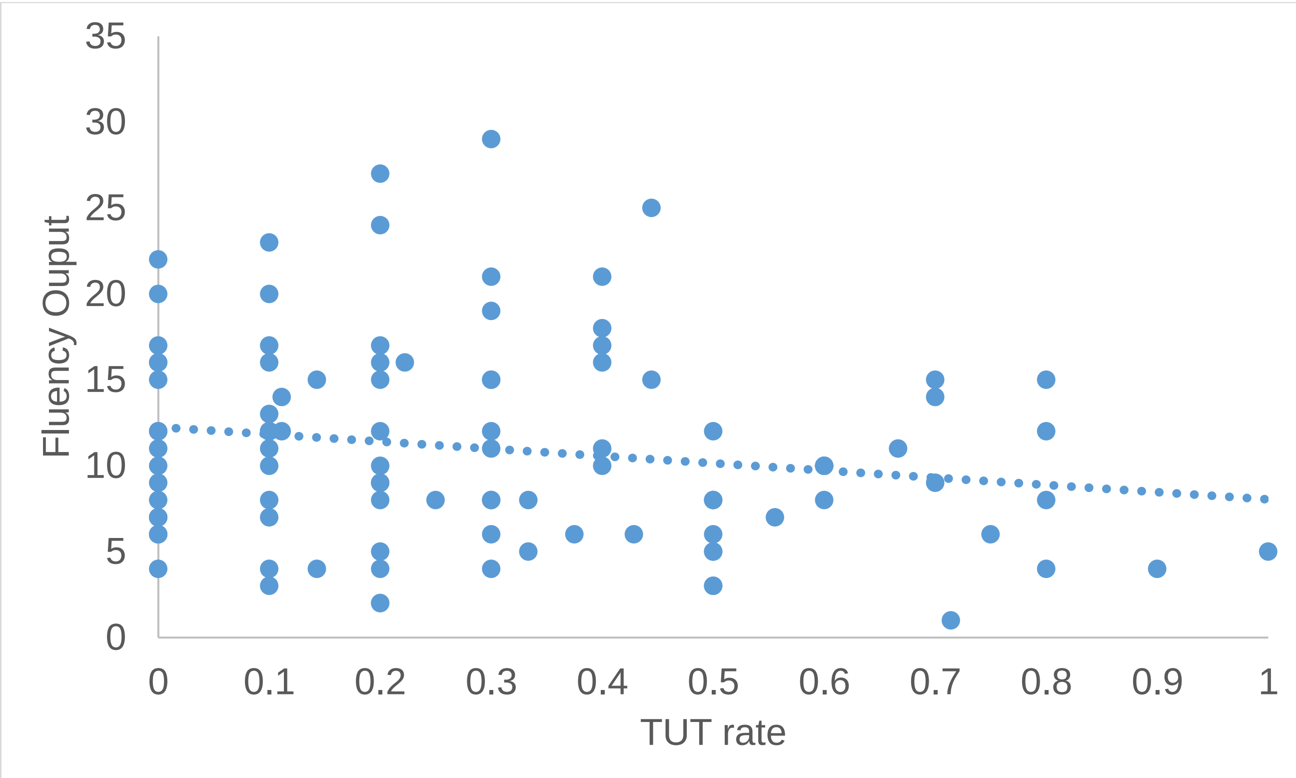


Figure 3. TUT Rates and Second Block Fluency in the No Incubation Condition



Secondary Questions and Exploratory Analyses

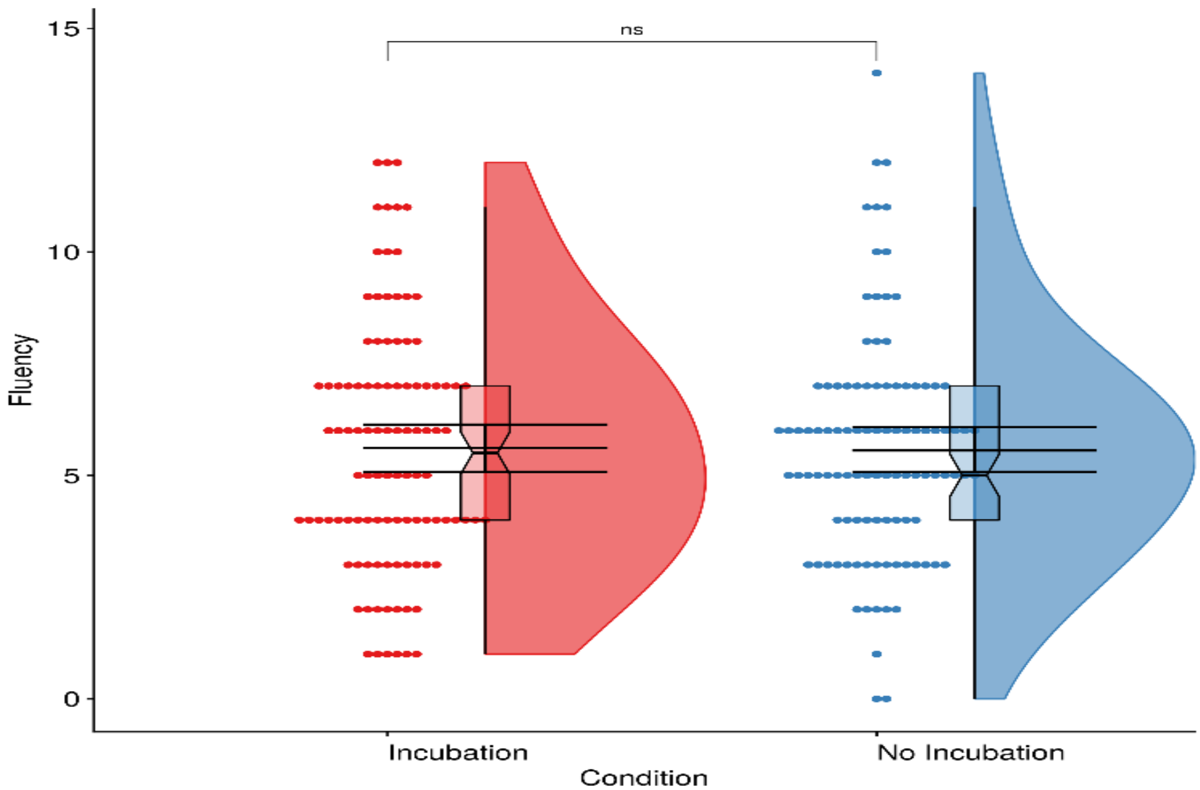
Controlling for Block-1 fluency

A potential confounding variable for second block fluency group comparison could be initial (Block-1) fluency ability. To examine this possibility, I conducted an ANCOVA to control for Block-1 fluency scores on recall from the second fluency block. There is no significant difference between the two conditions (incubation or no incubation) on Block-2 fluency when controlling for Block-1 fluency: $F(2, 200) = 0.94$, $MSE = 16.24$, $p = 0.391$,

First 60 s of Block-2 fluency

Given that there was no incubation effect on second block recall, overall, I considered whether any subtle or short-lived effect had been missed when analyzing the entire 2 min Block-2 task. Here, then, I analyzed only the first 60 s of output of the second fluency block. Again, as illustrated in Figure 4, there was no significant difference in fluency between the incubation condition (*M* = 5.6, *SD* = 2.5) and the no incubation condition (*M* = 5.6 *SD* = 2.8); $t(203) = 0.13$, $p = 0.895$.

Figure 4. First 60 s of Block-2 Fluency by Condition



Note. The lines indicate the mean and 95% CI.

Informally Assessing Block-1 Fluency Impasse

My research question hinges on the assumption that participants incurred an impasse towards the end of the first fluency period. Could one of the reasons there were null results be a lack of impasse? I examined just the last 60 s of the first fluency block (see Table 3). Given how many items people still output in the last minute in Experiment 1 compared to previous literature, which suggests around three animals per min at the 7-min mark (Bousfield & Sedgewick, 1944;

Fitzgerarld, 1983; Gronlund & Shiffrin, 1986; Kail & Nippold, 1984), there might not have been a significant impasse for many participants at the 7-min mark.

Table 3. Mean Number of Animals Recalled During Last 60 s of First Fluency Block

Condition	Mean	SD	Range
Incubation	7.54	4.76	0-23
No Incubation	8.23	4.85	0-21

Reading Time Estimates

To further examine possible mental context change, I assessed the correlation between TUT rate and time estimates. There was no significant correlation for TUT rates and reading task time estimate for the incubation condition, $r(100) = -0.08$ [−0.27, 0.12], $p = .429$, but the no-incubation condition was significant, though weak, $r(100) = 0.25$ [.06, 0.42], $p = .012$. (For archival purposes, I note that there was no significant difference in time estimates between the incubation condition ($M = 9.04$, $SD = 2.86$) and no-incubation condition ($M = 8.94$, $SD = 3.05$), $t(209) = -1.14$, $p = 0.814$.

Daydreaming

Since TUT rate did not correlate to fluency in either condition, one additional consideration is if *type* of mind-wandering was related to fluency. Given that Delaney et al., 2010 did find significant context change effects during daydreaming, I sought a correlation between proportion of daydreaming responses to the thought probes and Block-2 fluency. However, there was no significant correlation for daydreaming TUT rates and Block-2 fluency for the incubation condition, $r(100) = -0.00$, [-0.2, 0.19], $p = 0.984$, nor the no-incubation condition, $r(100) = -0.00$, [-0.2, 0.19], $p = 0.992$.

Clustering and Switching

Since a primary prediction was that the incubation group would produce more names than the no-incubation group, I also predicted that those in the incubation group would produce larger cluster sizes, and more switching, as a means by which to produce an overall larger fluency output in the second period fluency task. However, for cluster size in the second fluency block, subjects in the incubation condition (M = 2.87, SD = 1.69) produced significantly smaller clusters than did those in the no incubation condition (M = 3.86, SD = 3.14), $t(199) = 2.75$, p <.001, d = 0.39. For number of switches in the second fluency block, there was no significant difference between the incubation condition (M = 28.69, SD = 9.04) and the no-incubation condition (M = 27.36, SD = 7.92); $t(199) = -1.11$, $p = 0.269$. Please refer to Table 1 for Block 1 clustering and switching scores and analyses.

Strategies

As further exploration into what might drive fluency output, I also analyzed strategy usage. Participants could report any number of the 16 retrieval strategies provided. I chose to analyze 11 of the 16, as they all had at least 40 participants selecting it or 40 participants not selecting it (see Table 4). That is, I did not analyze a strategy if more than 80% of people all chose to use or not use it, so that there was some potential for variability in fluency scores between strategy endorsers and non-endorsers.

Table 4. Strategies Analyzed and the Number of Endorsers/Non-Endorsers for Each

Strategy Number	Question	First Block Number of Endorsers/Non-Endorsers	Second Block Number of Endorsers/Non-Endorsers
1	Types or classifications of animals (for example, thinking about all the types of mammals)	159/45	145/59

5	Physical attributes of animals (for example, thinking of all the animals with wings or all the animals with four legs)	81/123	45/159
6	Types of environments animals live in (for example, thinking of the desert animals or all the jungle animals)	166/38	155/49
7	Geographic location of animals (for example, thinking of all animals that live in Africa and then all the animals that live in the jungle)	130/74	111/93
8	Frequency of animals (for example thinking of the most common animals or the rarest animals)	66/138	45/159
9	Location of animals in a zoo (for example, thinking of all the animals in Sea World or in the monkey house at a zoo)	68/136	53/151
10	Animals affected by humans (for example, thinking of all the animals that are endangered)	132/72	121/83
11	Animals that you do not like, or fear (for example, thinking of a snake you fear, then thinking of a rat because you also fear rats)	118/86	93/111
12	Animals that live with humans (for example, thinking of all domesticated animals or animals on a farm)	150/54	146/58
13	Animals personally relevant to you (for	138/66	136/68

15	example, animals that you have seen recently) Association with previous animals (for example, “cat reminds me of dog which reminds me of wolf”)	101/103	104/100

Out of the 22 comparisons between subjects who endorsed versus didn’t endorse a strategy (each of the 11 strategies were analyzed for both fluency Blocks), there were only five cases in which either endorsing or not endorsing a strategy was associated with higher fluency (see Table 5). Strategies 1 and 7 (classifications and geographical locations) were associated with increased fluency during Block 1. For Block 2, those who used Strategy 6 (habitat types) produced more animal names than people who did not use it, while those who used Strategies 8 and 11 (frequency of contact and disliked/feared) reported fewer animal names than those who did not use them.

Given the lack of consistency between first and second block fluency results here, there is no clear indication of which strategies may lead to greater fluency output.

Table 5. Fluency Scores by Strategy Endorsement

Strategy	Endorsed		Non-Endorsed		*t*(df)	*p*
	M	*SD*	*M*	*SD*		
1	59.71	17.5	52.71	7.27	2.18(72)	.032
6*	11.90	5.71	11.02	5.57	2.19(93)	.031
7	62.25	17.11	52.97	18.37	3.57(178)	.000
8*	10.26	5.47	12.03	5.4	2.11(119)	.037
11*	10.83	5.03	12.44	5.6	2.01(165)	.046

Note. The * indicates results are shown for the second fluency Block.

Discussion

There were two main questions for Experiment 1—did an incubation period increase fluency post-incubation, and was TUT rate during incubation positively associated with fluency after incubation? The results demonstrated that: (1) the incubation condition did not produce greater fluency-task recall after incubation on the second recall block than did the no incubation condition (this also held true when controlling for first block fluency), and (2) incubation-period TUT rates were not positively correlated with fluency in the second recall block. Therefore, no incubation benefits were shown, contrary to predictions. Secondarily, switching also did not differ between groups, though cluster size was larger for the no-incubation group. Since the overall fluency did not differ between groups, however, the meaning of this cluster size difference is not clear. Furthermore, strategy usage did not provide consistent answers regarding how strategy choice may impact fluency output.

There is an obvious concern from these data. The output in the final min of the first recall period was higher than expected. Participants were naming, on average, 7.9 animals in the last min of Block 1, despite the average in the literature being closer to 3 animals per min at the 7-min mark (Bousfield & Sedgewick, 1944; Fitzgerald, 1983; Gronlund & Shiffrin, 1986; Kail & Nippold, 1984). Perhaps no incubation effects could have taken place because there was no impasse from which participants needed an incubation.

Experiment 1 showed null effects for most analyses, except for two significant findings that, contrary to predictions, the no-incubation condition produced larger cluster sizes in recall Block 2, and only this group's TUT rates correlated significantly with their reading time estimates. Therefore Experiment 2 replicated Experiment 1, with slightly modified tasks, to determine whether no incubation effects occur in fluency, or if they were simply missed in Experiment 1.

Experiment 2

To address the concerns from Experiment 1, I changed the timing of two tasks in Experiment 2. I increased the Block-1 fluency recall period from 7 to 9 min, to increase the chance that participants would hit impasse before incubation. Also, I increased the presentation duration of each sentence in the reading task from 7 s per slide to 8 s per slide, thereby increasing the incubation period from 11 min to 12.5 min, and thus potentially giving more opportunity for mental context changes that may allow an incubation effect to occur. All other aspects of the methods (including scoring and inclusion/exclusion criteria) were identical to Experiment 1.

Once again, for my primary analyses, I tested whether: (1) the incubation group output more animals in recall Block 2 compared to the no-incubation group, and (2) TUT rates positively predicted post-incubation retrieval output in the incubation group.

Participants

I recruited 249 participants for pay ($11.26/hour) from Prolific Academic. Data from 38 participants were dropped for meeting the exclusion criteria (which were identical to Experiment 1). All 38 were dropped for getting 2 or more questions wrong on the reading comprehension check, and like in Experiment 1, that exclusion criterion proved to also eliminate all participants who met any other exclusion criteria. That left data from 211 participants (incubation condition = 109, no-incubation condition = 102) in the dataset. See Table 1 for the demographics of retained participants.

Tasks

Retrieval

The retrieval task was nearly identical to that used in Experiment 1. The only difference was that in the first block of retrieval, task duration increased from 7 min to 9 min. The fluency scores for myself and the second rater were again highly correlated, with both retrieval blocks' *r*s(39) = 0.99.

Strategy Assessment

The strategy assessment was the same as in Experiment 1.

Incubation Passage Reading

The incubation period reading task was nearly identical to that used in Experiment 1. The only difference was that each slide (sentence) was presented for 1 s longer, from 7 s to 8 s per slide, for an overall 90 s task time increase, from 11 min to 12.5 min.

Thought Probes

The probes were the same in appearance, number, and timing as in Experiment 1.

Procedure

The procedure for both groups was the same as that in Experiment 1.

Results

Preliminary Analyses of Pre-Incubation (Block 1) Fluency Scores

Before addressing my theoretical questions about post-incubation retrieval, I first assessed whether the experimental groups differed in fluency scores during Block 1, before the incubation manipulation (see Table 6). The conditions didn't differ in recall output, $t(209) = 0.69$, $p = 0.494$, cluster size, $t(209) = 1.01$, $p = 0.315$, or switches, $t(209) = 0.983$, $p = 0.325$.

Table 6. First Block Fluency

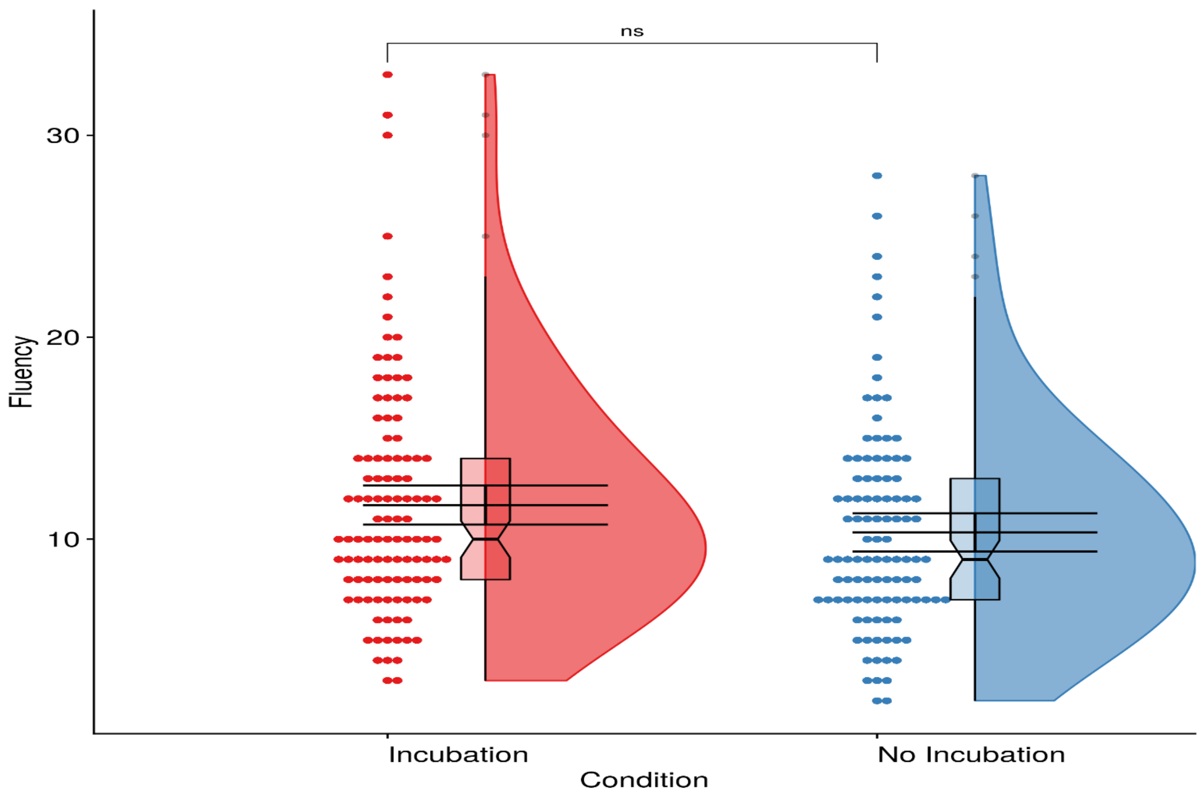
	Inc.			No Inc.		
DV	Mean	SD	Range	Mean	SD	Range
Output	67.39	0.05	19.0-117.0	65.45	21.50	24.0-127.0
Cluster size	3.20	0.90	1.6-8.0	3.80	1.24	1.1-12.8
Switches	32.20	9.90	3.0-57.0	30.90	9.20	7.0-57.0

Primary Analyses

Fluency Scores

Like in Experiment 1, my predictions of group differences on fluency did not occur in the Block-2 retrieval task; there was no significant difference between the incubation group (M = 11.69, SD = 5.73, skewness was 1.30 and kurtosis was 2.36) and no-incubation group (M = 10.34, SD = 5.12, skewness was 1.06, and kurtosis was 1.52); $t(209) = -1.79$, $p = 0.074$, d = 0.25. Although the mean difference was in the predicted direction, to detect a true population effect of this magnitude with 80% power ($\alpha = .05$, 2-tailed) would require 253 participants per group. For completeness, given that I re-ran the Student's t-test in Experiment 1 as Welch's test, I did so again here, and again found similar null results; $t(209) = -1.80$, $p = 0.073$.

Figure 5. Number of Unique Animal Names Recalled in Second Fluency Task Block by Condition



Note. The lines indicate the mean and 95% CI.

TUT Rates

Unlike in Experiment 1, participants reported similar TUT rates in the no incubation condition, (M = 0.23, SD = 0.22), and the incubation condition (M = 0.21, SD = 0.18) t(209) = 0.84, p = 0.400. Though I predicted that TUT rates and fluency would be positively correlated in the incubation condition, there was no correlation between TUT rates during the reading period and second block fluency, for either the incubation group, r(108) = .10 [−0.09, 0.28], p = 0.226, or the no-incubation group r(101) = −.02 [−0.21, 0.18], p = 0.389. To detect a true population correlation of ρ = .10 with 80% power (α = .05, 2-tailed) would require 782 participants in the incubation group.

Figure 6. TUT Rates and Second Block Fluency in the Incubation Condition

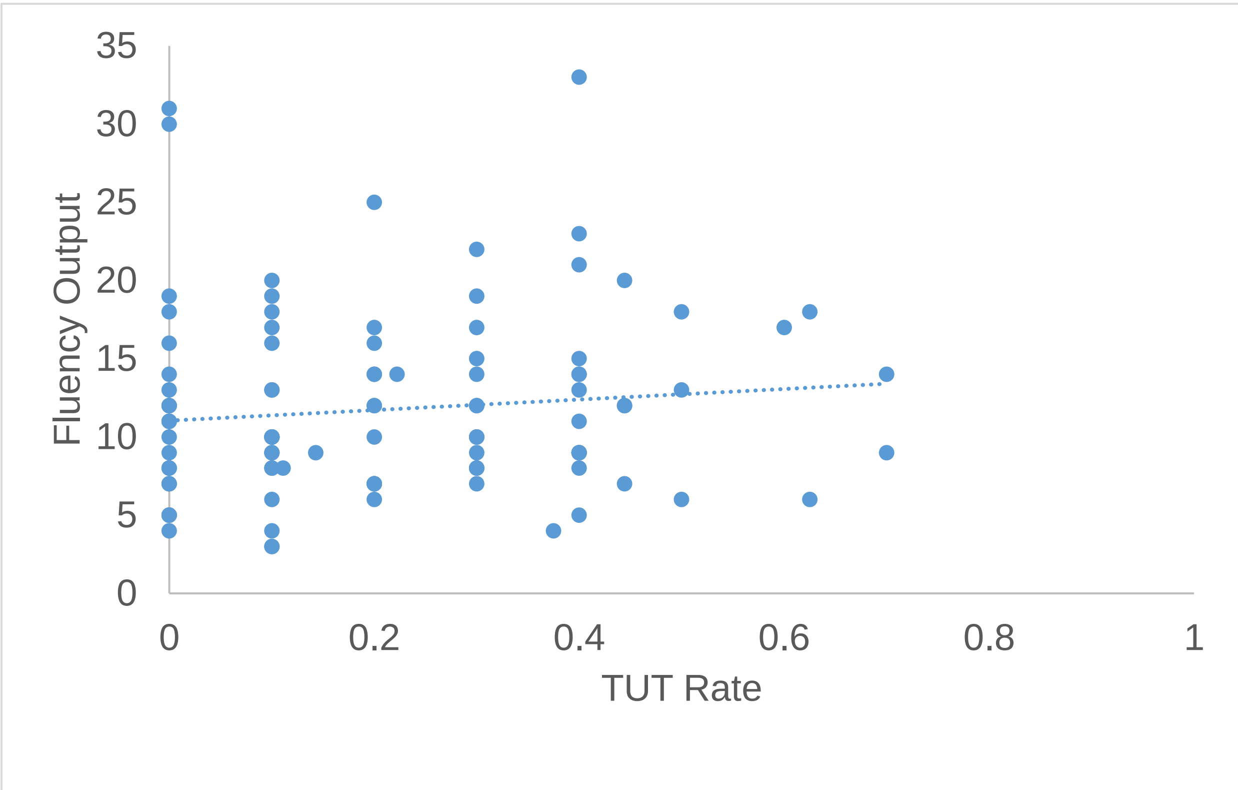
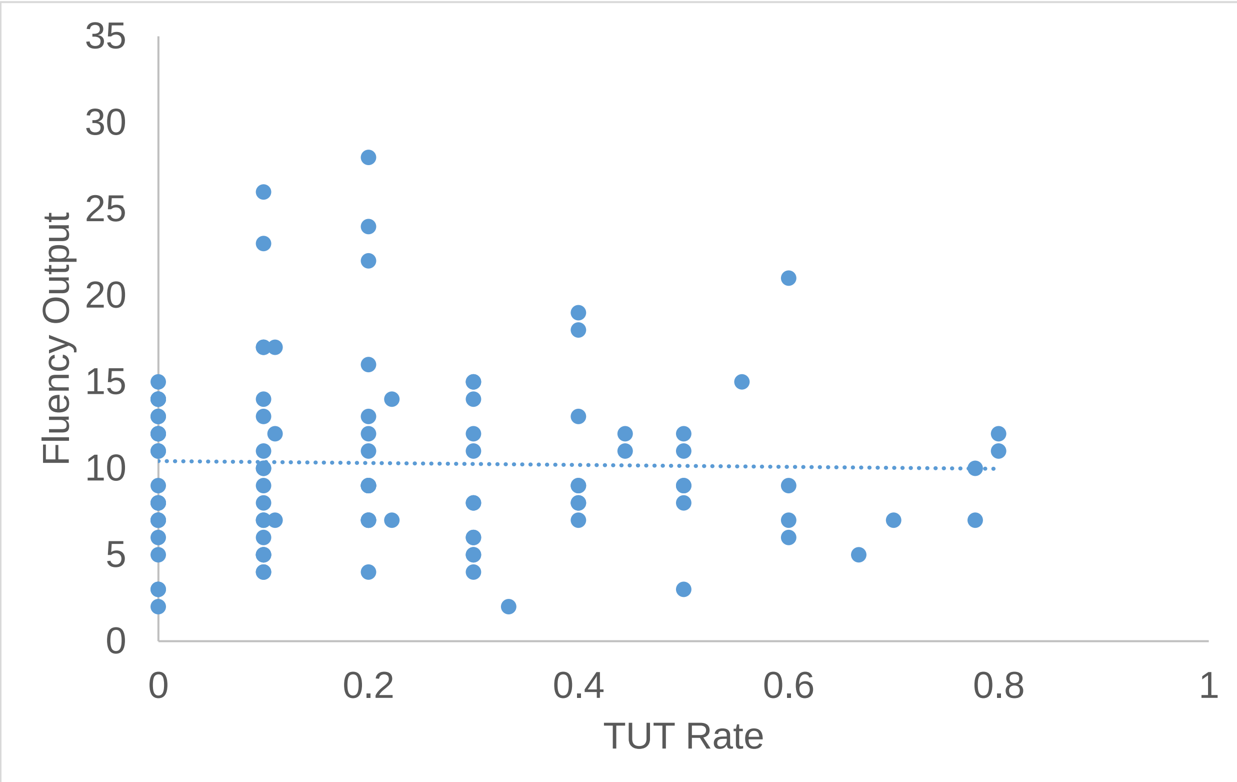


Figure 7. TUT Rates and Second Block Fluency in the No Incubation Condition



Secondary Questions and Exploratory Analyses

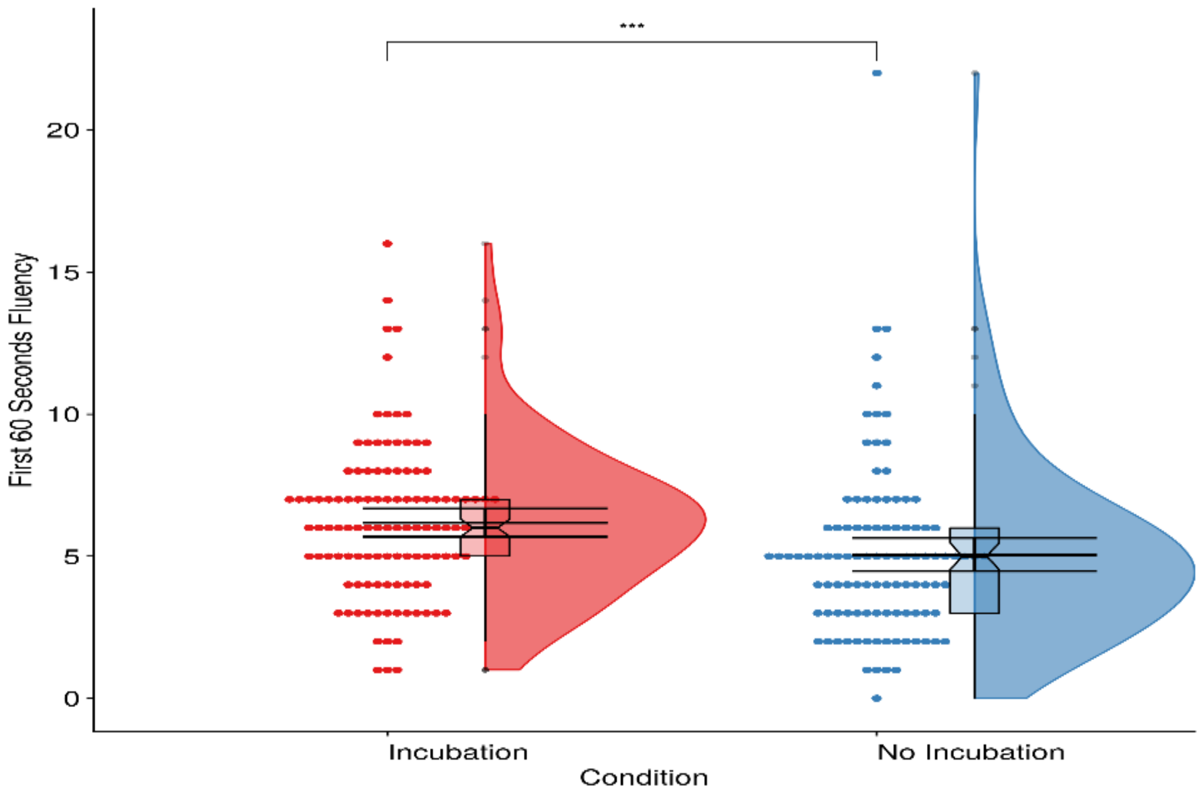
Controlling for Block-1 Fluency.

Just as in Experiment 1, an ANOVA on Block-2 fluency that controlled for Block-1 fluency indicated no significant difference between the incubation and no incubation conditions: $F(2, 208) = 2.78$, $MSE = 23.88$, $p = 0.097$.

First 60 s of Block-2 Fluency.

Since there was not an incubation effect again for Block-2 recall, overall, I again considered whether the first 60 s of the Block may have shown a short-lived effect. Unlike in Experiment 1, here in Experiment 2 there was a significant difference between incubation (M = 6.19, SD = 2.67) and no incubation conditions (M = 5.05, SD = 3.10) in the first 60 s of the second recall task; $t(209) = -2.86$ $p < .005$, $d = 0.39$ (even with an outlying high fluency score in the no incubation condition). TUT rates, however, still did not correlate significantly with fluency for the incubation condition, $r(108) = 0.05$ [-0.14, 0.21], $p = 0.606$.

Figure 8. First 60 s of Block-2 Fluency by Condition



Note. The lines indicate the mean and 95% CI.

Informally Assessing Block-1 Fluency Impasse.

Experiment 2 added 2 min to end of Block 1 compared to Experiment 1. Did this create an impasse, as intended? In Experiment 1, participants produced a mean of about 8 words in the last 60 s, with a maximum of 23. Table 7 presents the means, standard deviations, and ranges for responses in the last 60 s of Block 1 for Experiment 2. Considering that the output is so low, and closer to the literature norm, it seems safe to infer that impasse was reached by most participants.

Table 7. Mean Number of Animals Recalled During Last 60 s of First Fluency Block

Condition	Mean	SD	Range
Incubation	3.82	0.29	0.0-11.0
No Incubation	3.75	0.28	0.0-11.0

Reading Time Estimates

There was no significant correlation for TUT rates and reading task time estimate for the incubation condition, $r(109) = 0.08$ [−0.11, 0.26], $p = .408$, or for the no-incubation condition, $r(100) = 0.04$ [−0.16, 0.23], $p = .693$. (Again, for archival purposes, I note that there was no significant difference between the incubation condition ($M = 8.84$, $SD = 2.86$) and no-incubation condition ($M = 8.38$, $SD = 3.03$), in participants' perception of how many minutes the reading task lasted, $t(209) = -1.14$, $p = 0.255$.)

Daydreaming

Again, to determine if type of mind-wandering was related to fluency, I sought a correlation between proportion of daydreaming responses to the thought probes and Block-2 fluency. However, there was no significant correlation for daydreaming TUT rates and Block-2 fluency for the incubation condition, $r(109) = 0.02$, [-0.17, 0.21], $p = 0.836$, nor the no-incubation condition, $r(100) = 0.05$, [-0.15, 0.24], $p = 0.620$.

Clustering and Switching

Unlike in Experiment 1, for cluster size in the second fluency block, there was not a significant difference between the incubation ($M = 3.23$, $SD = 1.23$) and no-incubation ($M = 3.78$, $SD = 1.63$) conditions, $t(195) = 0.91$, $p = 0.362$. For number of switches in the second fluency Block, there was also not a significant difference between the incubation ($M = 30.92$, $SD = 9.62$) and no-incubation ($M = 32.16$, $SD = 9.14$) conditions, $t(195) = 0.91$, $p = 0.346$.

Strategies

The same method of choosing strategies to analyze in Experiment 1 was applied here in Experiment 2. The 11 analyzed strategies here (see Table 8) differed from those in Experiment 1, in that strategies 3 and 14 were added to analysis, and strategies 6 and 9 were no longer analyzed. These were again analyzed for both first and second recall periods.

Table 8. Strategies Analyzed and the Number of Endorsers/Non-Endorsers for Each

Strategy Number	Question	First Block Endorsers/Non-Endorsers	Second Block Endorsers/Non-Endorsers
1	Types or classifications of animals (for example, thinking about all the types of mammals)	152/57	162/47
3	Going through the alphabet (for example, thinking of all the animals that start with "A", and then that start with "B"	79/130	68/139
5	Physical attributes of animals (for example, thinking of all the animals with wings or all the animals with four legs)	79/130	65/142
7	Geographic location of animals (for	139/70	109/100

	example, thinking of all animals that live in Africa and then all the animals that live in the jungle)		
8	Frequency of animals (for example thinking of the most common animals or the rarest animals)	61/148	41/168
10	Animals affected by humans (for example, thinking of all the animals that are endangered)	134/75	113/96
11	Animals that you do not like, or fear (for example, thinking of a snake you fear, then thinking of a rat because you also fear rats)	103/106	88/121
12	Animals that live with humans (for example, thinking of all domesticated animals or animals on a farm)	144/65	79/130

13	Animals personally relevant to you (for example, animals that you have seen recently)	144/45	145/44
14	Using free association or random generation (such as saying whatever pops into your head)	45/164	38/169
15	Association with previous animals (for example, "cat reminds me of dog which reminds me of wolf")	107/102	114/95

Out of the 22 comparisons between subjects who did versus didn't endorse a strategy (each strategy was analyzed for both periods), there were only three cases in which either choosing or not choosing a strategy was associated with higher fluency, all from the Block-1 task (see Table 9). For strategies 7, 12, and 13, (geographic location, domesticated, personal relevance) those who choose to use them also had higher fluency. Given that in Experiment 2 there were no strategy effects on second block fluency and, given the inconsistency between Experiment 1 and Experiment 2 in overall strategy-usage patterns and effects (only strategy 7 [geographic location] was associated with higher Block-1 fluency in both experiments), it is not clear how (or whether) strategy usage may impact fluency.

Table 9. Fluency Scores by Strategy Endorsements

Strategy	Endorsed		Non-Endorsed		t(df)	p
	M	SD	M	SD		
7	70.23	19.57	62.41	21.37	2.75(204)	.006
12	69.72	20.01	60.27	20.96	3.04(151)	.002
13	69.29	20.91	60.19	19.23	2.76(141)	.003

CHAPTER III: DISCUSSION

General Discussion

Previous research has demonstrated incubation benefits for problem-solving fixation (Dodds et al., 2004; Sio & Ormerod, 2009) and output interference in episodic memory (Smith & Vela, 1991), with incubation especially helping when there is a mental context change during incubation (Smith & Beda, 2020). Therefore, I predicted that an incubation period would increase semantic memory fluency by relieving output interference, particularly when a mental context change was likely.

Broadly, this study was motivated by a lingering question in the incubation literature concerning mind wandering's possible role in stimulating creativity. Whereas Baird et al. (2012) established a connection between divergent thinking performance and mind-wandering during an incubation period, several studies since have not (Murray et al., 2021; Smeekens & Kane 2016; Steindorf et al., 2020). To address this controversy, the current study sought to break down the complex divergent thinking task into one of its more straightforward components—retrieval from semantic memory. Divergent thinking requires individuals to create uses for familiar objects, but before they can move on to generate novel uses, participants must first recall the object and its everyday uses from memory. Determining whether mind wandering may be associated with retrieval fluency after impasse should thus shed light on the more dynamic and complex relationship between divergent thinking and mind wandering. That is, if evidence suggested that TUTs could facilitate overcoming output interference by creating mental context changes that made previously recalled items less accessible, then TUTs might facilitate creative incubation in the same way. Rather than reflecting subconscious processing or "looser" spreading activation across semantic networks, perhaps mind-wandering boosts creativity simply by helping us to stop thinking about previously exhausted dead ends.

The current study therefore examined whether an incubation period could facilitate overcoming an impasse in memory retrieval, especially (or only) when the incubation task promoted mind-wandering experiences (i.e., TUTs) that may create mental context changes. In two experiments, participants recalled animal names (for 7 min in Experiment 1 and 9 min in Experiment 2) followed by either continued recall for another 2 min, or an incubation block of a reading task with thought probes, before continuing to recall for another 2 min.

The Present Findings

Across experiments, and contrary to predictions, fluency in the second recall block did not differ significantly by condition, indicating that incubation did not increase fluency performance compared to continuous recall effort; these non-significant results also held when controlling for first-block fluency. When analyzing only the first 60 s of Block-2 fluency, however, Experiment 2 did show that the incubation condition produced significantly more animal names than the no-incubation condition. So, perhaps with enough time to generate an impasse (i.e., 9 min in Experiment 2), there can be a short-lived incubation effect on fluency. This intriguing exploratory result will require replication.

Whereas the overall output of Block-2 recall did not differ between incubation and non-incubation conditions, in Experiment 1, the no-incubation condition did produce larger cluster sizes than the incubation condition. This cluster size difference did not replicate in Experiment 2. There were also no switching differences between conditions in either experiment. Thus, overall, the way participants were semantically outputting their recall did not seem to differ between groups.

Furthermore, TUT rates in the incubation reading task did not correlate positively with fluency output in Block 2 in either experiment, as would be expected if TUTs created mental context changes that reduced output interference from Block 1. In Experiment 1, these correlations, though nonsignificant, trended towards a negative relationship, but in Experiment 2, they were positive, though still nonsignificant. Considering that across experiments, there were no significant correlations, and the directions of the correlations were opposed between experiments, this study finds no support for mind wandering during incubation promoting post-incubation fluency performance. Additionally, I measured the participant's self-reported time estimate for the reading task, as an indirect measure of mental context change experienced during reading. I predicted, therefore, that TUT rate would positively correlate with the reported time estimation. However, there was a weak significant correlation between TUT rates and reading estimates only for the no-incubation group, only in Experiment 1 (this unpredicted finding was not replicated in Experiment 2).

Null Effects and Potential Task Differences

Except for the exploratory Experiment 2 finding of an incubation-condition benefit for fluency in the first 60 s post-incubation, the results from this study failed to demonstrate an overall difference between continuously attempting retrieval and experiencing an incubation block in between retrieval attempts. This is a surprise, given my predictions based on the relevant literatures, but perhaps the tasks used and their timing can explain the null findings. Both the fluency and incubation tasks might not have served the purposes that I intended.

Primary Tasks

The incubation-sensitive tasks cited in the introduction largely were not direct matches to the tasks used in the current study, and those differences might have impacted the study results in unanticipated ways.

Smith and Beda (2020) demonstrated that a mental context change during incubation relieved fixation on the RAT and was related to increased solution rates. But whereas the RAT is a verbal task that requires participants to conduct a memory search to find potential solution words, it is fundamentally a problem-solving task that requires more than memory retrieval for successful performance.

The task used by Smith and Vela (1991) was more like that in the current study. They found that an incubation period relieved output interference on recall, but their recall paradigm required subjects to memorize a list of pictured objects and, after differing amounts of time, write down in any order what objects they could remember. Thus, their task tapped into episodic memory rather than the semantic memory retrieval required in fluency tasks. Although I am unaware of any research suggesting that output interference operates differently in episodic than in semantic memory tasks, perhaps incubation-releasing interference is more task-dependent than I had thought.

For instance, consider Smith et al. (2017), which suggested that at least some forms of incubation may not benefit retrieval from at least some kinds of categories. Their paradigm assessed semantic memory, in which participants named items from given categories in one of two ways, either switching between the two categories or completing one category fully before moving on to the other category. This was not a typical incubation design, since the “incubation” was taking a break from one part of the retrieval task to work on another part of the retrieval task. It still demonstrated incubation benefits, however: those who switched back and forth produced more items overall than the continuous recall group.

Of key interest, though, the Smith et al. (2017) study also raised the problem of "flexibly defined categories." They gave participants category pairs in three experiments, with a different category pair in each experiment. Some categories were taxonomic (birds and clothing), and some of them were flexibility-defined (cold things, heavy things, equipment you take camping, and fattening foods). Whereas "flexible" categories were more ad hoc and could draw from multiple taxonomic categories, bird and clothing were defined as not flexible because participants could only answer from one taxonomic category. Of importance, here, Smith et al. (2017) found that only category switching between flexible categories yielded incubation benefits. They argued that flexible categories possess less category structure and therefore are represented moment-to-moment by changing contexts. These changing contexts are then more prone to forgetting interfering items compared to less flexible categories because participants are likely to restructure their memory search for a flexible category. In sum, only flexible categories lend themselves to mental context change and thus incubation-based reductions in output interference.

The current study used animals as the category. While Smith et al. (2017) found that "birds" was not a flexible category that lent itself to incubation effects or context change, the "animals" category does allow for ad hoc generation of many subcategories, like birds and farm animals, safari animals, mammals, etc. At the same time, perhaps the general category of animals is still not flexible or contextually divergent enough to demonstrate incubation effects. The present study did not consider whether the particular fluency task chosen was likely to induce context change, only whether it was likely to induce impasse, and then the incubation task was meant to facilitate mind wandering and context change.

There may be another, less theoretically interesting reason for the null incubation effects in the present study. Perhaps participants decreased their retrieval responses as time passed during fluency the task for reasons other than output interference. That is, perhaps they simply grew less motivated or engaged by the task over time. Smith et al. (2017), for example, had participants name items for only 6 min total, whereas my participants named items for either 9 or 11 min total. Instead of output interference blocking new items, then, perhaps participants in my experiments were simply too tired and unmotivated to continue. And if there were no interference from which to be freed, then incubation would likely not have helped; instead, it only gave participants yet another task to do.

As a counterargument against the idea that motivation and interest drove decreasing recall in Block 1, however, I increased the Block-1 duration between experiments because there was no evidence of impasse at 7 min, much less at 6 min in Experiment 1. That is, it did not appear that participants were generally struggling to recall new items at the 7 min mark. Moreover, it was in Experiment 2, with the longer first retrieval block, in which the data showed an effect of incubation on the first 60 s of Block-2 retrieval. This exploratory finding suggests the possibility that incubation did briefly reduced output interference.

Incubation Task Durations

The duration of the incubation period may or may not have played a role in the null results reported here. Dodds et al. (2004) found that most studies showed an incubation effect in their narrative review of problem solving and creativity, but the current study's incubation periods were shorter than those that generated the largest effects identified by Dodds et al.

(2004). The present study imposed 11- and 12.5-min incubation periods, but Dodds et al. (2004) found that 15–30-min incubation periods produced the largest incubation effects.

Sio and Ormerod's (2009) quantitative meta-analysis, in contrast, explained the influence of task time differently than did Dodds et al. (2004). Sio and Ormerod (2009) found that, whereas incubation period duration did *not* predict incubation effects, the length of the problem-solving period *before* incubation (the "preparation period") did: Longer preparation periods were associated with larger incubation effects. The preparation period of the present study was long by the standards of the meta-analysis (7 min in Experiment 1 and 9 min in Experiment 2), so I expected an incubation effect following that preparation duration. And in fact, when the pre-incubation recall task increased from Experiment 1 to Experiment 2, there was an incubation effect for the first 60 s of Block 2 of retrieval, suggesting that long preparation periods might allow for short-lived incubation benefits to memory retrieval.

Exploratory Findings of a Short-Lived Incubation Effect in Experiment 2

As noted above, an exploratory analysis found that there was an incubation effect for the first 60 s of post-incubation retrieval fluency in Experiment 2, suggesting that incubation effects may be short-lived. Smith and Vela (1991) similarly found that their incubation effects were strongest in the first minute of their retesting in both of their experiments. Unfortunately, Sio and Ormerod (2009) did not analyze the effects of post-incubation task durations in their meta-analysis, so this might prove to be an interesting avenue of questioning for researchers interested in incubation effects.

Even if we take the present exploratory finding of a fleeting incubation benefit at face value, however, post-incubation fluency did not correlate positively with incubation-period TUT rates, meaning that mind wandering during incubation was not associated with increased post-incubation fluency. This finding is reminiscent of Murray et al. (2021), Smeekens and Kane (2016), and Steindorf et al. (2020), all of whom were unable to establish a correlation between incubation-period TUT rates and post-incubation creative problem-solving, as measured by divergent thinking tasks. Since divergent thinking tasks require memory retrieval first before being able to employ creative uses for objects, and require controlling interference from these retrieved uses (see Gilhooly et al., 2007), the present study asked whether mind wandering could facilitate the mental search for memory. It did not appear to do so.

Delaney et al. (2010) found that daydreaming produced mental context changes, but daydreaming was defined differently than the current study's definition of mind wandering. Delaney et al (2010) asked participants to memorize a word list and then to think of previous or current homes (in Experiment 1) or near or far vacations (in Experiment 2) before memorizing a second list. Participants in the far-thought conditions subsequently forgot more learned List 1 items than did those in the near conditions. So, whereas Delaney et al. (2010) found experimental differences for forgetting when participants were directed to think about one category over the other (and found correlations the between psychological "distance" of directed thought and forgetting), the current study looked at the correlation between naturally occurring off-task thoughts during an ongoing task and overcoming generating new items from memory. Delaney et al. (2010) thus introduced a concentrated period of specific thought content that was unrelated to their memory task, and this could have encouraged more mental context change in their participants compared to my paradigm. (And, beyond the differences between our mind wandering definitions and assessments, this is another comparison of episodic and semantic memory contexts, which further differentiates our studies.)

Future Directions

A first step to better understand the current study's results would be replicating the exploratory finding reported from Experiment 2, that an incubation period increased post-incubation retrieval in the first 60 s after an initial 9 min block of retrieval. If the finding replicates, then incubation may prove to offer (only) temporary benefits to memory.

Additionally, there are outstanding questions about the potential influence of retrieval tasks (Smith & Beda 2020; Smith & Vela 1991; Smith et al., 2017). First, which semantic categories for retrieval count as "flexible" (and facilitate incubation), and which count as "inflexible" (and don't facilitate incubation)? Second, are there genuine differences between episodic and semantic memory tasks for producing incubation effects? If so, one of those differences may be that a semantic memory category must naturally lend itself to a mental context change (through its flexibility) for an incubation period to benefit retrieval.

And for incubation itself, does the incubation task need to purposefully direct the participants' thoughts to task-unrelated topics, as in Delaney et al. (2010), instead of allowing naturally occurring mind wandering to prompt a mental context change, as in my paradigm? Future research may also explore other incubation tasks: The current study used a reading task as a low cognitive demand task, as Sio and Ormerod (2009) showed that linguistic problems benefitted the most from low demand tasks like reading or drawing a picture. Perhaps drawing would have been even more low demand to allow more mind wandering and more of a mental context change. Likewise, Baird et al., (2012) found their incubation effects when using a very simple choice reaction time task (0-back). So, perhaps modelling my incubation period linguistic problems was not helpful, and examining incubation tasks more often used in the divergent thinking literature would show more robust incubation effects.

REFERENCES

Baird, B., Smallwood, J., Mrazek, M. D., Kam, J. W., Franklin, M. S., & Schooler, J. W. (2012). Inspired by distraction: Mind wandering facilitates creative incubation. *Psychological Science*, *23*(10), 1117-1122.

Beaty, R. E., Silvia, P. J., Nusbaum, E. C., Jauk, E., & Benedek, M. (2014). The roles of associative and executive processes in creative cognition. *Memory & Cognition, 42*, 1186-1197.

Beeftink, F., Van Eerde, W., & Rutte, C. G. (2008). The effect of interruptions and breaks on insight and impasses: Do you need a break right now? *Creativity Research Journal*, *20(4),* 358-364.

Bousfield, W. A., & Sedgewick, C. H. W. (1944). An analysis of sequences of restricted associative responses. *The Journal of General Psychology*, *30(2)*, 149-165.

Brown, R., & McNeill, D. (1966). The “tip of the tongue” phenomenon. *Journal of Verbal Learning and Verbal Behavior, 5*(4), 325-337.

Cheyne, J. A., Solman, G. J., Carriere, J. S., & Smilek, D. (2009). Anatomy of an error: A bidirectional state model of task engagement/disengagement and attention-related errors. *Cognition, 111*(1), 98-113.

Choi, H., & Smith, S. M. (2005). Incubation and the resolution of tip-of-the-tongue states. *The Journal of General Psychology*, *132*(4), 365-376.

Delaney, P. F., Sahakyan, L., Kelley, C. M., & Zimmerman, C. A. (2010). Remembering to forget: The amnesic effect of daydreaming. *Psychological Science*, *21(7)*, 1036-1042.

Dodds, R. A., Ward, T. B., & Smith, S. M. (2004). A review of experimental research on incubation in problem solving and creativity. Unpublished doctoral thesis (Texas A&M University).

Faul, F., Erdfelder, E., Lang, A.-G., & Buchner, A. (2007). G*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods*, *39*, 175-191

Fitzgerald, J. M. (1983). A developmental study of recall from natural categories. *Developmental Psychology*, *19*(1), 9.

Gilhooly, K. J., Fioratou, E., Anthony, S. H., & Wynn, V. (2007). Divergent thinking: Strategies and executive involvement in generating novel uses for familiar objects. *British Journal of Psychology, 98(4)*, 611-625.

Goeke, C., Diekamp, D., Standvoß, K., & König, P. (2017). LabVanced: a unified JavaScript framework for online studies.

Gronlund, S. D., & Shiffrin, R. M. (1986). Retrieval strategies in recall of natural categories and categorized lists. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 12(4), 550.

Gruenewald, P. J., & Lockhead, G. R. (1980). The free recall of category examples. *Journal of Experimental Psychology: Human Learning and Memory, 6*(3), 225.

Hart, J. T. (1967). Memory and the memory-monitoring process. *Journal of Verbal Learning and Verbal Behavior*, *6*, 685-691.

Kail, R., & Nippold, M. A. (1984). Unconstrained retrieval from semantic memory. *Child Development*, *55*, 944-951.

Koppel, R. H., & Storm, B. C. (2014). Escaping mental fixation: Incubation and inhibition in creative problem solving. *Memory*, *22*(4), 340-348.

McVay, J. C., & Kane, M. J. (2010). Does mind wandering reflect executive function or executive failure? Comment on Smallwood and Schooler (2006) and Watkins (2008). *Psychological bulletin*, *136*(2), 188–197. https://doi.org/10.1037/a0018298

McVay, J. C., & Kane, M. J. (2012). Why does working memory capacity predict variation in reading comprehension? On the influence of mind wandering and executive attention. *Journal of Experimental Psychology: General*, *141*(2), 302.

Murray, S., Liang, N., Brosowsky, N., & Seli, P. (2021). What are the benefits of mind wandering to creativity? Psychology of Aesthetics, Creativity, and the Arts. Advance online publication. https://doi.org/10.1037/aca0000420

Reingold, E. M., Yang, J., & Rayner, K. (2010). The time course of word frequency and case alternation effects steon fixation times in reading: Evidence for lexical control of eye movements. *Journal of Experimental Psychology: Human Perception and Performance, 36*(6), 1677.

Roediger, H.L., Tekin, E., Uner, O. (2017). Encoding-Retrieval Interactions. *Cognitive Psychology of Memory Vol*. 2 *of Learning and Memory: A Comprehensive Reference, 2nd edition*, Byrne, J. H. (ed.). pp. 5–26. Oxford: Academic Press.

Rosen, V. M., & Engle, R. W. (1997). The role of working memory capacity in retrieval. *Journal of Experimental Psychology: General*, *126*(3), 211.

Sahakyan, L., & Kelley, C. M. (2002). A contextual change account of the directed forgetting effect. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *28*(6), 1064.

Sahakyan, L., & Smith, J. R. (2014). “A long time ago, in a context far, far away”: Retrospective time estimates and internal context change. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *40*(1), 86.

Schelble, J.L., Therriault, D.J. & Miller, M.D. Classifying retrieval strategies as a function of working memory. *Memory in Cognition 40*, 218–230 (2012). https://doi.org/10.3758/s13421-011-0149-1

Seli, P., Konishi, M., Risko, E. F., & Smilek, D. (2018). The role of task difficulty in theoretical accounts of mind wandering. *Consciousness and Cognition*, *65*, 255-262.

Sio, U. N., & Ormerod, T. C. (2009). Does incubation enhance problem solving? A meta-analytic review. *Psychological Bulletin, 135(1),* 94.

Smallwood, J., Nind, L., & O’Connor, R. C. (2009). When is your head at? An exploration of the factors associated with the temporal focus of the wandering mind. *Consciousness and Cognition, 18*(1), 118-125.

Smeekens, B. A., & Kane, M. J. (2016). Working memory capacity, mind wandering, and creative cognition: An individual-differences investigation into the benefits of controlled versus spontaneous thought. *Psychology of Aesthetics, Creativity, and the Arts, 10*(4), 389.

Smith, S. M., & Beda, Z. (2020). Old problems in new contexts: The context-dependent fixation hypothesis. *Journal of Experimental Psychology: General, 149*(1), 192.

Smith, S. M., & Blankenship, S. E. (1991). Incubation and the persistence of fixation in problem solving. *The American journal of psychology*, 61-87.

Smith, S. M., Gerkens, D. R., & Angello, G. (2017). Alternating incubation effects in the generation of category exemplars. *The Journal of Creative Behavior, 51(2),* 95-106.

Smith, S. M., & Vela, E. (1991). Incubated reminiscence effects. *Memory & Cognition, 19(2),* 168-176.

Smith, S. M., & Vela, E. (2001). Environmental context-dependent memory: A review and meta-analysis. *Psychonomic bulletin & review*, *8*(2), 203-220.

Smith, S. M. (1995). Fixation, incubation, and insight in memory and creative thinking. *The creative cognition approach*, *135*, 156.

Steindorf, L., Hammerton, H. A., & Rummel, J. (2020). Mind wandering outside the box—About the role of off-task thoughts and their assessment during creative incubation. *Psychology of Aesthetics, Creativity, and the Arts.* Advance online publication. https://doi.org/10.1037/aca0000373

Thurstone, L. L. (1938*). Primary mental abilities* (Vol. 119). Chicago: University of Chicago Press.

Troyer, A. K., Moscovitch, M., & Winocur, G. (1997). Clustering and switching as two components of verbal fluency: evidence from younger and older healthy adults. *neuropsychology*, *11*(1), 138.

Wren, J. (1891). *Lazy Thoughts of a Lazy Girl.* Good Press.

APPENDIX A: SELECTED VERBATIM INSTRUCTIONS

Cheating Reduction Measures

Below are the instructions regarding non-cheating for the experiment and the following affirmation of non-cheating. This will be given to participants at the start of the experiment.

“This study aims to understand how people think; it is not aimed to see how well you do on tasks. Therefore, please try your best without outside resources, e.g., help from friends and family, search engines, or personal electronic devices.

There are no bonuses for high performance. You cannot receive extra rewards for doing better on tasks than you would have done naturally. Instead, your use of external help would hurt the integrity and validity of our study.

Please try your best without outside assistance and answer any questions honestly.”

Below is the compensation instructions quiz all participants had to answer before seeing task-specific instructions.

“What reward will you receive based on your performance?

1. More if I perform better

2. Same compensation no matter how I perform

3. Less if I perform better”

Below is a non-cheating affirmation participants were asked to answer before seeing the rest of the instructions.

“It is important to us that you do NOT use outside sources like the Internet to search for the correct answer. Will you complete the rest of the study without help from outside sources?

Press 1 for 'Yes, I will answer alone'

Press 2 for 'No, I will answer with outside help'”

Fluency Task Instructions

Block 1

"In this first task you will be given the name of a category. We will ask you to recall as many examples from this category that you can think of in (BLOCK 1: 7 or 9 BLOCK 2: 2) minutes.

We will tell you the category right before we begin. As soon as you think of any member of the category, please type its name. You may recall and report these category examples in any order that they occur to you. After you type the name of each category example, please click the SUBMIT button to record your response. Your response will then disappear and you can type your next example.

Please keep trying to recall new category examples throughout the entire 9 minute period, even if you start having trouble recalling new examples. Your goal is to report as many different examples as you can.

Click Continue to get the category name and begin.

The category is ANIMALS. Please type as many ANIMALS as you can think of in the next task."

Block 2

"Your category is ANIMALS. Please type as many *new* ANIMALS as you can think of in the next 2 minutes, without repeating ones that you recalled earlier in the study."

APPENDIX B: STRATEGY ASSESSMENT

The questionnaire was presented to participants after their second recall task, and they completed it twice, once for the most recent (2 min) recall task and once for the previous, (7/9 min) first recall task.

Which of the following things did you think about to help you think of animals? Please check all that apply:

☐ **Types or classifications of animals** (for example, thinking about all the types of dogs or types of mammals)

☐ **Types of make-believe animals** (for example, thinking about all the animals you've seen in cartoons)

☐ **Going through the alphabet** (for example, thinking of all animals that start with "A", then all that start with "B"…)

☐ **Rhyming** (for example, thinking of dog, then thinking of frog because it rhymes with dog)

☐ **Physical attributes of animals** (for example, all thinking of all the animals with wings or all animals with 4 legs)

☐ **Types of environments animals live in** (for example, thinking of all the desert animals or all the jungle animals)

☐ **Geographic location of animals** (for example, thinking of all the animals that live in Africa or Florida)

☐ **Frequency of animals** (for example, thinking of the most common animals or the rarest animals)

☐ **Location of animals in a zoo** (for example, thinking of all animals the at Sea World or in the monkey house at the zoo)

☐ **Animals affected by humans** (for example, thinking of all the animals that humans eat or that are endangered)

☐ **Animals you like, do not like or fear** (for example, thinking of a snake you fear, then thinking of a rat because you also fear rats)

☐ **Animals that live with humans** (for example, thinking of all domesticated animals or all the animals on a farm)

☐ **Animals personally relevant to you** (for example, animals you have had as pets, animals you have seen recently)

☐ **Using free association or random generation** (such as saying whatever animals popped into your head)

☐ **Association with previous animal** (for example, "cat reminds me of dog, which reminds me of wolf")

☐ **Other** (Please explain if not listed on other side) ____________________________________

____

____

APPENDIX C: READING COMPREHENSION QUIZ QUESTIONS

1. Why would the author’s dog not lie in front of the fire?
    a. He simply did not want to
    b. He was not allowed because of his liver *
    c. He was too old to lie down
    d. He ran warm so it was uncomfortable
2. Does the author think of dogs as close to humans?
    a. Yes, because they have the vices and virtues of humans *
    b. No, because they don’t have the vices and virtues of humans
    c. Yes, because it would be sad if they weren’t
    d. No, because it would be silly if they were
3. What did the author call the activity where the dog would jump over obstacles?
    a. Steeplechases*
    b. Obstacle course
    c. Fence jumping
    d. Water jumps
4. Has the author ever been growled at?
    a. Yes, by one dog
    b. No, despite meeting several dogs*
    c. Yes, but not by dogs
    d. No, city dogs are well-behaved.