JCB

Publication of the Creative Education Foundation ®

STEVEN M. SMITH
DAVID R. GERKENS
GENNA ANGELLO

Alternating Incubation Effects in the Generation of Category Exemplars

ABSTRACT

Four experiments tested the forgetting fixation hypothesis of incubation effects, comparing continuous vs. alternating generation of exemplars from three different types of categories. In two experiments, participants who listed as many members as possible from two different categories produced more responses, and more novel responses, when they alternated back and forth between the two categories, as compared to continuous uninterrupted listing from each of the two categories. This incubation effect was not found in Experiment 1, when participants were given taxonomic categories (*birds* and *clothing*) for the generation task, but was found in Experiment 2 with sense impression categories (*cold things* and *heavy things*), and in Experiment 3 with *ad hoc* categories (*equipment you take camping* and *fattening foods*). A similar incubation effect was observed in Experiment 4 when a non-verbal task was given between category generation tasks, but only for flexibly defined categories. The results suggest that forgetting from one alternating listing period to the next in the form of altering category cue representations was consistent with the observed incubation effects. These *alternating incubation* effects have implications for understanding cognitive processes that underlie creative cognition.

Keywords: creativity, cognition.

When people try to think of creative ideas, they often become stuck on ordinary ideas, or ones encountered recently, a phenomenon known as *fixation* (e.g., Jansson & Smith, 1991; Smith & Blankenship, 1989, 1991). If additional work on a fixated problem is done, it has been found that more solutions or ideas can be produced after a delay (an *incubation interval*), as compared to immediate additional work—an *incubation effect* (e.g., Browne & Cruse, 1988; Christensen & Schunn, 2005; Dodds, Smith & Ward, 2002; Guilford, 1979; Kohn & Smith, 2009; Moss, Kotovsky, & Cagan, 2007; Norlander & Gustafson, 1996; Olton, 1979; Patrick, 1986; Penney, Godsell, Scott, & Balsom, 2011; Sio & Ormerod, 2009; Smith & Blankenship, 1989, 1991; Vul & Pashler, 2007). Olton (1979) defined incubation as “a facilitation of thinking that is evident after an interval during which no conscious work was done on the task, assuming an earlier period of substantial conscious work” (p. 11). There have been several theories concerning the cause(s) of incubation effects, including the unconscious work theory (e.g., Bowers, Regehr, Balthazard, & Parker, 1990; Yaniv & Meyer, 1987), the opportunistic assimilation theory (e.g., Seifert, Meyer, Davidson, Patalano, & Yaniv, 1994), and the forgetting fixation theory (e.g., Smith, 1995a,b; Smith & Blankenship, 1989, 1991). Incubation effects may be multiply caused (e.g., Smith, 1995a,b), that is, there could be more than a single cause of various observed effects.

The goal of the present study was not to address creativity in an ecologically realistic context, but rather to examine a very narrow aspect of creative thinking with controlled experimental studies, using what has been termed a *creative cognition* approach to the study of creativity and to our understanding of the creative aspects of cognition (e.g., Finke, Ward, & Smith, 1992; Smith, Ward, & Finke, 1995). The creative cognition approach acknowledges that cognition represents only one slice of creativity—others include motivation to create, social and collaborative aspects of creativity, cultural considerations, historical contexts, and individual differences. Furthermore, there are many cognitive processes that collaborate in many ways to produce creative ideas and products, not a singular *creative process* (e.g., Smith et al., 1995). The many processes include analogical reasoning and transfer (e.g., Christensen & Schunn, 2005), conceptual combination and extension (e.g., Ward, 1995; Ward, Smith, & Vaid, 1997), visualization (e.g., Finke, 1996; Shah, Millsap, Woodward, & Smith, 2012), and memory processes (e.g., Angello, Storm, & Smith, 2014; Smith & Vela, 1991). Understanding how these cognitive processes operate is important for a greater understanding of creativity, and for a better understanding of the ways in which cognition is inherently creative (Smith et al., 1995). In the present study, we examine cognitive processes, including memory retrieval, attentional switching, and conceptual structures, that may be involved in incubation effects.

In the present study, we do not test competing predictions of various theories of incubation; rather, we test predictions of *one* of those theories, the forgetting fixation theory (e.g., Smith, 1995a,b; Smith & Blankenship, 1989, 1991). The *forgetting fixation* hypothesis states that when fixated ideas (ones that initially block creative ideation) are reduced in accessibility (forgotten), then other, more novel ideas become more accessible. In comparison to theories that explain incubation in terms of serendipitous stimuli that can trigger creative ideas, such as the opportunistic assimilation theory, the forgetting fixation theory explains incubation as something that undermines counterproductive work. The forgetting fixation theory, alone, cannot explain the insight that may occur at the end of an incubation interval; cognitive re-structuring (e.g., Knoblich, Ohlsson, Haider, & Rhenius, 1999; Metcalfe & Weibe, 1987), that is, thinking about the fixated problem in a different way, must still occur before an incubation effect is successful. Forgetting fixation, according to this view, is an important step that enables cognitive re-structuring.

Empirical tests of the forgetting fixation hypothesis have supported the idea that diminished accessibility of inappropriate information that blocks problem-solving can lead to incubation effects (e.g., Kohn & Smith, 2009; Smith & Blankenship, 1989, 1991; Vul & Pashler, 2007). Smith and Blankenship (1989) found correlational evidence consistent with the forgetting fixation hypothesis; longer delays improved re-test performance on initially unsolved insight problems, and delays also decreased recall of words (blockers) used initially to block problem-solving. Furthermore, experimentally reducing the accessibility of blockers via directed forgetting (Koppel & Storm, 2014) or via repeated suppression (Angello et al., 2014) improved performance on verbal insight problems; forgetting blockers benefited resolution of unsolved insight problems. These studies provide experimental evidence of the forgetting fixation hypothesis.

Many of the reported studies of incubation effects have used various insight problems, such as Remote Associate Problems (e.g., Dodds et al., 2002; Kohn & Smith, 2009; Smith & Blankenship, 1991; Vul & Pashler, 2007). The few published incubation studies involving divergent thinking tasks have reported robust incubation effects (e.g., Fulgosi & Guilford, 1968; Gilhooly, Georgiou, Garrison, Reston, & Sirota, 2012; Kohn & Smith, 2010; Shah, Smith, Vargas-Hernandez, Gerkens, & Wulan, 2003). Divergent thinking tasks, such as the alternate uses test (Torrance, 1974), are often used to measure creativity as an individual difference measure (e.g., Goff & Torrance, 2002) or as a measure of creative production (e.g., Kerne & Smith, 2004; Shah et al., 2003), and they serve as the basis for traditional group brainstorming (e.g., Osborn, 1957; Paulus, Nakui, Putman, & Brown, 2006). Although the four different measures of divergent thinking tasks that are usually examined are fluency (number of ideas), originality, flexibility (number of categories of ideas), and quality, experimental studies using divergent thinking tasks usually report only fluency, and occasionally novelty or originality (e.g., Gilhooly et al., 2012; Kohn & Smith, 2010; Shah et al., 2003).

Barsalou, who pioneered the study of *ad hoc* categories (Barsalou, 1982, 1985), defined them as natural categories that are "created spontaneously for use in specialized contexts," giving as examples "things to take on a camping trip," "possible costumes to wear to a Halloween party," and "places to look for antique desks" (p. 211). Such categories, Barsalou noted, are no different from divergent thinking tasks, such as alternate uses tests, which measure creative production. These *ad hoc* types of categories, according to Barsalou, are particularly sensitive to contextual influences, lacking the stable structure and the high correlation of features among category members seen in most taxonomic categories, whose structures are well-established, and whose exemplars are likely to share a particular set of features. One of the questions asked in the present study was whether category generation (i.e., the task of listing as many members of a category as possible) would show incubation effects, as have been found for divergent thinking tasks, and if so, whether those incubation effects would depend on the types of categories used.

The forgetting fixation theory predicts incubation effects in category generation because drawing some category members from memory causes fixation in the form of output interference, and a subsequent incubation interval engenders forgetting of the already-output items, thereby relieving output interference (Smith & Vela, 1991). The buildup of output interference during category generation, and the subsequent reduction in interference when interfering items are forgotten, can be explained by a memory retrieval theory, such as the SAM model (e.g., Raaijmakers & Shiffrin, 1981). The theory states that retrieval begins by delimiting a search set to items in memory with non-negligible probabilities of being retrieved by a particular cue; items within that set are sampled and replaced (Raaijmakers & Shiffrin, 1981, 1992). Retrieval functions like a study trial, increasing the subsequent probability of retrieving already-retrieved items when the same retrieval cues are used (e.g., Raaijmakers & Shiffrin, 1981; Rundus, 1973). After many items are sampled from a search set, sampling attempts are more likely to produce already-retrieved items than items not previously retrieved, a *biased search set* (e.g., Raaijmakers & Shiffrin, 1981). Once a search set is biased, the chance of finding as yet unretrieved members of the search set becomes diminishingly small.

This study involves category generation tasks, in which participants retrieve multiple members from categories. Generating exemplars from categories resembles the creative ideation processes that underlie the way people generate a set of creative ideas for solving a problem, consistent with Ward's theory of structured imagination (e.g., Ward, 1994), which links the graded structure of categories with generation of creative ideas based on those categories. Because dominant category members are likely to be sampled first, a biased search set should consist of the dominant, or commonly listed members (e.g., Barsalou, 1985), making novel category members effectively inaccessible. Therefore, relief from retrieval bias can be caused by an incubation interval, and should be measurable in terms of the number of category members retrieved, and the mean normative novelty of the generated exemplars, measures used to assess incubation effects in studies that used divergent thinking tasks (e.g., Kohn & Smith, 2010; Shah et al., 2003).

Incubation effects in memory and problem-solving might be caused by decay of initially retrieved exemplars, or by re-delimiting a search set. When a search set is delimited, relevant items may initially fail to be included. A redefined search set might include relevant items not in the originally defined set. If search set delimitation is influenced by context (e.g., Raaijmakers & Shiffrin, 1981), then changing temporal contexts that accompany incubation intervals could lead to a redefined search set, a new one that might include previously inaccessible items. Raaijmakers and Shiffrin (1992) referred to this contextually altered way of searching memory as "a change in the nature of the cue between time A and time B—i.e., a change in the (functional) stimulus" (pp. 220–221). Either or both of these mechanisms, redefining a search set and/or decay, could account for incubation effects in a category generation task.

We used three different types of categories: taxonomic categories (*birds* and *clothing*), *sense impression* categories (*cold things* and *heavy things*), and *ad hoc* categories (*foods that are fattening*, and *equipment you take camping*). Taxonomic categories are well-known, with relatively stable representations that should not be affected by brief changes in temporal context. Sense impression and *ad hoc* categories are flexible, with exemplars that can be drawn from multiple taxonomic categories. If retrieved members of any of these types of categories are retrieved, that should cause output interference because of an increasing bias to retrieve already-retrieved items. If incubation intervals cause forgetting of interfering (already-retrieved) items via *decay*, then our experiments should find incubation effects for all three types of categories. If, however, incubation intervals cause forgetting of interfering items via *restructuring of the search set*, then *ad hoc* and sense impression categories should be more likely than taxonomic categories to show incubation effects, because *ad hoc* and sense impression categories have less stable category structure than taxonomic categories.

THE PRESENT EXPERIMENTS

In the present experiments, we explored a new type of incubation effect that we refer to as an *alternating incubation effect*. In many studies of incubation effects (e.g., Browne & Cruse, 1988; Dodds et al., 2002; Kohn & Smith, 2009; Smith & Blankenship, 1989, 1991; Vul & Pashler, 2007), there are two intervals of work on a problem, including the initial period of work on the problem, which is typically the same for control and incubation conditions, and a second interval of work that either comes immediately after the initial period (the control condition) or after an incubation interval (the incubation condition). This type of procedure ensures that the total time given to work on a problem is equal for the two conditions, but the incubation condition takes longer, overall, because of the addition of the incubation interval. In the present experiments, the control and incubation conditions provide the same amount of time for participants to work on a problem, and the two treatments also provide the same overall amount of time. That is, we looked for incubation effects without providing extra time for the incubation condition.

In the present experiments, participants were given either continuous work, first on one, and then on a second category generation task (the control condition), or they switched back and forth between the two different tasks. Output interference in one category (i.e., from temporarily strengthened items that are already retrieved) should dissipate while attention is switched to generating members of the other category. We tested this *alternating incubation* hypothesis; given equal amounts of time listing members of two different categories, alternating between the two different categories was predicted to lead to more responses, and more novel responses, than a condition in which category generation is first continuous for one category, followed by continuous generation from the second category. In addition to our primary hypothesis, which predicted more responses and more novel responses in the alternating condition relative to the continuous (control) condition, we also predicted that the alternating incubation effect would be greater for less stable types of categories, such as sense impression and *ad hoc* categories, relative to categories with more stable category structures, such as taxonomic categories. The reason for this secondary prediction was that less stable category representations should be more susceptible to forgetting of interfering items via search set restructuring.

METHOD FOR ALL EXPERIMENTS

In our comparison of alternating vs. continuous work on two category generation tasks, participants in both conditions had 6-minute to list members from two different categories. Those in the continuous condition listed members of category A for 3-minute, then category B for 3-minute. Participants in the alternating condition switched back and forth from category A to category B, alternating every minute for 6-minute. Thus, in the alternating condition, the periods of time spent listing members of category B served as incubation intervals for category A, and vice versa.

In Experiment 4, we examined whether alternating between category generation and mazes would cause the same alternating incubation effects.

PARTICIPANTS

A total of 49 student volunteers participated in Experiment 1, 60 in Experiment 2, 59 in Experiment 3, and 180 in Experiment 4. Students self-enrolled for the studies, and had the choice of participating in any of several psychology experiments, or writing a short paper for partial course credit. Assignment to treatment groups was random. In Experiment 1, 20 participants enrolled in the continuous listing condition, and 29 in the alternating condition. In Experiment 2 there were 28 continuous and 31 alternating. In Experiment 3 there were 29 continuous and 31 alternating. In Experiment 4, 30 participants were randomly assigned to each of the six treatment conditions.

MATERIALS

In Experiments 1, 2, and 3, category names were presented at the top of the screen with a text box in which participants could type responses. In Experiment 1, two taxonomic category names were presented, *birds* and *clothing*. In Experiment 2, two sense impression category names were presented, *cold things* and *heavy things*. In Experiment 3, two *ad hoc* category names were presented, *things to take camping* and *fattening foods*. In Experiment 4, all three types of categories were used, but the presentations were given on paper rather than on a computer, so that maze puzzles could be given in the alternating conditions. One taxonomic category (*birds*), one *ad hoc* category (*fattening foods*), and one sense impression category (*heavy things*) were used for the category generation task.

DESIGN AND PROCEDURE

Experiments 1, 2, and 3 used a 2 (category) × 2 (listing method) design. Category was a within-subjects factor. Each participant listed members for both categories, and category order was randomly determined for each participant. Listing method, a between-subjects factor, was either continuous or alternating; participants either listed members for a given category for 3-minute without interruption (continuous), or listed for a total of 3-minute in 1-minute intervals separated by listing members for the other category (alternating). In Experiment 4, a 3 (category type) × 2 (listing method) design was used. Category type was taxonomic, *ad hoc*, or sense impression, and listing method was continuous or alternating; both were between-subjects factors.

In Experiments 1, 2, and 3 participants saw a category name (e.g., *birds*) on a computer screen and began listing as many category members as they could. Once time expired for a category, an alert appeared on the screen ending the listing session. In Experiment 4, participants listed members for only one category for 3-minute. Listing method, a between-subjects variable, was continuous or alternating; participants either listed members for a single category for 3-minute continuously, or listed in three 1-minute intervals, with each interval separated by a minute of working on maze problems (alternating). After each minute, participants turned a page. In the continuous condition, they continued with the same category. In the alternating condition participants had a maze to complete (as an incubation task) for two 1-minute periods, once after the first minute of category generation, and again after the second minute. Following each maze, participants tried to list additional category members.

The dependent measures were quantity (number listed per participant), and mean novelty (total number of participants in a condition divided by the number who listed a specific category member). Larger novelty scores were more novel. If a participant listed a particular category member twice or more, it was counted only once.

EXPERIMENT 1

The primary forgetting fixation hypothesis predicted that listing category members from two different taxonomic categories would result in more responses (greater *quantity*), and more *novel* responses, under alternating conditions than with continuous uninterrupted blocks of category generation. Our secondary hypothesis, however, stated that such incubation effects may occur only for relatively unstable categories whose momentary representations are contextually influenced, such as *ad hoc* and sense impression categories. Because representations of taxonomic category cues, such as *birds* and *clothing*, should be relatively stable, less influenced by transitory contexts, then incubation effects should be small or absent with taxonomic categories.

RESULTS

A 2 × 2 mixed analysis of variance (ANOVA) was computed, using category (*birds* or *clothing*) as a within-subjects variable, listing method (continuous or alternating) as a between-subjects variable, and quantity (number of responses) as the dependent variable. The main effect of category was significant [$F(1, 47) = 56.37$, $p < .001$ $\eta^2 = .55$]; more responses were listed for *clothing* than for *birds* (Table 1). Neither the main effect of listing method, nor the interaction of listing method × category were significant [both Fs < 1.0].

Another 2 × 2 mixed ANOVA was computed, using category (*birds* or *clothing*) as a within-subjects variable, listing method (continuous or alternating) as a between-subjects variable, and mean response novelty as the dependent variable. Novelty for each *response* was calculated as the total number of participants in the experiment divided by the number of those participants who listed a specific category member; thus, the novelty score of each response was a measure of its infrequency of occurrence among participants in our sample. The mean novelty score of all of a *participant's* responses for a category was used in this analysis. Greater novelty scores corresponded to less commonly given responses. This method for scoring novelty does not rely on subjective scoring, but rather on the empirically determined statistical infrequency of every response, a method used by others (e.g., Kohn & Smith, 2010; Shah et al., 2003). Neither main effect of listing method or category were significant, nor was the interaction of the two variables [all Fs < 1.0].

In Experiment 1, there were no cases in which a particular response was given for both *birds* and *clothing* categories, either in the continuous or alternating conditions.

TABLE 1. Mean Quantity and Novelty as a Function of Listing Condition in Experiments 1, 2, and 3

Category	Condition	Quantity	Novelty
Taxonomic categories (Experiment 1)			
Clothing	Continuous	26.79 (1.34)	2.66 (.29)
Clothing	Alternating	25.80 (1.62)	3.08 (.35)
Birds	Continuous	17.72 (1.30)	3.04 (.18)
Birds	Alternating	16.30 (1.56)	3.02 (.21)
Sense impression categories (Experiment 2)			
Heavy things	Continuous	16.45 (1.00)	4.54 (.26)
Heavy things	Alternating	21.48 (1.03)	5.40 (.28)
Cold things	Continuous	16.03 (.95)	2.78 (.14)
Cold things	Alternating	19.31 (.98)	3.73 (.14)
Ad hoc categories (Experiment 3)			
Fattening foods	Continuous	17.00 (1.12)	3.33 (.19)
Fattening foods	Alternating	20.29 (1.18)	3.81 (.19)
Camping equipment	Continuous	20.23 (1.15)	3.27 (.14)
Camping equipment	Alternating	24.17 (1.21)	3.53 (.15)

Note. Standard Errors are shown in parentheses.

DISCUSSION

No evidence of incubation effects was found in Experiment 1, neither in terms of number of responses, nor response novelty. These results are consistent with our secondary hypothesis that alternating incubation effects in category generation will not be found for taxonomic categories because the relative stability of such categories makes it less likely that a brief task switch would lead to forgetting and re-structuring of the search set used to search for category members.

EXPERIMENT 2

Sense impression categories (e.g., *cold things*, *heavy things*) are defined in terms of features that all of the category members have in common, and that are coded by the senses, such as visual properties (e.g., *blue things*, *round things*), auditory properties (e.g., *loud things*, *musical things*), or tactile properties (e.g., *hot things*, *slippery things*). Flexibility in the representation of a sense impression category makes it contextually sensitive. Therefore, during a period of incubation, the functional category cue (i.e., the way one momentarily thinks about that category) can change. The forgetting fixation hypothesis predicted incubation effects in alternating listing conditions, relative to continuous listing, for these flexibly defined categories; better performance for the alternating listing condition was predicted for both quantity and novelty measures.

RESULTS

A 2 × 2 mixed ANOVA was computed, using category (*cold things* or *heavy things*) as a within-subjects variable, listing method (continuous or alternating) as a between-subjects variable, and quantity (number of responses) as the dependent variable. The main effect of listing method on quantity was significant [$F(1, 58) = 14.25$, $p < .001$, $\eta^2 = .20$]; the alternating condition produced more responses than did the continuous condition (Table 1). Neither the main effect of category [$F(1, 58) = 2.24$, $p = .14$, $\eta^2 = .04$], nor the category × listing method interaction [$F(1, 58) = 1.02$, $p = .32$, $\eta^2 = .02$] were significant.

A second 2 × 2 mixed ANOVA was computed, using category as a within-subjects variable, listing method as a between-subjects variable, and mean response novelty as the dependent variable. The main effect of listing method on novelty was significant [$F(1, 58) = 16.62$, $p < .001$, $\eta^2 = .22$]; the alternating condition produced more novel responses than did the continuous condition (Table 1). The main effect of category on novelty was also significant [$F(1, 58) = 74.06$, $p < .001$, $\eta^2 = .68$]; more novel responses were given for heavy things than for cold things. The category × listing method interaction was not significant [$F(1, 58) < 1.0$].

Out of a total of 405 exemplars generated by participants in Experiment 2, there were only 6 cases in which a particular response was given for both *cold things* and *heavy things* in the alternating condition, as compared with 7 cases in the continuous listing condition.

DISCUSSION

As predicted by the forgetting fixation hypothesis, a benefit for the alternating listing condition, relative to the continuous listing procedure, was found both in terms of the number of responses, and the novelty of those responses. The only difference between

Experiment 1, in which no incubation effect was found, and Experiment 2, which found the predicted incubation effect, was the use of sense impression categories, rather than taxonomic categories. Because sense impression categories can be represented in multiple ways, the advantage of the alternating listing condition in Experiment 2 can be attributed to changes in functional category cues, that is, changes in the cognitive representations of cues, from one 60-second listing period to the next.

EXPERIMENT 3

In Experiment 3, we repeated the experimental procedure using *ad hoc* categories (e.g., *foods that are fattening* and *equipment you take camping*), which, like sense impression categories, are open-ended, and can contain members from multiple taxonomic categories. Because *ad hoc* categories can be cognitively represented in multiple ways, they are susceptible to context changes, and should show incubation effects. Because *ad hoc* categories, like sense impression categories, are flexibly structured, it was predicted that more category members, and more novel responses, would be generated in the alternating condition than in the continuous listing condition.

RESULTS

A 2 × 2 mixed ANOVA was computed, using category (*foods that are fattening* or *equipment you take camping*) as a within-subjects variable, listing method (continuous or alternating) as a between-subjects variable, and quantity (number of responses) as the dependent variable. The main effect of listing method on quantity was significant [$F(1, 57) = 7.17$, $p < .01$, $\eta^2 = .11$]; the alternating condition produced more responses than did the continuous condition (Table 1). The main effect of category was also significant [$F(1, 57) = 14.25$, $p < .001$, $\eta^2 = .20$]; more responses were listed for *equipment you take camping* than for *foods that are fattening* (Table 1). The category × listing method interaction was not significant [$F(1, 57) < 1.0$].

A second 2 × 2 mixed ANOVA was computed, using category as a within-subjects variable, listing method as a between-subjects variable, and mean response novelty as the dependent variable. The main effect of listing method on novelty was significant [$F(1, 57) = 3.99$, $p < .05$, $\eta^2 = .07$]; the alternating condition produced more novel responses than did the continuous condition (Table 1). Neither the main effect of category [$F(1, 57) = 1.08$, $p = .30$, $\eta^2 = .02$], nor the category × listing method interaction [$F(1, 57) < 1.0$] were significant.

Of a total of 412 exemplars generated by participants in Experiment 2, there were 7 cases in which a particular response was given for both *camping equipment* and *fattening foods* categories in the alternating condition, as compared with 5 cases in the continuous condition.

DISCUSSION

An alternating incubation effect was again found in Experiment 3. As in Experiment 2, this incubation effect was seen in the quantity of responses and the average novelty of responses. Because *ad hoc* categories, like sense impression categories, can be flexibly represented, these results can be attributed to the likelihood that alternating listing conditions encouraged varied representations of *ad hoc* category cues from one 60-second period to the next. Given changes in the functional category cues, a type of forgetting, the result was that different category members were accessed after a break from category generation, resulting in an incubation effect.

EXPERIMENT 4

Experiments 1, 2, and 3 demonstrated that the benefits of alternation surpass the costs in category generation, but only for loosely defined categories, such as *ad hoc* or sense impression categories. Benefits were seen in measures of the quantity of responses, and to some degree, the average normative novelty of the generated category exemplars. Those in the alternating condition were interrupted and switched to a new problem of the same type. That is, although they switched from generating members of one category to members of another category, they never switched to an altogether different task, the traditional way that incubation intervals are operationally defined in experimental studies (e.g., Smith & Blankenship, 1989, 1991). In Experiment 4, we examined the alternating incubation effect, using this more traditional definition of an incubation task. Rather than alternating between one category generation problem and another, as in Experiments 1, 2, and 3, participants in Experiment 4 alternated between a category generation task and a maze-solving task.

Based upon the results of the first three experiments, it was predicted that planned comparisons would show incubation effects in Experiment 4 for *ad hoc* and sense impression categories, but not for taxonomic categories. It was predicted that comparisons of continuous vs. alternating listing conditions would show more responses, and more novel responses for the alternating condition, but that a comparison of those listing conditions for taxonomic categories would not show effects for either measure.

RESULTS

A 2 × 3 between-subjects ANOVA was computed, using listing method (continuous vs. alternating) and category type (taxonomic vs. *ad hoc* vs. sense impression) as independent variables, and quantity of responses as the dependent measure. There was a significant effect of listing method [$F(1, 174) = 5.82$, $p = .02$, $\eta^2 = .03$]; more responses were given in the alternating condition ($M = 21.93$, $SE = 0.65$) than the continuous listing condition ($M = 19.96$, $SE = 0.52$), an alternating incubation effect (Table 2). There was also a main effect of category type [$F(2, 174) = 3.71$, $p = .03$, $\eta^2 = .04$]; fewer responses were listed for the taxonomic category ($M = 19.43$, $SE = 0.62$) than for the *ad hoc* ($M = 21.30$, $SE = 0.74$) and sense impression categories ($M = 22.10$, $SE = 0.78$, Table 2).

The interaction of category type and listing method was not significant [$F(2, 174) = 1.22$, $p = .30$, $\eta^2 = .01$]. Planned comparisons contrasted alternating vs. continuous

TABLE 2. Mean Quantity and Novelty for Taxonomic, Sense Impression, and *Ad Hoc* Categories in Experiment 4 as a Function of Listing Condition

Category type	Listing condition	Quantity	Novelty
Taxonomic	Continuous	19.20 (0.76)	6.01 (0.63)
Taxonomic	Alternating	19.67 (1.00)	5.08 (0.56)
Sense impression	Continuous	20.30 (1.01)	17.10 (1.23)
Sense impression	Alternating	23.90 (1.12)	18.82 (1.40)
Ad hoc	Continuous	20.37 (0.92)	14.08 (0.99)
Ad hoc	Alternating	22.23 (1.16)	14.62 (1.35)

Note. Standard Errors are shown in parentheses.

listing conditions, one for each of the three category types. The effect of listing condition was significant for the sense impression category [$t(58) = 2.39$, $p = .02$]; 18% more responses were listed in the alternating condition ($M = 23.90$, $SE = 1.12$) than in the continuous listing condition ($M = 20.30$, $SE = 1.01$), an alternating incubation effect. For the *ad hoc* category condition, although there was a 9% increase in the number of responses in the alternating condition relative to the continuous condition, the effect was not significant [$t(58) = 1.26$, $p = .21$]. For the taxonomic category condition, the effect of listing condition was not significant [$t(58) = .37$, $p = .71$].

A second 2 × 3 between-subjects ANOVA was computed, using listing method (continuous vs. alternating) and category type (taxonomic vs. *ad hoc* vs. sense impression) as independent variables, and mean novelty of responses as the dependent measure. The same norm for category novelty from Experiments 1–3 was used to score responses in Experiment 4. There was a significant effect of category type [$F(2, 174) = 69.73$, $p < .001$, $\eta^2 = .45$]; less novel responses were given for the taxonomic category ($M = 5.54$, $SE = 0.42$) than for sense impression ($M = 17.96$, $SE = 0.93$) and *ad hoc* categories ($M = 14.35$, $SE = 0.83$, Table 2). Neither the main effect of listing method nor the listing method × category type interaction approached significance (both Fs < 1.0). Although numerically the mean novelty was slightly better for the alternating conditions with *ad hoc* and sense impression categories (Table 2), planned comparisons failed to find a significant difference between alternating vs. continuous conditions for any category type [for the taxonomic category, $t(58) = 1.10$, $p = .28$; for the sense impression category, $t(58) = .93$, $p = .36$; and for the *ad hoc* category, $t(58) = .32$, $p = .75$].

DISCUSSION

Experiment 4 found an alternating incubation effect using an incubation task (solving mazes) that differed considerably from category generation. This effect was found for the quantity of ideas, but was not reliable for the average novelty of ideas. The incubation effect (in quantity of responses) was not seen at all for the taxonomic category, and it was significant for the sense impression category. Although not significant, a trend in the expected direction was seen for the *ad hoc* category. The main effect of incubation on quantity of ideas across all category types, and the simple effect of incubation on quantity for the sense impression category, show that shifting to a different type of task can lead to an alternating incubation effect in category generation.

Although incubation effects were found in Experiment 4, they were not as strong as they were in Experiments 2 and 3. The reasons for this are not clear. The most obvious difference between Experiment 4 and the previous experiments concerns the incubation task. It is possible, for example, that listing members of other categories during a break (Experiments 2 and 3), rather than engaging in a non-verbal task (Experiment 4), may have inspired new ideas. A participant frustrated in attempts to think of additional members of one category might find inspiration, or even more subtle spreading activation, triggered by their listing of members of the alternating category. Such verbally based spreading activation, and subsequent opportunistic assimilation (Seifert et al., 1994), might not occur for the non-verbal mazes used in Experiment 4. Furthermore, participants might have found the maze task more fun than category listing, which might have influenced responses. These alternative explanations, however, are speculative, and are not critically tested by the present experiments.

It is also apparent that, for whatever reason, the novelty scores were considerably higher, and more variable, for flexible categories in Experiment 4 than in Experiments 2 and 3. This greater variability could have diluted the incubation effect in novelty scores for flexible categories in Experiment 4. One possible reason for this greater variability was that the norms used to compute novelty scores in Experiment 4 were derived from responses given in Experiments 1, 2, and 3, a different sample of participants. Furthermore, the norms assembled in Experiments 1, 2, and 3 engaged moderate-sized samples of 49, 60, and 59 participants, respectively, larger than the samples used by Barsalou in his original *ad hoc* category norms (Barsalou, 1982, 1985), but not as large as the sample of participants used elsewhere (e.g., Van Overschelde, Rawson, & Dunlosky, 2004). If a sample size is too small, then the corresponding norm might not include rare or idiosyncratic responses; a participant listing such responses might, therefore, have a disproportionately high novelty score. Idiosyncratic responses given by participants in Experiment 4 of this study, that is, responses that were not included in our smaller sample norms, could have artificially inflated novelty scores for some participants in Experiment 4, thereby increasing the variability among novelty scores, and dampening the effect size for that measure.

Finally, it should be noted that Experiment 4 engaged a sample of participants that were different from those who participated in earlier experiments. Trying to understand differences among students in different experiments, who were sampled from different semesters with different instructors, would be highly speculative.

GENERAL DISCUSSION

Our results both support and qualify the forgetting fixation hypothesis of incubation effects. When participants alternated between two different category generation tasks involving flexible *ad hoc* or sense impression categories, or if they alternated between category generation and maze problems, compared to an equal amount of uninterrupted category generation, they generated more category exemplars (Experiments 2, 3, and 4), and more novel ones (Experiments 2 and 3). Alternating incubation effects were found only for flexibly represented sense impression and *ad hoc* categories, not for more rigidly structured taxonomic categories. The forgetting fixation theory supports these results, explaining that initial retrieval results in a biased retrieval set, where dominant, already-retrieved category members block retrieval of less dominant members; time away from that generation task permits enough forgetting to mitigate the biased retrieval set, thereby allowing retrieval of less dominant exemplars.

Because *ad hoc* and sense impression categories are flexible, and restructuring of cognitive representations might occur during an experimental session, forgetting from one listing period to the next in the form of changes in functional category cues could explain the alternating incubation effects observed in category generation. Our failure to observe alternating incubation effects with taxonomic categories is consistent with findings reported by Glaser (2008), where participants generated members of taxonomic categories for 3-minute either continuously, or switching between categories. Switch costs were observed, both in terms of the number of items generated and latency after a switch. When switch costs, in terms of latency, were taken into account, then a production advantage could be observed as a result of switching. Given that there are both costs and benefits of switching between taxonomic categories, the net effect of alternating conditions was apparently nullified. Longer periods of listing taxonomic category members might accommodate the time costs of switching, and might result in observable alternating incubation effects.

The benefits of alternation for more flexibly represented categories may be due to reasons other than, or in addition to our explanation that alternation increases opportunities for category restructuring. One alternative explanation states that cross-fertilization between generation tasks could occur; participants might opportunistically assimilate ideas generated from one category, thereby providing (directly or via spreading activation) novel ideas for the other category generation tasks (Seifert et al., 1994). For example, if a participant thought of "ice" as a response for *COLD THINGS*, that might trigger "iceberg" as a response for *HEAVY THINGS*. This explanation might also explain the weaker alternating benefit for the maze condition in Experiment 4, if mazes are not as effective as word generation in terms of activating semantically useful exemplars for category generation. Counter to this explanation, however, was the finding that very few responses were given in both of a participant's category listings (e.g., listing "iceberg" for both *COLD THINGS* and *HEAVY THINGS*); there were none for taxonomic categories, and only 1.5% of responses for *ad hoc* and sense impression categories. Furthermore, those doubly listed responses were about equally frequent for alternating and continuous listing conditions. Of course, doubly listed responses may not reflect all influences of spreading semantic activation in alternating conditions; thinking of "snow" as a member of *COLD THINGS* could trigger thoughts of "iceberg" for *HEAVY THINGS*. It is not clear, however, why cross-fertilization could not also occur in taxonomic categories; it seems equally reasonable that "boots" listed for *CLOTHING* might trigger the response "ostrich" for *BIRDS*, or that "penguin" listed for *BIRDS* could activate "tuxedo" for *CLOTHING*, yet no alternating listing advantage was found for taxonomic categories. More empirical research will be necessary to pin down the mechanism or mechanisms underlying the alternating incubation effects we have reported.

Our method for assessing novelty differs from methods used in many studies of divergent thinking tasks (see, for example, Plucker, Qian, & Wang, 2011; Silvia, 2008; Sio & Ormerod, 2009). A common method for measuring novelty is with subjective ratings (e.g., Silvia, 2008), in spite of the human bias that is inherent in such ratings, particularly with such an abstract concept as novelty that raters must judge. A less subjective method for measuring novelty (or originality) has been to tally the number of unique or unusual ideas generated by an individual. As Plucker et al. (2011) noted, however, such methods confound fluency (measured by the quantity of ideas) with novelty, because the more ideas that are generated, the more unique or unusual responses are likely to be generated. The method for measuring novelty in the present experiments was used to avoid the use of subjective measures, but it does not confound novelty and fluency the way that tallies of unique ideas does. In the present experiments, each response (compiled across all participants in an experiment) had a normative infrequency (novelty) score (calculated for all of the experiment's participants). The novelty score for a participant was calculated as the mean novelty for the set of all of the participant's responses; thus, every response contributes to the participant's novelty score. With the tally method, only a percentage of responses (i.e., the unique or unusual ones) are counted toward the novelty score, and if that percentage is similar at various levels of fluency, it makes sense that novelty and fluency correlate so highly. With the present method, no such correlation was expected between novelty and fluency (quantity) measures, and no obvious pattern could be seen in the correlations of novelty with quantity in the present experiments; those correlations were strongly positive for the categories *Things to Take Camping* (r = .70) and *Birds* (r = .31), near zero for *Heavy Things* (r = −.04), and strongly negative for *Cold Things* (r = −.91), *Fattening Foods* (r = −.72), and *Clothing* (r = −.70). The present method for calculating response and participant novelty scores has also been used in other published studies of incubation effects (Kohn & Smith, 2010; Shah et al., 2003).

Two characteristics of the present research are particularly new and potentially interesting to researchers of creative cognition. The first is the methodological innovation of studying aspects of creative thinking using a category generation task, a task often used to study cognitive aspects of conceptual structures in both creative (e.g., Ward, 1994, 1995) and non-creative cognition (e.g., Barsalou, 1982, 1985). Studying divergent thinking as an activity analogous to category generation (at least for flexibly structured categories) could lead to insights into methods of enhancing creative production. For example, Angello (2014) showed that repeated attempts to retrieve non-existent category members inhibited access to dominant category members, but increased retrieval of low frequency (i.e., novel) exemplars in a category generation task. Although these exploratory findings are not ecologically valid in terms of real-world creativity, they help establish causal relationships among variables, and they may lead to methods that could enhance creativity in divergent production tasks. The second aspect of our results that is new and potentially useful concerns the way we obtained the benefits of an incubation interval without having participants take any time off from work. Simply alternating between two tasks, rather than working continuously on each task for the same amount of time, increased measures of creative divergent production. This suggests that the mechanism that enables incubation effects, re-structuring after abandoning one's initial approach (see Sio & Ormerod, 2009), does not necessarily require an idle period of time between bouts of work on a problem.

REFERENCES

Angello, G. (2014). *Failed retrieval attempts foster generation of novel responses*. Unpublished doctoral dissertation, Texas A&M University.

Angello, G., Storm, B.C., & Smith, S.M. (2014). Overcoming fixation with repeated memory suppression. *Memory*, *23*, 381–389; doi:10.1080/09658211.2014.889167.

Barsalou, L.W. (1982). Context-independent and context-dependent information in concepts. *Memory & Cognition*, *10*, 82–93.

Barsalou, L.W. (1985). Ideals, central tendency, and frequency of instantiation. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, *11*, 629–654.

Bowers, K.S., Regehr, G., Balthazard, C.G., & Parker, K. (1990). Intuition in the context of discovery. *Cognitive Psychology*, *22*, 72–110.

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

Christensen, B.T., & Schunn, C.D. (2005). Spontaneous access and analogical incubation effects. *Creativity Research Journal*, *17*(2), 207–220.

Dodds, R.A., Smith, S.M., & Ward, T.B. (2002). The use of environmental clues during incubation. *Creativity Research Journal*, *14*, 287–304.

Finke, R.A. (1996). Imagery, creativity and emergent structure. *Consciousness and Cognition*, *5*, 381–393.

Finke, R.A., Ward, T.B., & Smith, S.M. (1992). *Creative Cognition: Theory, Research, and Applications*. Cambridge, MA: MIT Press.

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

Gilhooly, K.J., Georgiou, G.J., Garrison, J., Reston, J.D., & Sirota, M. (2012). Don't wait to incubate: Immediate versus delayed incubation in divergent thinking. *Memory and Cognition*, *40*(6), 966–975; doi:10.3758/s13421-012-0199-z.

Glaser, D.S. (2008). *Foraging memory: Retrieving words from one and from two semantic categories*. Master's thesis, Rice University.

Goff, K., & Torrance, E.P. (2002). *Abbreviated Torrance test for adults*. Bensenville, IL: Scholastic Testing Services.

Guilford, J.P. (1979). Some incubated thoughts on incubation. *The Journal of Creative Behavior*, *13*, 1–8.

Jansson, D.G., & Smith, S.M. (1991). Design fixation. *Design Studies*, *12*(1), 3–11.

Kerne, A., & Smith, S.M. (2004). The information discovery framework. *Proceedings of the Association for Computing Machinery, Designing Interactive Systems (DIS)* (pp. 357–360). Cambridge, MA: Conference on Designing Interactive Systems: Processes, Practices, Methods, and Techniques.

Knoblich, G., Ohlsson, S., Haider, H., & Rhenius, D. (1999). Constraint relaxation and chunk decomposition in insight problem solving. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *25*, 1534–1555.

Kohn, N.W., & Smith, S.M. (2009). Partly versus completely out of your mind: Effects of incubation and distraction on resolving fixation. *Journal of Creative Behavior*, *43*(2), 102–118.

Kohn, N.W., & Smith, S.M. (2010). Collaborative fixation: Effects of others' ideas on brainstorming. *Applied Cognitive Psychology*, *24*, 1–22; doi:10.1002/acp.1699.

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

Metcalfe, J., & Weibe, D. (1987). Intuition in insight and noninsight problem solving. *Memory & Cognition*, *15*(3), 238–246.

Moss, J., Kotovsky, K., & Cagan, J. (2007). The influence of open goals on the acquisition of problem relevant information. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *33*, 876–891.

Norlander, T., & Gustafson, R. (1996). Effects of alcohol on scientific thought during the incubation phase of the creative process. *The Journal of Creative Behavior*, *30*, 231–248.

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

Osborn, A. (1957). *Applied imagination*. New York: Scribner.

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

Paulus, P.B., Nakui, T., Putman, V.L., & Brown, V.R. (2006). Effects of task instructions and brief breaks on brainstorming. *Group Dynamics: Theory, research, and practice*, *10*, 206–219.

Penney, C.G., Godsell, A., Scott, A., & Balsom, R. (2011). Problem variables that promote incubation effects. *The Journal of Creative Behavior*, *38*, 35–55; doi:10.1002/j.2162-6057.2004.tb01230.x.

Plucker, J.A., Qian, M., & Wang, S. (2011). Is originality in the eye of the beholder? Comparison of scoring techniques in the assessment of divergent thinking. *The Journal of Creative Behavior*, *45*, 1–22; doi:10.1002/j.2162-6057.2011.tb01081.x.

Raaijmakers, J.G., & Shiffrin, R.M. (1981). Search of associative memory. *Psychological Review*, *88*, 93–134; doi:10.1037/0033-295X.88.2.93.

Raaijmakers, J.G.W., & Shiffrin, R.M. (1992). Models for recall and recognition. *Annual Review of Psychology*, *43*, 205–234; doi:10.1146/annurev.ps.43.020192.001225.

Rundus, D. (1973). Negative effects of using list items as recall cues. *Journal of Verbal Learning & Verbal Behavior*, *12*, 43–50; doi:10.1016/S0022-5371(73)80059-3.

Seifert, C.M., Meyer, D.E., Davidson, N., Patalano, A.L., & Yaniv, I. (1994). Demystification of cognitive insight: Opportunistic assimilation and the prepared-mind perspective. In R.J. Sternberg & J.E. Davidson (Eds.), *The nature of insight* (pp. 65–124). Cambridge, MA: MIT Press.

Shah, J.J., Millsap, R.E., Woodward, J., & Smith, S.M. (2012). Applied tests of design skills – Part I: Divergent thinking. *American Society of Mechanical Engineering (ASME) Journal of Mechanical Design*, *134*(2), 134–143; doi:10.1115/1.4005594.

Shah, J.J., Smith, S.M., Vargas-Hernandez, N., Gerkens, R., & Wulan, M. (2003). *Empirical studies of design ideation: Alignment of design experiments with lab experiments*. Proceedings of the American Society for Mechanical Engineering (ASME) DTM.

Silvia, P.J. (2008). Another look at creativity and intelligence: Exploring higher-order models and probable confounds. *Personality and Individual Differences*, *44*, 1012–1021.

Sio, U.N., & Ormerod, T.C. (2009). Does incubation enhance problem solving? A meta-analytic review. *Psychological Bulletin*, *135*, 94–120.

Smith, S.M. (1995a). Fixation, incubation, and insight in memory, problem solving, and creativity. In S.M. Smith, T.B. Ward & R.A. Finke (Eds.), *The creative cognition approach* (pp. 135–155). Cambridge: MIT Press.

Smith, S.M. (1995b). Getting into and out of mental ruts: A theory of fixation, incubation, and insight. In R. Sternberg & J. Davidson (Eds.), *The nature of insight* (pp. 121–149). Cambridge, MA: MIT Press.

Smith, S.M., & Blankenship, S.E. (1989). Incubation effects. *Bulletin of the Psychonomic Society*, *27*, 311–314.

Smith, S.M., & Blankenship, S.E. (1991). Incubation and the persistence of fixation in problem solving. *American Journal of Psychology*, *104*, 61–87; doi:10.2307/1422851.

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

Smith, S.M., Ward, T.B., & Finke, R.A. (1995). *The creative cognition approach*. Cambridge: MIT Press.

Torrance, E.P. (1974). *Torrance tests of creative thinking*. Lexington, MA: Personnel Press.

Van Overschelde, J.P., Rawson, K.A., & Dunlosky, J. (2004). Category norms: An updated and expanded version of the Battig and Montague (1969) norms. *Journal of Memory and Language*, *50*, 289–335; doi:10.1016/j.jml.2003.10.003.

Vul, E., & Pashler, H. (2007). Incubation benefits only after people have been misdirected. *Memory & Cognition*, *35* (4), 701–710.

Ward, T.B. (1994). Structured imagination: The role of category structure in exemplar generation. *Cognitive Psychology*, *27*, 1–40; doi:10.1006/cogp.1994.1010.

Ward, T.B. (1995). What's old about new ideas? In S.M. Smith, T.B. Ward & R.A. Finke (Eds.), *The creative cognition approach* (pp. 157–178). Cambridge: MIT Press.

Ward, T.B., Smith, S.M., & Vaid, J. (1997). *Creative thought: An investigation of conceptual structures and processes*. Washington, DC: American Psychological Association Books.

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

Steven M. Smith, Texas A&M University

David R. Gerkens, California State University

Genna Angello, Michigan State University

Correspondence concerning this article should be addressed to Steven M. Smith, Department of Psychology, Texas A&M University, College Station, TX 77843. E-mail: stevesmith@tamu.edu