Running head: INCUBATION AND PROBLEM SOLVING

Does incubation enhance problem-solving? A meta-analytic review

Ut Na Sio

Thomas. C. Ormerod

Abstract

A meta-analytic review is reported of empirical studies that investigate incubation effects on problem-solving. Although some researchers report increased solution rates after an incubation period, a period of time in which a problem is set aside prior to further attempts to solve, others have failed to find effects. The analysis examined contributions to incubation effect sizes of moderators such as problem type, presence of solution-relevant or misleading cues, and lengths of preparation and incubation periods. A positive incubation effect was identified, divergent thinking tasks benefiting more than linguistic and visual insight tasks from incubation. Longer preparation periods gave a greater incubation effect, while filling an incubation period with high cognitive demand tasks gave a smaller incubation effect. Surprisingly, low cognitive-demand tasks yielded a stronger incubation effect than rest during an incubation period when solving linguistic insight problems. The existence of multiple types of incubation effect provides evidence for differential invocation of knowledge-based versus strategic solution processes across different classes of problem, and suggests that the conditions under which incubation can be used as a practical technique for enhancing problem-solving must be designed with care.

Keywords: Incubation, problem-solving, insight, meta-analysis

Does incubation enhance problem-solving? A meta-analytic review

Anecdotal reports of the intellectual discovery processes of individuals hailed as geniuses (e.g. Wallas, 1926; Woodworth & Schlosberg, 1954; Ghiselin, 1985) share a common theme: a flash of insight pops unexpectedly into the mind of the individual after they have put an unsolved problem aside for a period of time, having failed in initial attempts to solve it. This temporary shift away from an unsolved problem that allows a solution seemingly to emerge as if from no additional effort is termed an ‘incubation period’ (Wallas, 1926). Its importance in current thinking and practice is illustrated by a yield of 5510 articles that mention the term ‘Incubation’, along with one of ‘Creativity’, ‘Insight’ or ‘Problem’, from a recent search of Google Scholar™, the search restricted to the years 1997 to 2007 and the subject areas of Social Sciences, Arts and Humanities. An additional 1970 articles were yielded by including Business, Administration and Economics. Yet there are many conflicting accounts of incubation, some studies reporting strong effects (e.g., Smith & Blankenship, 1989), others failing to find any effect at all (Olton & Johnson, 1976). This paper aims to resolve the uncertainties surrounding the phenomenon by providing a statistical meta-analytic review of empirical studies of incubation. We suggest that it is only armed with the results of an integrative and quantitatively based review that progress in understanding the mechanisms that might underlie the phenomenon can be made.

One theoretical reason for studying incubation is because it is closely associated with insightful thinking. Indeed, Wallas (1926) proposed incubation as the second of four phases in problem-solving (the others being preparation, illumination and verification). Insight may be characterized as a sudden, unpredictable and non-verbalizable solution discovery (e.g., Metcalfe & Weibe, 1987). Some researchers see the apparently unconscious nature of solution discovery as evidence that the processes required to achieve insight in problem-solving are qualitatively

different from those used to tackle problems that do not require insight (e.g., Wertheimer, 1985; Jung-Beeman & Bowden, 2000). Incubation might serve a valuable role in arbitrating between theories of insight, in particular between special-process theories based on unconscious mechanisms of spreading activation (e.g., Knoblich, Ohlsson & Reinhus, 1999) and theories of insight as a normal problem-solving processes based on conscious mechanisms of search (e.g., MacGregor, Ormerod & Chronicle, 2001).

Understanding the role of incubation periods may also allow us to make use of them effectively to promote creativity in areas such as individual problem-solving, classroom learning, and work environments. Educational researchers have tried to introduce incubation periods in classroom activity, and positive incubation effects in fostering students' creativity have been reported (Lynch & Swink, 1967; Medd & Houtz, 2002; Rae, 1997; Webster, Campbell, & Jane, 2006). However, in the absence of a comprehensive theory or model that can explain how and why positive incubation effects might emerge and under what conditions they are best fostered, no general pedagogic recommendations can be made.

Several hypotheses have been proposed to account for the alleged positive effects of incubation periods on problem solving, and they can be divided into two main kinds; conscious-work and unconscious-work. The conscious-work hypothesis holds that incubation effects are due to issues such as reduction of mental fatigue (Posner, 1973) or additional covert problem solving during the incubation period (Browne & Cruse, 1988; Posner, 1973). Both sources implicate changes in consciously controlled problem-solving activities during an incubation period. In contrast, the unconscious-work hypothesis suggests that positive incubation effects are the result of gradual and unconscious problem-solving processes that occur during an incubation period (Bower, Regehr, Balthazard, & Parker, 1990; Simon, 1966; Smith, 1995;

Smith & Blankenship, 1991; Seifert, Meyer, Davidson, Patalano, & Yaniv, 1995; Yaniv & Meyer, 1987).

Three different unconscious processes have been proposed to account for incubation effects. The first involves eliciting new knowledge: over time, activation will spread towards previously-ignored but relevant memory items. Even if relevant items do not receive above-threshold activation, this process can still sensitize individuals to related concepts, and thus they will be more likely to make use of external cues to solve a problem. In addition, partially-activated concepts may combine with others to yield fortuitous insightful ideas (Bower et al., 1990; Smith, 1995; Smith & Blankenship, 1991; Yaniv & Meyer, 1987). The second hypothesis is selective forgetting: an incubation period will weaken the activation of inappropriate solution concepts that distract individuals during initial attempts, allowing a fresh view of the problem (Smith, 1995; Smith & Blankenship, 1991). The third hypothesis is problem restructuring, in which an individual's mental representation of a problem will be re-organized into a more appropriate and stable form after initial unsuccessful attempts. The individual is then more able to capitalise upon relevant external information or to re-arrange problem information in a manner that allows a solution to be found more readily (Seifert et al., 1995). Problem restructuring might emerge from either switching the strategy used to search for moves to attempt (e.g., MacGregor, Ormerod, Chronicle, 2001) or from relaxing self-imposed inappropriate constraints on the problem representation (Knoblich, Ohlsson, Haider, & Rhenius, 1999). Studies of meta-cognition indicate that strategy switching can be unconscious (Newton & Roberts, 2005; Reder & Schunn, 1996; Siegler & Stern, 1998) and that different strategies compete for activation during the strategy selection process (Siegler & Stern, 1998).

The conscious- and unconscious-work accounts generate different predictions concerning the effects of activities that individuals engage in during an incubation period. According to the conscious-work account, individuals benefit most from an unfilled incubation period, as this gives them an opportunity either to relax, reduce fatigue, or to continue working on the problem. In contrast, unconscious work accounts suggest that unconscious problem-solving processes occur when individuals shift their attention away from the problem to other mental activities. Thus, a certain level of involvement in other tasks during an incubation period should facilitate post-incubation problem-solving.

A number of experimental studies have examined the role of task type during an incubation period. The experimental paradigms of these incubation studies are fairly uniform: one group of participants is interrupted with an incubation period (having a break or performing other tasks) while solving a problem, whereas the other group works on the problem continuously. Performance differences between these two groups are then compared. The findings of the published studies do not give unconditional support to either the unconscious-work or the conscious-work accounts. Patrick (1986) found that participants who had a filled incubation period outperformed those who had an unfilled incubation period. However, Browne and Cruse (1988) reported the opposite pattern: participants who took a rest during an incubation period performed better than those who had to perform tasks during an incubation period. There are also studies that report the same level of performance by participants with filled and unfilled incubation periods (Olton & Johnson, 1976, Smith & Blankenship, 1989). However, these studies vary in terms of the length of incubation period, the target problems tackled, and the nature of the interpolated tasks during the incubation period.

Because of inconsistent findings concerning incubation, some researchers have doubted the existence of the effect, in particular rejecting the unconscious-work hypotheses (Brown & Cruse, 1988; Olton & Johnson, 1976; Perkins, 1995). However, one explanation for conflicting findings is that there are procedural moderators other than task type that influence the occurrence of problem solving processes during an incubation period, such as the length of the incubation period or the nature of the problem. The field lacks a comprehensive review to summarize and evaluate these studies. There have been two reviews to date of relations between different procedural variables and the incubation effect, and both are qualitative in nature. Olton's (1979) review of past incubation studies led him to question the existence of incubation effect, given that no experimental paradigm appeared to demonstrate incubation effect reliably. Yet, limited numbers of studies were available at that time: only 10 incubation studies were included in his review. A recent review by Dodds, Ward and Smith (in press) with more studies included suggested that several variables may interact to influence the effectiveness of an incubation period. However, the qualitative nature of their review led them to conclude that findings of past studies are too divergent and that more studies are needed to assess the impact of each variable and to identify the optimum settings for an incubation effect.

The wide variation in experimental parameters among studies makes it difficult to draw

cross-experiment conclusions from a qualitative review. To overcome these problems, a systematic meta-analytic review is needed. Meta-analytic review allows a quantitative evaluation of research domains that describes the typical strength of the effect or phenomenon, and also the relation of each moderator to the size of the effect by using statistical analysis methods (Rosenthal, 1995). The objectives of the current study were to carry out the first statistical meta-analysis of incubation studies to assess the effect size of the experimental incubation effect, and more importantly, the impact of potential moderators on the incubation effect size. However, in order to undertake the meta-analysis, we first need to identify the likely key moderators, which we achieve in the next section by reviewing the methods used in previous studies. A particular focus of this review is to identify moderators that might discriminate between conscious-work and unconscious-work hypotheses, and also between the mechanisms (reduction of fatigue, additional work; activation of new information, forgetting, restructuring) that might underlie each hypothesis.

The Interpolated Task Used During the Incubation Period

Various types of interpolated task have been used in past studies, and they can be divided into tasks of high or low cognitive demand. Examples of high cognitive demand tasks include mental rotation, counting backwards, and visual memory tests, whereas reading is commonly adopted as a low cognitive demand task. High demand tasks should fully occupy the individuals' mind, and prevent further conscious work on the unsolved problem. Some studies reporting that undertaking a high cognitive demand task during an incubation period is beneficial to the problem solving process (Kaplan, 1990; Patrick, 1986; Segal, 2004). Nonetheless, studies using low cognitive demand tasks that do not require individuals to focus their conscious attention on task performance report similar benefits (Beck, 1979; Silveira, 1972; Smith & Blankenship, 1989).

Length of the Incubation Period

Longer incubation periods may allow more additional problem-solving activity, or allow a greater degree of forgetting of misleading items or spreading of activation memory. Thus, problem solvers may show a larger performance improvement when they return to the problem after a long incubation period than after a short one. Some studies report evidence supporting this contention (Silveria, 1971, Smith & Blankenship, 1989, Beck, 1979, Fulgosi & Guilford, 1968). However, it is difficult to draw cross-experiment conclusions, since there is no standard operationalization of what constitutes “long” and “short” incubation periods. In Smith and Blankenship’s study (1989), for example, a 15-min incubation period was defined as a long incubation period, and they reported that participants receiving this length of incubation period performed better than those receiving a 5-min incubation period. However, in Beck’s (1979) study, a 20-min incubation period was considered to be short, and participants’ performance in this group did not differ from the control group. Kaplan (1989) suggested that, to judge whether the incubation period is short or long, the length of time that problem solvers spend on initial attempts to solve (named the “preparation period” by Wallas, 1926) should also be taken into account. Kaplan found that a larger incubation effect was observed after increasing the ratio of the length of preparation period to incubation period. Thus, in addition to including incubation and preparation periods as separate moderators in the meta-analysis reported below, a secondary analysis was also undertaken using the ratio of preparation to incubation time as an alternative moderator.

Length of the Preparation period

During the preparation period, problem solvers gather information to formulate a problem representation and make initial attempts to solve, which may lead to an impasse. Although a

problem may not be solved during the preparation period, this does not mean that the effort the problem solver spends on the problem is fruitless. Schank (1982, 1999) and VanLehn (1988) both suggest that failure in problem solving is important in the human learning process. Studies by Palatano & Seifert (1994) and Seifert, et al. (1995) have found evidence of a Zeigarnik effect in insight problem solving (Zeigarnik, 1927, 1938), where individuals remembered the problems on which they got “stuck” better than those solved immediately. Seifert et al. hypothesized that having a better memory for failed problems might help individuals return efficiently to the problem once relevant new information is encountered during an incubation period, thereby maximizing the chance of solving. Evidence concerning this prediction has been obtained in an empirical study carried by Silveira (1971), showing that problem solvers performed better with longer preparation and incubation periods.

Nature of the Problem

Various different types of problem have been used in incubation studies. Some problems, which we term ‘creative problems’ here, require individuals to produce multiple new ideas to meet a specific brief. For example, a verbal divergent-production task is the Consequences test (e.g., “What would be the result if everyone suddenly lost the ability to read and write”; Fulgosi & Guilford, 1967). Typically, there is no right or wrong answer to these kinds of problem and performance is assessed in terms of the numbers of solution ideas that are generated.

Other problems require individuals to discover a specific target solution that is known in advance by the experimenter. Problems of this kind studied in the literature on incubation are generally of a type described as ‘insight’ problems, in that they require the solver to reject initial solution ideas by achieving insight into an alternative strategy or knowledge domain. The insight

problems used in the incubation studies can be divided into visual problems, which typically require the solver to consider a visuo-spatial array of the problem (e.g., the nine-dot problem; Scheerer, 1963), and linguistic problems, which typically require the solver to consider linguistic information related to the problem. The Remote Associates Task (RAT; Mednick, 1962) is one of the most commonly used linguistic problems in incubation studies. In each RAT, three stimulus words are presented to individuals, who then have to think of a fourth word that can form association with each of the three words. For example, if the three stimulus words of a RAT are “electric”, “wheel”, and “high”, the forth word can be “chair”. Bowden & Jung-Beeman (2003) have developed a pool of remote associate problems, and collected the normative data regarding the resolution rate and response time for the problems. The classification of insight problems into visual- and linguistic-based is supported by research findings from Gilhooly and Murphy (2005) showing that solving visual and linguistic insight problems require different types of cognitive skills.

In the remainder of this paper, we refer to problem types as creative, visual, and linguistic. Descriptions of the types of problem used in incubation studies are illustrated in Appendix A. Problem type is likely to be an important determinant of incubation, since it seems likely that each type creates different task demands. For instance, the nine-dot problem appears to require the solver to restructure an initially faulty or incomplete problem representation in searching for a representation that allows solution, while the Consequences Task appears to require the activation of as wide a range of different concepts as possible. One question the meta-analysis allows us to address is whether an incubation period favours one type of problem more than another.

The Presence of Solution-Relevant Cues

Some unconscious processes proposed to explain incubation effects are purely internal and independent of the external environment, such as the inhibition of irrelevant memory (Smith, 1995; Smith & Blankenship, 1991) and the recombination of partially-activated concepts (Bower et al., 1990). Others stress interactions with the external environment, such as the proposal that spreading activation can partially activate previously-ignored relevant memory and therefore sensitize the problem solver to chance encounters with related stimuli (Seifert et al., 1995). A few studies have examined effects of the presence of cues during an incubation period (Browne & Cruse, 1988; Dodds, Smith, & Ward, 2002; Dorfman, 1990; Driestadt, 1969, Olton & Johnson, 1976). Most failed to find any positive effect of cues on the incubation effect. However, the failure reported in these studies may be due to other factors, such as the difficulty of the unsolved problems. In order to have a fair evaluation on the impact of this moderator, we first have to isolate the effect of other moderators on the incubation effect.

Misleading cues

Another factor that may influence the occurrence of incubation effects is the presence of misleading cues. Smith and Blankenship (1989) carried out a series of experiments to examine the effect of an incubation period on solving Remote Associates tasks (RATs), in which participants had to find a word that might accompany each of three presented words. Smith & Blankenship presented cues (shown here in italics) comprising misleading associates and the target word next to each of the three stimulus words. An example of a misleading RAT is: SHIP *ocean*, OUTER *space*, CRAWL *floor*. The target solution is “space”. Performance improvements after an incubation period were observed only when participants solved tasks containing misleading cues. They concluded that a problem solver who is fixated on misleading information benefits more from an incubation period. The misleading cues data provide critical support for forgetting-based explanations of incubation. The presence of misleading cues is therefore one of the potential moderators examined in this meta-analysis.

The Meta-analysis

The variables mentioned above were the potential moderators of incubation chosen for this

meta-analysis. Note that other potential moderators might have been included (e.g., number of trials, participant characteristics), but we focused on those we believe fundamental to discriminating between differing theoretical accounts of incubation.

The statistical meta-analysis that follows addressed two questions: first, is there reliable evidence for incubation; and second, what are the most influential moderators? To address the first question, the effect size of the incubation effect reported in each available study was computed. Given that the variability among effect sizes is likely to be greater than that resulting from subject-level sampling, a random effects model was adopted in this meta-analysis. A heterogeneity test was carried out to verify this assumption, and then the weighted mean under the assumption of random effects model was computed and assessed to determine if it was significantly larger than zero.

To address the second question, weighted least-squares linear regressions were carried out using the aforementioned moderators as predictor variables and the incubation effect size in each study weighted by the inverse of its variance as the criterion variable. The results of the regression report the independent contributions of each potential moderator to the incubation effect, while controlling for all other moderators. This approach allows us to summarize the past studies systematically even though they vary widely in numbers and type of experimental parameters. In addition, interactions between different moderators, such as the nature of the interpolation task during the incubation period and the nature of the problem, were examined.

Literature Search

Publications that contained studies relevant to a meta-analysis of incubation were collected through a search of the ERIC, PsycInfo, PsycArticles, and MEDLINE databases using the keyword *incubat**, intersected with one of *fixation*, *creativ**, *divergent**, *insight**, or *problem*. Then, references given in all the obtained articles were systematically searched for additional relevant publications. There is a concern that studies with statistically significant results are more likely to get published than those without significant results, and this may lead to a biased retrieval of studies. To ameliorate this to some extent, similar literature searches were carried out

in the ProQuest Digital Dissertations databases and using Google Scholar™ for retrieving PhD dissertations, unpublished papers, and conference papers concerning the incubation effect. In total, 37 relevant publications were identified and obtained. Studies meeting the following criteria were assimilated in the analysis.

1. The settings and difficulty of the problems were the same among all the experimental conditions;
2. The total length of time that participants could spend on solving the problem consciously was the same among all the conditions;
3. The study included a control (no-incubation) group, and participants in that group worked on the problem continuously;
4. Participants' problem solving performance in pre and post- incubation periods was measured;
5. The study reported information that allowed the computation of an effect size.

The 1st and 2nd selection criteria ensured that tasks were presented in an identical way among different conditions, and any between-condition performance differences could be attributed to differences in settings of the incubation period. The inclusion of criterion 3, a control condition (no break between the first and the second attempts at the problem), is essential to provide a baseline against which performance in incubation conditions can be compared. Only publications that assessed the problem solving performance in both first and second attempts were included in the analysis (criterion 4). Therefore, some studies (e.g. Sio & Rudowicz, 2007; Yaniv & Meyer, 1987) that did not assess post –incubation problem solving performance were excluded. The information required for computing effect sizes is discussed in the section “Estimation of effect sizes”. Eight publications were excluded because the experimental studies contained within them

failed to meet one or more of the mentioned criteria. The specific reasons for excluding the publications are described in Table 1, which also describes the settings of studies included in the meta-analysis. Of the remaining 29 publications, 20 were refereed journal articles, 8 were PhD dissertations, and 1 was a conference paper. The ratio of the refereed to other studies is 2.2:1, which is within the suggested range of between 128:1 and 1:1 for including unpublished studies in an effort to avoid publication bias (Thornton & Lee, 2000). Most publications included multiple experiments, thereby allowing a reasonable sample size of independent studies ($n = 117$) to be achieved.

[Insert Table 1 about here]

Coding procedure

Many of the experiments reported in the selected publications had two or more experimental conditions, such as incubation periods of different length or different types of task in the incubation period. For the sake of the meta-analysis, experiments with more than one incubation condition were broken down into independent studies with one incubation condition and one control condition. The same control group may be included in more than one independent study, and compared with more than one incubation condition. For example, in Goldman, Wolters & Winogard's (1992) experiment, there were control, short-incubation-period, and long-incubation-period conditions. The experiment was decomposed into two studies, one consisting of the control and short-incubation-period conditions, and the other consisting of the control and long-incubation-period conditions. To avoid inflating the degrees of freedom available, the number of participants in the control condition was split across studies entered into the analysis, a method advocated by Bar-Haim, Lamy, Pergamin, Bakermans-Kranenburg, & van Ijzendoorn (2007).

There were also studies having more than one control condition. In such cases, the control condition that had the most similar setting to the incubation condition was chosen. For example, in Hansberry's (1998) third experiment, participants had to solve a list of RATs under one of three conditions: two control and one incubation. In one control condition, the RATs were presented individually for 60 seconds. In the other control condition, as well as in the incubation

condition, each RAT was presented in two separate 30-second blocks. Data from the latter control condition were therefore used in computing the effect size, because this control condition and the incubation condition had the closest settings in terms of RAT presentation.

After separating the experiments into numbers of independent studies, a standard system was used to code each study. Background information on each independent study (author, publication year, and the number of participants in each condition), as well as potential moderator variables, were extracted. Table 2 presents the coding system used in this meta-analysis. Appendix B presents the information extracted from each independent study by using the coding system.

[Insert Table 2 about here]

Estimation of effect sizes

The effect size, Cohen's d , was computed for each study entered into the meta-analysis. Cohen's d in this meta-analysis comprised the difference in mean problem-solving performance scores between the control and incubation conditions divided by their pooled standard deviation (Hedges & Olkin, 1985). In some cases, effect sizes had to be calculated from t - and f -values, frequencies or p -values. If a p -less-than value was given instead of a exact p -value, the p -less-than value was treated as an exact value, and an estimate of Cohen's d was generated. For studies that did not include any of the above-mentioned information but only provided statements of non-significant differences between the control and the incubation groups, then Cohen's d was assumed to be zero. Among the studies that included multiple incubation conditions, some provided a statement of non-significant performance differences among the incubation conditions, and only reported the overall performance difference between the control and the incubation conditions. In such cases, all incubation conditions were assumed to generate the same magnitude of incubation effect sizes. Of the 117 effect sizes, 88 were extracted directly from the means and standard deviations, t -value, f -value, frequencies, or p -value; 8 were computed from a p less than value; and 21 were estimated from statements of significance.

In some incubation studies, problem solving performance was assessed along more than one dimension. For example, in the study carried out by Vul and Pashler (2007), participants' performance on RATs was measured in terms of the time spent on solving RATs and the number of correct solutions. In such cases, a single effect sized was computed by averaging the effect

size from each measure (cf. Durlak & Lipsey, 1991).

Following Hedge and Olkin's (1985) suggestion for removing bias caused by small sample studies, an unbiased effect size estimate was computed by multiplying the effect size of each single study by a factor $1 - 3/(4(\text{total } N - 2) - 1)$, where total N is the total number of participants of that study. Any unbiased effect size larger than 2 standard deviations from the group mean was considered an outlier, and was recoded to the value of the effect size found at 2 standard deviations, following a procedure for reducing the bias caused by extreme effect sizes reported by Lipsey & Wilson (2001).

Heterogeneity Analysis

In this analysis, we predicted that the variance in magnitude of the unbiased effect sizes among studies was not due simply to sampling error but instead to the difference in settings of each study (e.g., length of incubation period, nature of incubation task, presence of cues). Therefore, analyses of the effect sizes should be carried out under the assumption of random effects model. To confirm the assumption of heterogeneous distribution of effect sizes, a heterogeneity test was carried out before running any analyses on the effect sizes. The standard measure of heterogeneity is the Cochran's Q test. The Q statistic is the weighted¹ sum of squared differences between the unbiased effect size estimate of each independent study and the weighted average unbiased effect size estimate across studies. Q is distributed as a chi-square statistic with $k - 1$ degree of freedom, where k is the number of independent studies. If the Cochran's Q test for heterogeneity is statistically significant (Q is larger than the chi-square value with $k - 1$ degree of freedom), the assumption of the random effects model is supported.

Publication Bias

Prior to investigating the impact of potential moderators, a preliminary analysis was undertaken to assess if a publication bias existed in the selection of studies despite the inclusion of unpublished studies. A funnel plot of sample size against unbiased effect size estimates was created. In the absence of any publication bias, it is expected that the plot would be a funnel shape, such that the amount of scatter about the mean effect size decreases with increasing sample size. In addition to checking the presence of publication bias qualitatively, a weighted least-squares linear regression was carried out using the unbiased effect size estimates as the

dependent variable, and sample sized weighted by the inverse of variance in a random effects model, which is the sum of the between-studies variance² (random variance component) and within-study variance of the unbiased effect size. The regression slope (unstandardized coefficient of the predictive variable) would be expected to approach zero if there is no publication bias (Macaskill, Walter, & Irwing, 2001). The outcome of this analysis is reported below.

Regression Model Testing

Due to the wide variation in experiment settings among incubation studies, observed incubation effect size differences may reflect the combined impact of different moderators. Hence, weighted least-squares regression analyses were carried out to reveal the true impact of each moderator on incubation effects. The regression analyses were organized into two main sections. In the first section, the incubation studies were first classified into different groups, in terms of the types of problem used, the cognitive load of the incubation tasks, the presence of misleading cues, and the presence of relevant cue during an incubation period. Within each sub-group, the random variance component of the studies was computed. A larger-than-zero random variance component implies that the variability of effect sizes within these studies is not simply due to subject-level sampling error (Lipsey & Wilson, 2001). A weighted least-squares regression analysis was carried out as a follow-up analysis to model the effect sizes. The unbiased effect size estimate weighted by the inverse of variance was the outcome variable of the regression analysis. The predictor variables included Problem Type, Misleading Cues, Cues, Incubation Task, Preparation Period, and Incubation Period³. The categorical variables were represented with the appropriate number of dummy coded vectors. The categories “Creative Problem” (Problem Type), “Rest” (Incubation Task), “No misleading cues” (Misleading Cues) and the “No Cue” (Cues) were used as reference groups in the analysis, and their coefficients were restricted to zero.

The second section of the analyses investigated the general impact of the moderators on the incubation effect sizes. In this section, all the incubation studies were group together, and a weighted least-squares regression analysis was carried out to investigate the general impact of each moderator. Again, the weighted unbiased effect size was the outcome variable of the regression analysis. The predictor variables were Problem Type, Misleading Cues, Cues, Incubation Tasks, Preparation Period, and Incubation Period. Another weighted least-squares regression was carried out to examine the interaction between the categorical variables Problem Type and Incubation Task. The predictor variables were the appropriate number of dummy coded vectors and the multiplicative terms of these two categorical variables, as well as the variables Misleading Cues, Preparation Period, Incubation Period, and Cues. A more detailed description of the selection of the dummy coded vectors and the multiplicative terms of “Problem Type” and “Incubation Task” is presented in Appendix C.

Results

One hundred and seventeen studies were included in this meta-analysis. The total number of participants was 3606, and the median number of participants per study was 25. An unbiased effect size estimate was computed for each independent study. Among these studies, 85 of them report positive effect sizes. The unbiased effect size estimates range from $-.71$ to 4.07 , and the median was $.26$. The unweighted mean of the unbiased effect size estimate was $.41$, with a standard deviation of $.71$. The upper and the lower bound 95% confidence interval were $.54$ and $.28$. Unbiased effect sizes larger than 2 standard deviations from the mean were recoded to the value of the effect size found at 2 standard deviations. Table 3 gives the stem-and-leaf display showing the distribution of the unbiased effect sizes. The unweighted mean of the adjusted unbiased effect size estimate was $.36$, with standard deviation $.51$. The upper and the lower bound 95% confidence interval were $.26$ and $.45$. The confidence interval does not include zero, implying that the estimate of mean unbiased effect size is significantly larger than zero.

[Insert Table 3 about here]

The heterogeneity statistic, Cochran's Q , was 173.99, and was significantly larger than the Chi-Square critical value, $df = 113$, $p < .001$. This supports the use of random effects model. The variance of each unbiased effect size in random effects model was the sum of the between-studies variance and within-study variance of the unbiased effect size. The between-studies variance, also called the random variance component, among these incubation studies was .0834. The mean of the weighted unbiased effect size was .29, with .04 standard deviation³, and the 95% confidence interval was (.21, .39). The non-zero confidence interval implies that the weighted mean is significantly larger than zero. This answers our first question, showing the existence of a positive incubation effect.

Figure 1 presents the funnel plot of sample size against estimated unbiased effect size of each study in the meta-analysis. A weighted least-squares regression using unbiased effect sizes weighted by the inverse of the variance as the dependent variable, and sample size as the predictive variable, was carried out. The regression coefficient of the predictive variable was not significantly different from zero, standardized $\beta = -.08$, $p = .41$, suggesting the absence of publication bias. Thus no correction has been made for publication bias.

[Insert Figure 1 about here]

Table 4 presents the weighted mean, standard deviation, 95% confidence interval, and random variance component in each sub-group of each categorical moderator.

[Insert Table 4 about here]

Six of the sub-groups (linguistic problems, creative problems, absence of misleading cues, absence of relevant cues, high cognitive load task, unoccupied incubation period) had larger-than-zero random variance components. New weightings, under the assumption of a random effects model, were computed for each of the sub-groups. Weighted least-squares regression analyses were carried out to find the moderators that accounted for the effect size variability among these sub-groups. Small numbers of studies using creative tasks and studies having unoccupied incubation periods preclude the possibility of regression analyses with these

moderators. An effect of applying a weighting to this regression analysis is to under-estimate the original standard error of each unstandardized coefficient. Thus, an adjusted standard error was computed by dividing the original standard error by the square root of the mean square residual, a procedure suggested by Lipsey & Wilson (2001). The corrected standard error was used in the significance test (z-test) of each unstandardized coefficient.

Tables 5 to 8 present a summary of the regression analysis results of each sub-group.

[Insert Table 5 about here]

[Insert Table 6 about here]

[Insert Table 7 about here]

[Insert Table 8 about here]

With the sub-group of studies using linguistic problems, low cognitive load tasks generated larger incubation effects than rest alone, $\beta = .54^*$, $p < .05$. Also, there was an interaction between Problem Type and Incubation Task with this sub-group, such that that low cognitive load tasks facilitated the incubation effect only when solving linguistic problems.

With the “absence of misleading cue”, and “absence of relevant cue” sub-groups, regression analyses reveal that, in the absence of these cues, individuals solving visual problems had smaller incubation effects than those solving creative and linguistic problems. There was also a positive impact of longer Preparation Periods on the incubation effect sizes. The cognitive load of the incubation tasks did not have any impact on the magnitude of the incubation effects in these two sub-groups. Note, however, that the presence or absence of these effects with these specific sub-groups does not necessarily imply the converse effects found with other sub-groups, hence a shift to analysis of general impacts of each moderator.

In the second stage of the analysis, a weighted least-squares regression analysis was carried

out to look at the general impact of each moderator on the incubation effect sizes. A summary of the regression analysis results is presented in Table 9. The negative coefficients associated with “visual problem” and “linguistic problem” indicate that individuals solving these two types of insight problem showed a smaller incubation effect than individuals solving creative problems. A z-test was carried out to compare the coefficients of “visual problem” and “linguistic problem”. The result was not statistically significant, $z\text{-score} = -1.25, p > .05$, suggesting the magnitude of the incubation effect for visual and linguistic insight problems was comparable.

[Insert Table 9 about here]

The length of preparation period was found to have a significant impact on the magnitude of the incubation effect, $\beta = .03, p < .05$. Three bivariate correlations were carried out to check for positive relationships between the length of preparation period and the magnitude of the weighted incubation effect when solving the three types of problem. There was a statistically significant positive correlation between the weighted incubation effect size and the length of preparation period with visual problems, $r(35) = .40^*, p = .02$, and creative problems, $r(14) = .60^*, p = .03$, but not with linguistic problems, $r(65) = -.04, p = .75$.

Another weighted least-squares regression analysis was carried out to examine the interaction between Problem Type and Incubation Task, using Misleading Cues, Preparation Period, Incubation Period, Cues, and an appropriate number of dummy vectors and the multiplicative terms of the variables Problem Type and Incubation Task as predictor variables. Table 10 presents the results of this regression analysis. To examine the interaction effects, the coefficient differences between their multiplicative terms were examined to see if they were significantly larger than zero by using z-tests. The details of equations for computing the coefficient difference can be found in Appendix C. Table 11 presents the coefficient differences

between multiplicative terms.

[Insert Table 10 about here]

[Insert Table 11 about here]

With creative problems, undertaking high cognitive load tasks was associated with smaller incubation effects than with low cognitive load tasks or rest during the incubation period, $C_j * H_j - C_j * R_{j..} = -.91, p < .05$, and $C_j * H_{j..} - C_j * L_{j..} = -.79, p = .08$. When solving linguistic and visual problems, no differences were found for incubation periods filled with low or high cognitive load tasks or rest. However, this regression model has a problem in exploring the interaction between Problem Type and Incubation Task. Among studies examining the role of an incubation period with creative problems, 12 out of 14 studies employed high cognitive load tasks and the remaining 2 studies employed rest during the incubation period. This unbalanced distribution may cause bias when examining the interaction between Incubation task and Problem Type. Hence, another regression analysis was carried out that excluded studies using creative problems. This third regression model included the variables Misleading Cues, Preparation Period, Cues, the dummy variables of Problem Type (excluding Creative problem) and Incubation Task, and their multiplicative variables. Appendix D presents a detailed description of the selection of the dummy coded vectors and the multiplicative terms. The results of this analysis is presented in Table 12, and the coefficient differences between the multiplicative terms are presented in Table 13

[Insert Table 12 about here]

[Insert Table 13 about here]

The regression results indicated an interaction between Problem Type and Incubation Task. When solving linguistic problems, a low cognitive load task generated significantly larger

incubation effects than rest, $L_{ij} * L_j - L_{ij} * R_j = .45, p = .05$. The difference between low and high cognitive loads was in the same direction but did not reach significance, and there was no significant difference between the rest condition and high cognitive load condition. When solving visual problems, the effect sizes among the three incubation conditions were comparable. This pattern of findings is consistent with the previous regression analysis results. In addition, the exclusion of creative problem studies makes the positive impact of a low cognitive load incubation period on linguistic problems more significant.

The model using the variables “Preparation Period” and “Incubation Period” was found to be not significant, and the variable “Preparation Period” was not significant in the analysis. This may be due to the decrease in the numbers of studies included in the current regression model. Moreover, as mentioned above, a positive association between the length of preparation period and incubation effect size was found when solving visual and creative problems. Thus, the exclusion of creative problem studies appears to decrease the significance of the variable “Preparation Period”, and the significance level of the regression model using the “Preparation Period” as one of the predictive variables.

Discussion

The meta-analysis supports the existence of incubation effects, and also identifies some potential moderators, including the problem type, length of preparation period, and the incubation task. Individuals solving creative problems were more likely to benefit from an incubation period than individuals solving linguistic and visual problems. Longer preparation periods gave rise to larger incubation effects. When solving linguistic problems, a low cognitive load task gave the strongest incubation effects.

We suggest that the positive incubation effects found with creative problems are a direct reflection of their multi-solution nature. When solving a creative problem, individuals benefit from performing a wide search of their knowledge to identify as many relevant connections as possible with the presented stimuli. Each time an individual re-approaches the problem, performance is improved by extending search to previously unexplored areas of their knowledge network. Incubation appears to facilitate the widening of search of a knowledge network in this fashion.

Linguistic and visual problems typically have only one possible solution. In order to solve them, individuals have to explore their memory or environment to look for specific relevant knowledge or to adapt a specific strategy. Widening search to new items of knowledge may not be facilitative if the solution to a problem lies within already activated knowledge that is currently represented inappropriately. Under this account, incubation supports knowledge activation, but it does not support restructuring.

Another finding of the meta-analysis was the beneficial effect of an incubation period filled with low demand tasks on solving linguistic problems. A positive effect of a filled incubation period on problem-solving compared with rest during the incubation period undermines the conscious-work hypothesis that incubation effects are due to the mental fatigue reduction (Posner, 1973). There remains a possibility, of course, that a sufficiently light load might allow additional covert problem solving compared with a heavier task load (Browne & Cruse, 1988; Posner, 1973), but this does not explain why a light load should be more facilitative than rest alone.

The positive effect of a light cognitive load may indicate competition between controlled and automatic processes in solving linguistic problems. It has been suggested with remote

associates task performance that only strong (and in this context incorrect) associates are accessed when individuals focus their attention on seeking solutions, whereas remote associates are more likely to be accessed when an individual's cognitive resources are allocated in a diffuse manner (Ansburg & Hill, 2001; Finke, Ward, & Smith, 1992; Martindale, 1995). During an incubation period, low-demand tasks may occupy part of the problem solvers' attention, preventing the focused concentration that yields strong associates. Resting during an incubation period may allow individuals to continue consciously working on the problem, while performing high-demand tasks may shift attention entirely to that interpolated task, leading to a narrow rather than diffused attentional focus. The impact of performing high-demand tasks is analogous to the 'verbal overshadowing' effect reported by Schooler and colleagues (Schooler, Ohlsson, & Brookes, 1993) in which the act of verbalizing can impair performance by focusing individuals' attention inappropriately on verbalisable components of a task. The suggested role of the light-load incubation tasks receives indirect support from recent findings which show that that visual search can be more efficient when performed concurrently with an unrelated task than when performed alone (Smilek, Enns, Eastwood, & Merikle, 2006). They suggest that the dual-task condition prevents a narrow attentional focus in searching stimuli.

With visual problems, the magnitude of the incubation effects was independent on the setting of an incubation period (filled or unfilled). Differences between visual and linguistic insight problems may arise through a greater reliance on strategic search rather than knowledge activation in the former than the latter. MacGregor, Ormerod, & Chronicle (2001) proposed that in solving the nine-dot problem, individuals select and execute moves that maximally reduce the distance between current and goal states, essentially drawing lines that connect as many dots as possible. While there remain moves available that satisfy a criterion of satisfactory progress (in

this case, the ratio of dots cancelled to lines available), individuals will persevere with an initial representation of the problem that, in the case of the nine-dot problem, does not include consideration of space outside the dot array. According to MacGregor et al's account, individuals must experience a failure to find moves that meet a criterion of satisfactory progress before they change the initial representation of the problem, thereby including space outside the dot array in their attempts. An incubation period would be helpful only if the problem solvers became aware of the necessity of a strategy shift, but according to MacGregor et al they are unlikely to do so unless they encounter criterion failure as a result of reaching impasse. Siefert, Mayer, Davidson, Palatino, & Yaniv (1995) offer an alternative account that also points to the criticality of experiencing failure and impasse for eventual success in insight problem-solving.

If the hypothesis that visual problems require impasse for a strategy switch to occur is correct, a long preparation period (i.e., pre-incubation problem-solving) should be more likely to yield benefits from subsequent incubation with visual problems because it allows individuals to reach impasse prior to incubation. The results of the regression analysis and the follow-up bivariate correlations are consistent with this prediction, showing a statistically significant positive correlation between the incubation effect size and the length of preparation period with visual problems.

A positive correlation between length of preparation period and incubation effect size was also found with creative problems. A long preparation period may allow individuals to exhaust search in one domain, making it more likely for them to explore a new domain in the second phase of solving. A positive correlation was not found when solving linguistic problems, though this may simply reflect the small variability in length of preparation period among studies using linguistic problems (the preparation period of 82% of these studies were ranged from 0.5 to 1

minute).

The meta-analysis reveals that embedding misleading cues to the problems was not a significant predictor overall. This result is in contrast with previous reports (e.g., Smith et al), which suggest that incubation effects arise through forgetting of inappropriate information. The true effect of misleading cues is underestimated in our regression analyses, as we only examined the overall effect of misleading cues on problem solving in general. Twenty nine independent studies included in this meta-analysis examined the impact of misleading cues; 25 of them were examining the impact on linguistic problems, the rest on visual problems. The weighted mean of the effect size estimates of these studies for each problem type by the presence of misleading cues were: Linguistic with Misleading Cue: $M = .36$, $SD = .09$; Linguistic without Misleading Cue: $M = .17$, $SD = .06$; Visual without Misleading Cue: $M = .26$, $SD = .08$; Visual with Misleading Cue: $M = .18$, $SD = .39$. For studies employing linguistic problems, the presence of misleading cues induced a larger incubation effect, but an opposite pattern of results was found in studies employing visual problems, suggesting that the impact of misleading cues may be modality-specific. In order to test whether the presence of misleading cues affects incubation with linguistic problems alone, a weighted one-way analysis of variance was run to compare the incubation effect sizes of studies using linguistic problems that included misleading cues (25 studies) against studies using linguistic problems that did not include misleading cues (40 studies), $F(1,65) = 3.04$, $MSe = 3.00$, $p = .08$. Thus, where problem materials involve linguistic stimuli, getting rid of misleading concepts may be the key to task solution, in contrast with visual problems in which the key may be to restructure the knowledge that is currently active. The effect of misleading cues also offers some support for the selective-forgetting hypothesis, but this

effect may be task-specific.

In contrast with previous reports (e.g., Dominowski, 1972; Dreitadt, 1969; Mednick, Mednick, & Mednick, 1964), the moderator Presence of Cues was not found to be a significant predictor of incubation effects. It has been hypothesized that, during an incubation period, unconscious processes such as spreading activation sensitize individuals to solution concepts and make them more likely to utilise externally presented cues. To examine this hypothesis, researchers have presented the answers of unsolved problems during an incubation period, and compared post-incubation performance with participants not receiving any cues during the incubation period (e.g., Dominowski & Jenrick, 1972; Dodds, Smith, & Ward, 2002). Findings of these studies are equivocal, but it does appear that problem solvers do not always make use of solution relevant cues, even when the cue includes the solution itself. For example, Chronicle, Ormerod, & MacGregor (2001) found that presenting the nine-dot problem with a shaded background in the shape of the solution did not lead to significant levels of facilitation, even when the relevance of the shading was drawn to participants' attention. However, because of the small number of studies that present solution relevant cues (three with linguistic, seven with visual problems, none with creative problems) and the wide variation in experimental parameters among these studies, it is impossible to carry out further statistical analysis.

In summary, the meta-analysis results support the existence of incubation effects, though there appears to be a range of effects specific to particular tasks and performance conditions. When attempting creative problems that require a wide search of knowledge, individuals benefit from an incubation period. Problems that involve reaching some kind of insight into a unique solution do not always benefit from incubation under all conditions. In the case of linguistic problems, such as the RAT, there is a modest incubation effect but only where the incubation

period is filled with a low cognitive demand task. One possible explanation is that performing low cognitive load incubation tasks allows the occurrence of some unconscious problem solving processes, such as spreading-activation and selective forgetting. In the case of visual problems, incubation effects arise only where there has been a sufficiently long preparation period prior to incubation for the problem-solver to have entered a state of impasse. Only under these conditions can an incubation period contribute to the strategic shift needed to restructure a problem representation. Thus, the theoretical positions of spreading activation, selective forgetting and restructuring each receive support. However, evidence for each appears to be specific to particular problem types.

Spreading activation and strategic search are basic mechanisms underlying different types of general cognitive process. For instance, the ACT-R computational framework (Anderson, 1994) utilizes both these mechanisms. Different researchers have successfully adapted this model to simulate a wide range of non-insight problem solving processes. Examining the occurrence of incubation effects in terms of fundamental cognitive processes would offer important findings for developing a computational model of insight problem-solving.

While the conscious work hypothesis receives little support here, the meta-analysis leaves open the possibility that unconscious processes may reflect forgetting, activation of new knowledge, or restructuring. Further experimental studies might focus on comparing the occurrence of these different unconscious processes during an incubation period in different experimental settings. One methodology that might allow such comparisons was employed by Sio and Rudowicz (2007), who examined the occurrence of spreading-activation by measuring individuals' sensitivity to answers of the unsolved RATs before and after a filled and unfilled incubation periods. They found that the spreading-activation process occurred only in the filled

incubation period condition and in a fixated mind, though this study did not measure post-incubation period performance.

Given the outcome that both spreading activation and restructuring might arise from an incubation period, depending on the task, then both the current views of insight problem-solving may take some comfort from the results. The evidence from the linguistic problems is consistent with release from inappropriate constraints (Knoblich, et al., 1999), while evidence from the visual problems is consistent with restructuring that results from a strategic shift following impasse (MacGregor, et al., 2001). Given that these theoretical alternatives have to date been explored only with different task sets, it seems quite possible that a complete account of insight might need both theoretical components. Such a view may well be consistent with the account of insight offered by Kershaw and Ohlsson (2004), who propose a multiple-source account of the difficulties individuals encounter in insight problem-solving.

The finding of a positive impact of an incubation period on solving creative thinking problems supports the contention that incubation periods help the elicitation of new ideas. Incubation is a concept central to many methodologies for encouraging creative decision-making, especially among management science and business communities (e.g., Rickards, 1991), and this result may be taken as supporting the inclusion of an incubation phase in such methodologies.

It should be noted that, despite efforts to include studies from sources other than peer-reviewed journals, the meta-analysis may be influenced by a bias in favour of reporting significant effects at the expense of null effects. Thus, incubation effects may be to some extent overstated in this meta-analysis, a problem common to all meta-analyses. Nonetheless, the reasonably large effect sizes found with creative problems indicates that, with this class of problem at any rate, incubation is a potentially valuable mechanism for fostering creative

thought.

Clearly, the empirical data on incubation are not straightforward. As a consequence, the traditional narrative review approach is not amenable to drawing strong cross-study conclusions. It is perhaps disappointing that relatively few published studies met the necessary criteria for inclusion in the meta-analysis, since a failure either to measure post-incubation performance or to provide effect size information limits the conclusions that can be drawn from them. Nonetheless, sufficient studies remain for the meta-analysis to be undertaken and to reveal some intriguing results.

One remaining problem is the relatively narrow range of problem types that have been explored. For instance, the majority of studies that explore incubation effects with linguistic problems, which is the majority of studies overall, use the RAT. It is unclear that whether the RAT can be considered an insight problem or a linguistic completion task, suggesting it may not be representative of all linguistic problem-solving tasks. Bowden and Jung-Beeman (2003) have found that participants sometimes claimed that they solved RATs with insight, while sometimes reported that they solved them without insight. Further studies should aim to explore task-specific experimental settings for maximizing the incubation effect with a wider range of tasks. A further research issue of value might also be to explore individual differences in incubation effects. For example, if the role of incubation is to encourage diffused attention, then individuals who show a propensity towards allocating attention broadly (e.g., as measured via field dependence) may benefit differentially from an incubation period. Also, studies have revealed that strategy switching is related to working memory capacity (Geary, Hoard, Byrd-Craven, & Desoto, 2004). Thus, memory capacity may also interact with incubation effects in solving visual problems.

One the whole, the results of this meta-analysis support the existence of multiple types of problem-specific incubation effect. We suggest that the concept of incubation can only be understood through a close examination of the problems to which it is applied and the conditions under which it is elicited.

References

- Anderson, R. J. (1994). Representations and requirements: The value of ethnography in system design. *Human-Computer Interaction*, 9, 2, 151-182.
- Ansburg, P. I., & Hill, K. (2003) Creative and analytic thinkers differ in their use of attentional resources. *Personality and Individual Differences*. 34, 1141-1152.
- Bar-Haim, Y., Lamy, D., Pergamin, L., Bakermans-Kranenburg, M. J., & van Ijzendoorn M. J. (2007). Threat-related attentional bias in anxious and nonanxious individuals: A Meta-analytic study. *Psychological Bulletin*, 133, 1-24.
- Beck, J. (1979). *The effect of variations in incubation activity and incubation time on creative response production*. Unpublished doctoral dissertation, New York University.
- Bennett, S. M. (1975). *Effect of incubation on the associative process of creativity*. Unpublished doctoral dissertation, University of Notre Dame.
- Bowden, E.M., & Jung-Beeman, M. (2003). Aha! Insight experience correlates with solution activation in the right hemisphere. *Psychonomic Bulletin & Review*, 10, 730-737.
- Bowden, E. M., & June-Beeman, M. (2003). Normative data for 144 compound remote associate problems. *Behavior Research Methods, Instruments, & Computers*, 35(5), 634-639.
- 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(3), 177-185.
- Chronicle, E. P., Ormerod, T. C., & MacGregor, J. N. (2001). When insight just won't come: The failure of visual cues in the nine-dot problem. *Quarterly Journal of Experimental Psychology*, 54A, 903-919.

- Cooper, H., & Hedges, L. V. (1994). *The Handbook of Research Synthesis*. New York: Russell Sage Foundation.
- Dodds, R. A., Smith, S. M., & Ward, T. B. (2002). The use of environmental clues during incubation. *Creativity Research Journal*, 14(3/4), 287-305
- Dodds, R. A., Ward, T. B., & Smith, S. M. (in press). Incubation in problem solving and creativity. In M. A. Runco (Ed.), *Creativity Research Handbook, Vol. 3*. Cresskill, NJ: Hampton Press.
- Dorfman, J., Shames, V. A., & Kihlstrom, J. F. (1996). Intuition, incubation, and insight: implicit cognition in problem solving. In G. Underwood (Ed.), *Implicit Cognition* (pp. 257–287). Oxford: Oxford University Press.
- Dreistadt, R. (1969). The use of analogies and incubation in obtaining insights in creative problem solving. *The Journal of Psychology*, 71, 159-175.
- Dominowski, R. L., & Jenrick, R. (1972). Effect of hints and interpolated activity on solution of an insight problem. *Psychonomic Science*, 26(6), 335-338.
- Dreistadt, R. (1969). The use of analogies and incubation in obtaining insights in creative problem solving. *The Journal of Psychology*, 71, 159-175.
- Durlak, J.A., & Lipsey, M.W. A practitioner's guide to meta-analysis. *American Journal of Community Psychology*, 19, 291-332.
- Finke, R.A., Ward, T.B., & Smith, S.M. (1992). *Creative Cognition: Theory, Research, and Applications*, Cambridge, MA: MIT Press.
- Fulgosi, A., & Guilford, J. P. (1968). Short-term incubation in divergent production. *American Journal of Psychology*, 81, 241-246.
- Geary, D. C., Hoard, M. K., Byrd-Craven, J., & DeSoto, M. C. (2004) Strategy choices in simple

- and complex addition: Contributions of working memory and counting knowledge for children with mathematical disability. *Journal of Experimental Child Psychology*, 88, 121–151.
- Ghislin, B. (1985). *The Creative Process. A Symposium*. Berkeley: University of California Press.
- Gilhooly, K. J., & Murphy, P. (2005). Differentiating insight from non-insight problem. *Thinking and Reasoning*, 11, 279-302
- Goldman, W., Wolters, N., & Winograd, E. (1992). A demonstration of incubation in anagram problem solving. *Bulletin of the Psychonomic Society*, 30, 36-38.
- Hedges L. V., & Olkin, I. (1985). *Statistical Methods for Meta-Analysis*. Orlando: Academic Press.
- Jung-Beeman, M., & Bowden, E. (2000). The right hemisphere maintains solution-related activation for yet-to-be solved insight problems . *Memory & Cognition*, 28, 1231-1241.
- Kaplan, C. (1990). *Hatching a theory of incubation: Does putting a problem aside really help? If so, why?* Unpublished doctoral dissertation, Carnegie-Mellon University.
- Kershaw, T.C., & Ohlsson, S. (2004). Multiple causes of difficulty in insight: The case of the nine-dot problem. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, 30(1), 3-13.
- 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.
- Lipsey, M. W., & Wilson, D. B. (2001). *Practical Meta-analysis*. Thousand Oaks: Sage.
- Lynch, M. D., & Swink, E. (1967). Some effects of priming, incubation and creative aptitude on

- journalism performance. *Journal of Communication*, 17(4), 372-382.
- Macaskill, P., Walter, S. D., & Irwing, L. (2001). A comparison of methods to detect publication bias in meta-analysis. *Statistics in Medicine*, 20, 641-654.
- MacGregor, J. N., Ormerod, T. C., & Chronicle, E. P. (2001). Information-processing and insight: A process model of performance on the nine-dot and related problems. *Journal of Experimental Psychology: Learning, Memory and Cognition*, 27, 176-201.
- Martindale, C. 1995. Creativity and connectionism. In S.M. Smith, T. B. Ward, & R. A. Finke (Eds.) *The Creative Cognition Approach*, Cambridge (pp. 249-268), MA: MIT Press.
- Medd, E., & Houtz, J, C. (2002). *Educational Research Quarterly*, 26(2), 13-16.
- Metcalfe, J., & Weibe, D. (1987). Intuition in insight and non-insight problem-solving. *Memory and Cognition*, 15, 238-246.
- Mednick, M., Mednick, S., & Mednick, E. (1964). Incubation of creative performance and specific associative priming. *Journal of Abnormal and Social Psychology*, 69, 84-88.
- Newton, E.J., & Roberts, M.J. (2005). The window of opportunity: A model for strategy discovery. In M.J. Roberts and E.J. Newton (Eds.), *Methods of Thought: Individual Differences in Reasoning Strategies* (pp. 129-158). Hove, United Kingdom: Psychology Press.
- Olton, R. M. (1979). Experimental studies of incubation: Searching for the elusive. *Journal of Creative Behavior*, 13, 9-22.
- Olton, R. M., & Johnson D. M.(1976). Mechanism of incubation in creative problem solving. *American Journal of Psychology*, 89(4), 617-630.
- Patalano, A. L., & Seifert. C. M. (1994). Memory for impasses during problem solving. *Memory and Cognition*, 22(2), 234-242.

- Patrick, A. S. (1986). The role of ability in creative "incubation". *Personal and Individual Differences*, 7(2), 169-174.
- Perkins, D. N. (1995). Insight in minds and genes. In R. J. Sternberg & J. E. Davidson (Eds.), *The nature of insight*. (pp. 495-533). Cambridge, MA: The MIT Press.
- Posner, M.I. (1973). *Cognition: An Introduction*. Glenview, IL: Scott, Foresman.
- Rae, C. M. (1997). The creative power of doing nothing. *Writer*, 110(7), 13-15.
- Reder, L. M., & Schunn, C. D. (1996). Metacognition does not imply awareness: Strategy choice is governed by implicit learning and memory. In L. M. Reder (Ed.), *Implicit memory and metacognition* (pp. 45–78). Hillsdale, NJ: Erlbaum.
- Rickards, T. (1991). Innovation and creativity: woods, trees and pathways. *R & D Management*, 21(2), 97-108.
- Rosenthal, R. (1984). *Meta-analytic for Social Research*. Beverly Hills: Sage Publication.
- Rosenthal, R. (1995). Writing meta-analytic reviews. *Psychological Bulletin*, 118, 183-192.
- Schooler, J. W., Ohlsson, S., & Brooks, K. (1993). Thoughts beyond words: When language overshadows insight. *Journal of Experimental Psychology: General*, 122, 166–183.
- Seabrook, R., & Dienes, Z. (2003). Incubation in problem solving as a context effect. *Proceedings of the 25th meeting of the Cognitive Science Society, Boston, July 31-Aug 2, 2003*. Mahwah, NJ: Lawrence Erlbaum Associates.
- Schank, R. C. (1982). *Dynamic Memory: A Theory of Reminding and Learning in Computers and People*. Cambridge: Cambridge University Press.
- Schank, R. C. (1999). *Dynamic Memory Revisited*. Cambridge: Cambridge University Press.
- Seifert, M. C., Meyer, D. E., Davidson, N., Patalano, A. L., & Yaniv, I. (1995). Demystification of cognitive insight: Opportunistic assimilation and the prepared-mind perspective. In R.

- J. Sternberg & J. E. Davidson (Eds.). *The Nature of Insight* (pp. 65-124). Cambridge, MA: MIT Press.
- Siegler, R. S., & Stern, E. (1998). Conscious and unconscious strategy discoveries: A microgenetic analysis. *Journal of Experimental Psychology: General*, 127, 377-397.
- Silveira, J. (1971). Incubation: The effect of interruption timing and length on problem solution and quality of problem processing. Unpublished doctoral dissertation, University of Oregon.
- Sio, U. N. & Rudowicz, E. (2007). The Role of an incubation period in creative problem solving. *Creativity Research Journal*, 19(2-3), 307-318.
- Smilek, D., Enns, J. T., Eastwood, J. D., & Merikle, P. A. (2006). Relax! Cognitive style influences visual search. *Visual Cognition*, 14, 543-564.
- Simon, H. A. (1966). Scientific discovery and the psychology of problem solving. In R. G. Colodny (Ed.), *Mind and Cosmos: Essays in Contemporary Science and Philosophy* (pp. 22-40). University of Pittsburgh Press.
- Smith, S. M. (1995). Fixation, incubation, and insight in memory and creative thinking. In S. M. Smith, T. B. Ward, & R. A. Finke (Eds.). *The Creative Cognition Approach* (pp. 135-146). Cambridge, MA: MIT Press.
- Smith, S. M., & Blankenship, S. E. (1989). Incubation effects. *Bulletin of the Psychonomic Society*, 27(4), 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.
- Segal, E. (2004). Incubation in insight problem solving. *Creativity Research Journal*, 16(1), 141-149.

- Snyder, A., Mitchell, D. J., Ellwood, S., Yates, A., & Pallier, G. (2004). *Nonconscious idea generation. Psychological Reports (94)*, 1325-1330
- Thornton, A., & Lee, P. (2000). Publication bias in meta-analysis – its causes and consequences. *Journal of Clinical Epidemiology*, 53(2), 207-216.
- VanLehn, K. (1988). Towards a theory of impasse-driven learning. In H. Mandl & A. Lesgold (Eds), *Learning: Issues for Intelligence Tutor System* (pp. 19-41). New York: Springer-Verlag.
- Vul, E., & Pashler, H. (2007).Incubation is helpful only when people are misled. *Memory, and Cognition*, 35(4), 701-710.
- Wallas, G. (1926). *The Art of Thought*. London: Jonathan Cape.
- Webster, A., Campbell, C., & Jane, B. (2006). *Enhancing the creative process for learning in primary technology education*. *International Journal of Technology & Design Education*, 16(3), 221-235.
- Wertheimer, M. (1985). A Gestalt perspective on computer simulations of cognitive processes. *Computers in Human Behavior*, 1, 19–33.
- Woodworth, R. S., & Schlosberg, H. (1954). *Experimental Psychology*. New Delhi: Oxford & IBH.
- Yaniv, I., & Meyer, D. E, (1987). Activation and meta-cognition of inaccessible stored information: Potential bases for incubation effects in problem solving. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 13(2), 187-205.
- Zeigarnik, B. (1927). Das Behalten erledigter und unerledigter Handlungen. *Psychologische Forschung*, 9, 1-85.
- Zeigarnik, B. (1938). On finished and unfinished tasks. In W. D. Ellis (Ed.), *A sourcebook of*

Gestalt psychology (pp.300-314), New York: Humanities press.

The 37 Incubation Studies (in chronological order)

- Mednick, M., Mednick, S. & Mednick, E. (1964). Incubation of creative performance and specific associative priming. *Journal of Abnormal and Social Psychology*, 69, 84-88.
- Gall, M., & Mendelsohn, G.A. (1967). Effects of facilitating techniques and subject experimenter interaction on creative problem solving. *Journal of Personality and Social Psychology*, 5, 211-216.
- Fulgosi, A. & Guilford, J. P. (1968). Short-term incubation in divergent production. *American Journal of Psychology*, 81, 241-246.
- Dreistadt, R. (1969). The use of analogies and incubation in obtaining insights in creative problem solving. *The Journal of Psychology*, 71, 159-175.
- Murray, H. G. & Denny, J. P. (1969). Interaction of ability level and interpolated activity (opportunity for incubation) in human problem solving. *Psychological Reports*, 24(1), 271-276.
- Fulgosi, A., & Guilford, J. (1972). A further investigation of short-term incubation. *Acta Instituti Psychologici*, 64-73, 67-70.
- Dominowski, R. L. & Jenrick, R. (1972). Effect of hints and interpolated activity on solution of an insight problem. *Psychonomic Science*, 26(6), 335-338.
- Silveira, J. (1971). *Incubation: The effect of interruption timing and length on problem solution and quality of problem processing*. Unpublished doctoral dissertation, University of Oregon.
- Peterson, C. (1974). Incubation effects in anagram solution. *Bulletin of the Psychonomic Society*, 3, 29-30.
- Bennett, S. (1975). *Effect of incubation on the associative process of creativity*. Unpublished doctoral dissertation. University of Notre Dame.

- Olton, R. M. & Johnson D. M.(1976). Mechanism of incubation in creative problem solving. *American Journal of Psychology*, 89(4), 617-630.
- Beck, J. (1979). *The effect of variations in incubation activity and incubation time on creative response production*. Unpublished doctoral dissertation, New York University.
- Brockett, C. (1985). *Neuropsychological and cognitive components of creativity and incubation*. Unpublished doctoral dissertation, Virginia Commonwealth University.
- Patrick, A. S. (1986). The role of ability in creative "incubation". *Personal and Individual Differences*, 7(2), 169-174.
- Yaniv, I & Meyer, D. E, (1987). Activation and meta-cognition of inaccessible stored information: Potential bases for incubation effects in problem solving. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 13(2), 187-205.
- Browne, B. A. & Cruse, D. F.(1988). The incubation effect: Illusion or Illumination. *Human Performance*, 1(3), 177-185.
- Smith, S. M. & Blankenship, S. E. (1989). Incubation effects. *Bulletin of the Psychonomic Society*, 27(4), 311-314.
- Frankel, A. (1990). *The effect of incubation and eidetic imagery training on tasks of problem solving*. Unpublished doctoral dissertation, Fordham University.
- Kaplan, C. (1990). *Hatching a theory of incubation: Does putting a problem aside really help? If so, why?* Unpublished doctoral dissertation, Carnegie-Mellon University.
- Dorfman, J. (1990). *Metacognition and incubation effects in insight problem solving*. Unpublished doctoral dissertation, University of California, San Diego.
- Smith, S. M. & Blankenship, S. E. (1991). Incubation and the persistence of fixation in problem solving. *American Journal of Psychology*, 104, 61-87.

- Goldman, W., Wolters, N., & Winograd, E. (1992). A demonstration of incubation in anagram problem solving. *Bulletin of the Psychonomic Society*, 30, 36-38.
- Torrance-Perk, J. (1997). *The incubation effect: Implication for underlying mechanism*. Unpublished doctoral dissertation. University of Waterloo.
- Hansberry, M. T. (1999). *Fixation and incubation effects in problem solving*. Unpublished doctoral dissertation. University of Hampshire
- Henley, R. J. (1999). *Priming Incubated Problems*. Unpublished doctoral dissertation. University of Sussex.
- Jamieson, B. A. (1999). *Incubation and aging: The nature of processing underlying insight*. Unpublished doctoral dissertation University of Georgia.
- Dodds, R. A., Smith, S. M., & Ward, T. B. (2002). The use of environmental clues during incubation. *Creativity Research Journal*, 14(3/4), 287-305
- Medd, E. & Houtz, J. C. (2002). *Educational Research Quarterly*, 26(2), 13-16.
- Moss, S. A. (2002). The Impact of Environmental Clues in Problem Solving and Incubation: The Moderating Effect of Ability. *Creativity Research Journal*, 14(2), 207.
- Seabrook, R., & Dienes, Z. (2003). Incubation in problem solving as a context effect. *Proceedings of the 25th meeting of the Cognitive Science Society, Boston, July 31-Aug 2, 2003*. Lawrence Erlbaum Associates: Mahwah, NJ
- Both, L., Needham, D., & Wood, E. (2004) Examining tasks that facilitate the experience of incubation while problem solving. *Alberta Journal of Educational Research*, 50(1), 57-67.
- Penney, C. G., Godsell, A. Scott, A., & Balsom, R. (2004). Problem variables that promote incubation effects. *Journal of Creative Behavior*, 38(1), 35-55.
- Segal, E. (2004). Incubation in insight problem solving. *Creativity Research Journal*, 16(1),

141-149.

Snyder, A., Mitchell, D. J., Ellwood, S., Yates, A., & Pallier, G. (2004). *Nonconscious idea generation. Psychological Reports (94)*, 1325-1330

Christensen, B.T. & Schunn, C. D. (2005). Spontaneous access and analogical incubation effects. *Creativity Research Journal, 17(2&3)*, 207-220.

Sio, U. N. & Rudowicz, E. (2007). The role of an incubation period in creative problem solving. *Creativity Research Journal, 19 (2-3)*, 307-318.

Vul, E. & Pashler, H. (2007).Incubation is helpful only when people are misled. *Memory, and Cognition,35(4)*, 701-710.

Author Note

Ut Na Sio and Thomas C. Ormerod, Department of Psychology, Lancaster University, United Kingdom.

Correspondence concerning this article should be addressed to Thomas C. Ormerod, Department of Psychology, Lancaster University LA1 4YF, United Kingdom. (email: t.ormerod@lancaster.ac.uk).

Footnotes

¹ The weighting was the inverse of the within-study variance of the effect estimate, and the formula for the within-study variance was $((2 * \text{square root of total } N) + (N \text{ of experimental} * N \text{ of control} * \text{square root of the unbiased effect size})) / (2 * \text{total } N * N \text{ of experimental} * N \text{ of Control})$, where N is the number of participants in that condition (Cooper & Hedges, 1994). Three study were excluded when computing the weighted average unbiased effect size estimate and the Cochran's Q value because it was a within-subjects study and all participants were involved in both control and incubation conditions. Thus, the weighting formula could not apply to it.

²The between-studies variance was equal to $[Q-(k-1)]/c$, where Q is the Cochran's Q value and k was the number of studies. The formula for c was $((\text{the sum of the inverse of the within-study variance}) - (\text{the sum of the square of the inverse of the within-study variance}) / (\text{the sum of the inverse of the within-study variance}))$, and was suggested by Cooper & Hedges (1994).

³The standard deviation of the weighted mean, also known as the "standard error", was calculated as the square root of $1/\sum w_i$, and the 95% Confidence Interval was calculated as the weighted mean $\pm 1.96 * \text{standard deviation of the weighted mean}$. (Hedges & Olkin, 1985)

Table 1 Summary of Incubation Studies

Year	Author	Problem	Misleading cues in problems	Preparation period	Incubation period	Incubation Task	Cues during incubation period	Other Factors	Incubation Effect reported	Included in Meta-Analysis
1964	Mednick, M. T., Mednick, S. A., Mednick, E. (Experiment 1)	RAT	No	1 min	Not specified	Analogy vs Analogy + Cues	Yes	/	Yes, in Analogy +Cues condition	No, no control group
1964	Mednick, M. T., Mednick, S. A., Mednick, E. (Experiment 2)	RAT	No	1 min	Not specified	Analogy vs Analogy + Cues	Yes	Cue relevance and correctness; Problem solving ability	Yes, high ability group in Analogy + Cues condition	No, no effect size information
1967	Gall, M., & Mendelsohn, G. A.	RAT	No	2 min	0, 25 min	Non-verbal vs Free associates + Cues	Yes	/	No	Yes, estimated from $p < \text{value}$
1968	Fulgosi, A., Guilford, J. P.	Consequence Task	No	2 min	0, 10, 20 min	Number-Series (10 and 20 min)	No	/	Yes, in 20 min condition	Yes, estimated from statement of significance, and $p < \text{value}$

1969	Dreistadt, R	Farm & Tree						In control	
		Planting Problems	No	5 min	0, 8 min	Guess Card vs Guess Card + cues	Yes	condition, half the participants received cues.	Yes, with cues, Farm Problem only
1969	Murry, H. G. & Denny J. P.	Saugstad's ball problem	No	5 min	0, 5 min	Syllogisms, and tracing complex sequences of digits/numbers	No	Problem solving ability	Yes, in low ability group.
									Yes
1972	Fulgosi, A., Guilford, J. P.	Consequence Task	No	2 min	0, 30 min, 60 min	Number-Series (30 and 60 min)	No	/	Yes, in 30-min condition
									No, no effect size information
1972	Dominowski, R., & Jenrick, R. (Experiment 1)	Hat Rack Problem	No	5 min	0, 10 min	Free association vs Anagram	Yes	One group received cues throughout.	Yes, estimated from statement of significance
1972	Dominowski, R., & Jenrick, R. (Experiment 2)	Hat Rack Problem	No	3 min	0, 3 min	Free association vs Anagram	No	One group received cues. Problem solving ability	No Yes, estimated from statement of significance
1972	Silveira, J. M. (Experiment 1)	Necklace Problem	No	3 min vs 13 min	0, 30, 210 min	Read + Free Activity	No	/	Yes Yes

1972	Silveira, J. M. (Experiment 2)	Necklace Problem	No	13 min	210 min	Read + Free Activity	No	/	Yes	Yes
1972	Silveira, J. M. (Experiment 3)	Consequence Task	No	2 min	210 min	Read + Free Activity	No	/	No	Yes
1974	Peterson, C.	Anagram	No	20 sec	3.6 mins	Other Anagram	No	Task Difficulty	Yes	Yes
1975	Bennett, S. M.	RAT	No	1 min	10 min	Hear music vs mathematics problems Rest vs Stroop + count backward vs Review problem vs lecture vs Hear music	No	/	No	No, no control group
1976	Olton, R. M., & Johnson, D. M.	Farm Problem	No	10 min	0, 15 min	Relax vs Write essay	Yes	/	No	Yes
1979	Beck, J.	Verbal Divergent Thinking Task	No	12 min	0, 20, 30 min	Relax vs Write essay	No	/	Yes, with longer incubation period	Yes
1985	Brockett, C. A.	Brick and RAT	No	10 min	0, 20 min	Questionnaires	No	/	Yes	Yes

1986	Patrick, A. S.	RAT	No	2 min	0, 5 min	Conversation vs Mental rotation	No	/	Yes, in rotation Condition	Mental	Yes, estimated from $p < \text{value}$ and statement of significance
1987	Yaniv, I, & Meyer, D. E.	Rare-word Definition Task	No	Not specified	Not Specified	Questionnaire	No	/	Yes		No, no post-incubation performance measure
1988	Browne, B. A. & Cruse, D.F. (Experiment 1)	Farm Problem	No	20 min	0, 5 min	Hear music, Draw graph, Memorize text	No	/	No		Yes
1988	Browne, B. A. & Cruse, D.F. (Experiment 2)	Farm Problem	No	20 min or 25 min	0, 5 min	Hear music, Draw graph, Memorize text	No	/	No		Yes
1989	Smith, S. M. & Blankenship, S.E. (Experiment 1)	Rebus	Misleading or Useful cues	0.5 min	0, 5, 15 min	Rest vs Music Perception	No	Useful cues presented for second attempt	Yes, 15 min > 5 min incubation and control group.		Yes

	Smith, S. M. &	Misleading						Yes, 15 and 5		
1989	Blankenship,S.E. Rebus (Experiment 2)	or Useful cues	0.5 min	0, 5, 15 min	Rest vs Music Perception	No	/	min incubation > Yes control group.		
					Rebus with			Yes, no		
	Smith, S. M. &	Misleading		0, 10, 15	Music			difference	Yes, estimated	
1989	Blankenship,S.E. Rebus (Experiment 3)	or Useful cues	0.5 min	min	perception, Mathematics, or Rest	No	/	among incubation conditions	from p < value	
		Yes,							Yes, estimated	
	Smith, S. M. &	Misleading			Read vs			Yes, in Read	from p < value	
1989	Blankenship,S.E. Rebus (Experiment 4)	or Useful cues	0.5 min	0, 5 min	Mathematics	No	/	story condition	and statement of significance	
					Word problems			Yes, in both		
1990	Dorfman, J. (Experiment 3)	Word Puzzle Task	No	0.49 min	0, 5, 15 min	vs word problems + cues	No	/	incubation conditions	Yes
									Yes, estimated	
1990	Dorfman J (Experiment 4)	Word Puzzle Task	No	0.49 min	0, 3, 8, 13 delay	Number series + cues		No	from statement of significance	
1990	Kaplan, C. A. (Experiment 1)	Consequence Task	No	2 min	0, 30 min	Psychometric test battery	No	/	Yes	Yes

1990	Kaplan, C. A. (Experiment 2)	Consequence Task	No	2 min	0, 32 min	Mathematics + lecture	No	/	Yes	Yes,
1990	Kaplan, C. A. (Experiment 3)	Consequence Task	No	2 min	0, 30 min	Lecture	No	/	No	Yes, estimated from statement of significance
1990	Kaplan, C. A. (Experiment 4)	Consequence Task	No	2 min	40 min	Mathematics and insight problems	No	/	Yes	Yes
1991	Smith, S. M. & Blankenship, S.E. (Experiment 1).	RAT	Yes	0.5 min	0, 5 min	Read science fiction	No	Problem solving ability	Yes, with misleading cues > for low ability participants	Yes
1991	Smith, S. M. & Blankenship, S.E. (Experiment 2)	RAT	Yes	1 min	0, 5 min	Read science fiction	No	Problem solving ability	Yes, with misleading cues	Yes
1991	Smith, S. M. & Blankenship, S.E. (Experiment 5)	RAT	Yes	0.5 min	0, 0.5, 2 min	Free association	No	Problem solving ability	Yes, with misleading cues, < for low ability participants	Yes, estimated from $p <$ value and statement of significance

	Goldman, W. P.,					General				
1992	Wolters, N. & Winograd, E.	Anagram	No	40 min	0, 20, 1440 min	knowledge + free activity	No	/	Yes, in 1440 min condition.	Yes
1992	Houtz, J. C. & Franke, A. D.	Life-relevant problem	No	10 min	10 min	Anagram task	No	/	Yes	Yes
		RAT, Word Completion,				Rest +lexical decision vs Rest + lexical decision + cues				
1997	J. (Experiment 1)	Candle & Radiation problems.	Yes,	25 min	0, 10 min		Yes	/	No	Yes
1997	J. (Experiment 2)	Candle & Radiation Problems, RAT	Yes, Misleading cues	1 min	0, 8 min	Memory (Candle); Read (Radiation); Analogy (RAT)	Yes	/	No	Yes
1997	J. (Experiment 3)	Two-String & Hatrack Problems, Radiation	Yes, Misleading cues	3 min	0, 8 min	Analogy (Two-String); Paired associates (Hatrack); Read (Radiation)	Yes	/	Yes	Yes

1998	Hansberry, M.T. (Experiment 2)	Riddles	Yes,	1 min	0, 15 min	RAT	No	/	No	Yes
1998	Hansberry, M. T. (Experiment 3)	RAT	Yes,	0.5 min	0, 10 min	RAT	No	Problem solving ability	No	Yes
1999	Henley, R. J. (Experiment 3.2)	Anagram	No	0.25 min	0, 1440 min	Free Activity	No	/	No	Yes, estimated from statement of significance
1999	Henley, R. J. (Experiment 4)	Anagram	No	0.93 min	0, 1440 min	Free Activity	No	/	No	Yes, estimated from statement of significance
1999	Jamieson, B. A. (Experiment 1)	RAT	Yes	0.33 min	0, 5 min	Mathematics	No	/	No	Yes
1999	Jamieson, B. A. (Experiment 2)	RAT	Yes	0.33 min	0, 5 min	Mathematics	No	/	No	Yes
2002	Dodds , R., Smith, S. M., & Ward, T. B. (Experiment 1)	RAT	No	10 min	15 min	Insight problem with: Drawing vs Make a word test vs Make a word test + Cues	Yes	Cues (answer, relevant information or related word)	Yes, when answer or unrelated word presented during incubation	No, no control group

						Insight problem		Cues (answer, Yes, when	
						with: Drawing		relevant	answer
2002	Dodds , R., Smith, S. M., & Ward, T. B. (Experiment 2)	RAT	No	10 min	0, 15 min	vs Make a word test vs Make a word test + Cues	Yes	information or related word)	presented during incubation
						Unrelated		Prompt to think	Yes, if working
2002	Medd., E & Houtz, J.	Creative Writing	Yes,	10 min	0, 10 min	writing task vs Related Writing task	No	about problem during incubation	Yes, estimated from p < value and statement of significance
								Tasks described as solvable or unsolvable.	Yes. Interaction between ability and problem type
2002	Moss, S. A.	RAT	No	0.5 min	0, 15 min	Verbal reasoning task +Cues	Yes	Problem solving ability	No, no control group
2003	Seabrook, R., & Dienes, Z	Anagram	No	0.25 min	7 min	Word generation tasks	Yes	Cues (Irrelevant or relevant)	Yes, in Relevant-Cue group
2004	Both, L., Needham, D., & Wood, E.	Anagram	No	1.67 min	0, 6 min	Letter search; Questionnaire	No	/	Yes

	Both, L.,					Letter and digit				
2004	Needham, D., & Wood, E	Anagram	Yes	1.67	0, 6 min	search; Questionnaire	Yes	/	Yes	Yes
	Penny, C. G., Godsell, A.,					Word completion +				
2004	Scott, A., & Balsom, R.	Anagram	No	5.75 min	15 , 180 min	Cues vs Inactivity	No	/	Yes, in 15 min condition	No, no control group
	(Experiment 1)									
	Penny, C. G., Godsell, A.,					Word completion +				
2004	Scott, A., & Balsom, R.	Anagram	No	5.75 min	15, 1440 min	Cues vs Inactivity	No	/	Yes, in 15 min condition	No, no control group
	(Experiment 2)									
	Penny, C. G., Godsell, A.,					Word completion +				
2004	Scott, A., & Balsom, R.	Anagram	No	5.75 min	0, 30, 120 min	Cues vs Inactivity	No	/	Yes	Yes
	(Experiment 3)									

		Insightful								
2004	Segal, E.	Mathematic Puzzle	No	20 min	0, 4 , 12 min	Read vs Word puzzle	No	/	Yes, with word puzzle	Yes
	Snyder, A.,									
2004	Mitchell, J., Ellwood, S., & Yates A.	Divergent thinking task	No	5 min	5 min	Conversation	No	/	Unknown	No
	Christensen, B.									
2005	T. & Schunn, C.D.	Insight Puzzle	No	Not specified	Not specified	Analogous or distracter rating task	Yes	Cues (analogous cue or distracter cue)	Yes, in analogous cue condition	No, no control group
						Mental Rotation Task +		Chess expertise; Cues to chess	Yes, experts more sensitive to relevant	No, no post-incubation performance measure
2007	Sio, U. N & Rudowicz, E	RAT	Yes	1 min	0, 2 min	Mathematics Task vs Listening Music	No	expertise were related, neutral, or misleading.	concepts after incubation.	
2007	Vul, E. & Pashler,H.	Anagram	No	12 min	0, 5 min	Video game	No	Task difficulty	No	Yes
2007	Vul, E. & Pashler,H.	RAT	Yes	12 min	0, 5 min	Video game	No	Task difficulty	Yes, in the fixation group	Yes

Table 2. *Coding System*

Variables	Coding Description
Author	Author(s) of the study
Year	Year the study was published
Total	Total number of participants
Problem Type	0- Creative problem, e.g., consequence task(C) 1- Visual problem, e.g., Farm Problem, Radiation Problem (V) 2- Linguistic problem, e.g, Remote Associates Task, Anagram, Rebus (Li)
Misleading Cues	0- No misleading cues 1- Misleading cues embedded in the problem
Preparation Period	Amount of time spent on each problem before the incubation period (in minute)
Incubation Period	Length of the incubation period (in minute)
Incubation Task	0- Rest (R) 1- Low cognitive demand task, e.g., drawing picture, reading (L) 2- High cognitive demand task, e.g., mental rotation task, memory test (H)
Cues	Presence of relevant cues during the incubation period 0- No cue 1- Yes

Table 3. *Stem-and-Leaf Display of 117 Unbiased Effect Sizes (Outliers were recoded)*

Stem	Leaf
-.7	1
-.5	9, 8, 8
-.4	0
-.3	8, 5, 3
-.2	
-.1	8, 7, 4, 4, 0
-.0	3
.0	0, 3, 4, 5, 5, 6, 6, 7, 9
.1	0, 0, 0, 0, 0, 1, 1, 3, 4, 4, 4, 5, 7, 7
.2	1, 4, 5, 8, 8
.3	1, 3, 6, 6, 6
.4	0, 0, 1, 1, 2, 4, 5, 7, 7, 8, 8, 9
.5	0, 2, 2, 6, 6, 6, 6, 6, 6, 7
.6	2, 2, 4, 4, 4, 5, 6, 7
.7	0, 1, 2, 3, 4
.8	6
.9	0
1.0	5, 5, 7, 7, 8, 9
1.1	7
1.2	
1.3	
1.4	
1.5	
1.6	8
1.7	0
1.8	2, 2, 2, 2

Note. Each unbiased effect size is recorded with the first digit in the stem column, and the second digit in the leaf column.

Table 4. *The Random Variance Component, Weighted Mean, Standard Deviation, Standard Error, and 95% Confidence Interval of the Effect Size Estimate by Each Categorical Moderator*

	Problem Type			Misleading Cues		Incubation Task			Cues	
	Linguistic	Verbal	Creative ^a	Yes	No ^b	High Load	Low Load	Rest ^c	Yes	No Cue ^d
Number of studies	65	35	14	29	85	76	22	16	32	82
Random Variance Component	.00281	0	.37418	0	.12004	.06478	0	.30000	0	.10409
Mean	.22	.26	.29	.35	.32	.24	.52	.46	.24	.34
Standard Deviation	.05	.08	.14	.09	.06	.05	.10	.18	.09	.06
95% Confidence Interval	.32 & .13	.41 & .10	.55 & .02	.53 & .17	.43 & .20	.35 & .13	.72 & .32	.82 & .11	.41 & .07	.45 & .22
Mean comparison with the reference group	t(77) = -.43	t(47)=-.19		t(112)=.30		t(90)=1.17	t(36)=.28		t(112)=.90	

Note. The lower confidence intervals of all the weighted unbiased effects are larger than zero, suggesting that the mean is significant larger at zero, $p < .05$.

^{a,b, c, d} The reference groups in mean comparisons

Table 5. *Regression Model for “Linguistics Problem” Studies (N=65)**A. ANOVA significance test.*

	Sum of Squares	df	Mean Square	F	p-value
Model	13.42	5	2.68	3.15*	.01
Residual	5.24	59	.85		
Total	63.66	64			

B. Summary of the Regression Model

Predictor	Unstandardized β	Standardized β	Corrected SE_{β}	z-score
Incubation Task				
High cognitive load task (H)	.06	.06	.21	.27
Low cognitive load task (L)	.54*	.43	.25	2.15
Misleading Cues	.15	.17	.12	1.24
Ration of the preparation period to the incubation period	.15	.12	.19	.80
Relevant Cue	-.04	-.04	.13	-.30

Note. Random Variance Component = .00281, $R^2 = .21$, Another regression with the same predicting variables, except replacing the variable “ration of length of the preparation to the incubation period” by the variables “Incubation Period” and “Preparation Period”, was carried out. The pattern of the results was comparable but it had lower explanatory power, and lower significant level, $R^2 = .20$, $F(6, 58) = 2.47$, $p = .034$. Neither the variable “preparation period” nor “incubation period” was significant.

* $p < .05$.

Table 6. *Regression Model for “No Misleading Cue” Studies (N = 85)**A.. ANOVA significance test.*

	Sum of Squares	df	Mean Square	F	p-value
Model	2.25	7	2.89	3.56*	.002
Residual	62.62	77	.81		
Total	82.87	84			

B. Summary of the Regression Model

Predictor	Unstandardized β	Standardized β	Corrected SE_{β}	z-score
Problem Type				
Visual problem (V)	-.59*	-.51	.19	-3.02
Linguistic problem (Li)	-.31	-.29	.18	-1.74
Incubation Task				
High cognitive load task (H)	-.18	-.14	.22	-.80
Low cognitive load task (L)	.06	.04	.26	.23
Length of the Incubation Period	<.001	.055	<.001	.45
Length of the Preparation Period	.03*	.36	.01	2.16
Relevant Cue	-.03	-.02	.16	-.20

Note. Random Variance Component = .12004, $R^2 = .24$.

* $p < .05$,

Table 7. *Regression Model for “No Relevant Cue” Studies (N = 82)**A. ANOVA significance test.*

	Sum of Squares	df	Mean Square	F	p-value
Model	29.34	6	4.89	7.35*	<.001
Residual	49.94	75	.67		
Total	79.28	81			

B. Summary of the Regression Model

Predictor	Unstandardized β	Standardized β	Corrected SE_{β}	z-score
Problem Type				
Visual problem (V)	-.79**	-.66	.19	-4.10
Linguistic problem (Li)	-.22	-.21	.18	1.24
Incubation Task				
High cognitive load task (H)	-.25	-.22	.21	-1.20
Low cognitive load task (L)	.16	.12	.25	.64
Length of the Incubation Period	<.001	.02	0	.13
Length of the Preparation Period	.05**	.51	.02	3.01
Misleading Cue	.09	.07	.18	.53

Note. Random Variance Component = .10409, $R^2 = .37$.

* $p < .05$, ** $p < .001$.

Table 8. *Regression Model for “High Cognitive Load Incubation Task” Studies*

(N=75)

A. ANOVA significance test.

	Sum of Squares	df	Mean Square	F	p-value
Model	7.36	6	1.47	1.53	.19
Residual	66.19	69	.96		
Total	73.55	74			

B. Summary of the Regression Model

Predictor	Unstandardized β	Standardized β	Corrected SE_{β}	z-score
Problem Type				
Visual problem (V)	-.41*	-.34	.20	-2.01
Linguistic problem (Li)	-.28	-.30	.17	-1.58
Length of the Incubation Period	-.01	-.12	.01	-.87
Length of the Preparation Period	.02	.21	.02	1.31
Misleading Cue	.08	.06	.16	.51
Relevant Cue	.09	.08	.14	.62

Note. Random Variance Component = .06478, $R^2 = .10$.

* $p < .05$.

Table 9. *Regression Model for All Studies (N=114)**A. ANOVA significance test.*

	<i>Sum of Squares</i>	<i>df</i>	<i>Mean Square</i>	<i>F</i>	<i>p-value</i>
Model	25.56	8	3.20	4.08	<.001
Residual	82.22	105	.78		
Total	107.79	113			

B. Summary of the Regression Model

Predictor	<i>Unstandardized</i> β	<i>Standardize</i> β	<i>Corrected</i> SE_{β}	<i>z-score</i>
Problem Type				
Visual problem	-.60**	-.52	.15	-3.61
Linguistic problem	-.31*	-.31	.14	-1.96
Misleading Cues	.16	.13	.12	1.16
Incubation Task				
Low cognitive load	.14	.10	.18	.72
High cognitive load	-.16	-.15	.14	-1.02
Preparation Period	.03*	.35	.01	2.29
Incubation Period	<.001	.04	<.001	<.001
Cue Presented	-.01	-.01	.11	-.09

Note: Random Variance Component = .0834, $R^2 = .24$,

* $p < .05$, ** $p < .001$

Table 10. *Regression Model for all studies - interactions(N=114)**A. ANOVA significance test.*

	Sum of Squares	df	Mean Square	F	p-value
Model	35.60	12	2.97	4.15	<.001
Residual	72.18	101	.72		
Total	107.78	113			

B. Summary of the Regression Model

Predictor	Unstandardized β	Standardized β	Corrected SE_{β}	z-score
Problem Type* Incubation Task				
Visual problem (V) *Low cognitive load task (L)	.17	.08	.35	.50
Visual problem (V) *High cognitive load task (H)	.90*	.61	.35	2.60
Linguistic problem (Li)* Low cognitive load task (L)	.72	.43	.38	1.93
Linguistic problem(Li)*High cognitive load task (H)	1.00*	.97	.38	2.59
Problem Type				
Visual problem (V)	-1.29**	-1.10	.32	-4.00
Linguistic problem (Li)	-1.24**	-1.21	.40	-3.14
Incubation Task				
High cognitive load task (H)	-.91**	-.81	.30	-3.00
Low cognitive load task (L)	-.12	-.09	.33	-.38
Misleading Cues	.24	.19	.15	1.54
Length of the Incubation Period	<.001	.018	<.001	<.001
Length of the preparation period	.03*	.01	.01	1.97
Answer was presented	.03	.03	.13	.26

Note. Random Variance Component =.08340, $R^2 = .33$,

* $p < .05$, ** $p < .001$.

Table 11

Summary of Coefficient Differences of the Regression Model

The multiplicative terms	The coefficient difference	Corrected SE	z-score
$L_{ij} * H_j - L_{ij} * R_j$	.08	.49	.17
$L_{ij} * L_j - L_{ij} * R_j$	.60	.50	1.21
$L_{ij} * L_j - L_{ij} * H_j$	.52	.70	.74
$I_{nj} * H_j - I_{nj} * R_j$	-.01	.46	-.03
$I_{nj} * L_j - I_{nj} * R_j$	.05	.48	.10
$I_{nj} * L_j - I_{nj} * H_j$	.06	.66	.09
$C_j * H_j - C_j * R_j$	-.91*	.30	-3.00
$C_j * L_j - C_j * R_j$	-.12	.33	-.38
$C_j * L_j - C_j * H_j$	.79 ⁺	.45	1.77
$I_{nj} * H_j - C_j * H_j$	-.39	.47	-.82
$L_{ij} * H_j - C_j * H_j$	-.25	.55	-.45
$L_{ij} * H_j - I_{nj} * H_j$	.14	.72	.19
$I_{nj} * L_j - C_j * L_j$	-1.11*	.47	-2.36
$L_{ij} * L_j - C_j * L_j$	-.52	.55	-.95
$L_{ij} * L_j - I_{nj} * L_j$	.60	.72	.82
$I_{nj} * R_j - C_j * R_j$	-1.29**	.32	-4.00
$L_{ij} * R_j - C_j * R_j$	-1.24**	.40	-3.14
$L_{ij} * R_j - I_{nj} * R_j$	.04	.51	.08

⁺ $p = .08$, * $p < .05$, ** $p < .001$

Table 12. *Regression Model for studies excluding creative problem (N=100)**A. ANOVA significance test.*

	Sum of Squares	df	Mean Square	F	p-value
Model	15.04	8	1.88	2.04*	.05
Residual	83.92	91	.92		
Total	98.96	99			

B. Summary of the Regression Model

Predictor	Unstandardized β	Standardize β	Corrected SE_{β}	z-score
Problem Type* Incubation Task				
Visual problem (V) *Low cognitive load task (L)	-.21	-.12	.28	-.77
Visual problem (V) *High cognitive load task (H)	.09	.08	.24	.38
Problem Type				
Visual problem (V)	-.15	-.15	.21	-.68
Incubation Task				
High cognitive load task (H)	-.01	-.01	.18	-.05
Low cognitive load task (L)	.45*	.41	.20	2.27
Misleading Cues	.13	.13	.11	1.10
Length of the preparation period/incubation	.08	.20	.05	1.56
Answer was presented	.01	.01	.11	.13

Note. Random Variance Component = .00031, $R^2 = .15$, Another regression was

carried out with the same variables, replacing the variable “ratio of length of preparation to incubation period” with “Incubation Period” and “Preparation Period”,.

The regression model was not significant, $F(9, 90) = 1.27$, $p = .10$. Neither

“preparation period” nor “incubation period” was significant.

* $p < .05$.

Table 13. *Summary of Coefficient Differences of the Regression Model*

The multiplicative terms	The coefficient difference	Corrected SE	z-score
$Li_j * H_j - Li_j * R_j$	-.01	.17	-.05
$Li_j * L_j - Li_j * R_j$	.45*	.20	2.27
$Li_j * L_j - Li_j * H_j$	.46	.26	1.76
$V_j * H_j - V_j * R_j$	.08	.30	.28
$V_j * L_j - V_j * R_j$	.24	.34	.69
$V_j * L_j - V_j * H_j$	.15	.45	.34
$V_j * H_j - Li_j * H_j$	-.06	.32	-.17
$V_j * L_j - Li_j * L_j$	-.36	.35	-1.03
$V_j * R_j - Li_j * R_j$	-.15	.21	-.68

* $p < .05$,

Figure Captions

Figure 1. Funnel Plot of the Studies included in this Meta-Analysis.

Note. The dotted line indicates the mean unbiased effect size.

Appendix A

Task Description

Task Type	Description	Solution/Sample Task
Creative Problems		
Brick Task	Participant has to list all the uses he/she can think of for a brick.	/
Consequence Task	Participant has to list out as many consequences of an event as he/she can foresee	Sample Task: What would be the results if everyone suddenly lost the ability to read and write?
Creative Writing	/	Sample Task: Write about three concrete objects: a Koosch ball, a wooden type of propeller, and a triangular frisbee
Visual Insight Problems		
Candle Problem	Participant has to support a candle on wall by using some matches and a box of tacks.	Solution: Use a tack to attach box to the wall, then drip some wax onto the box and mount the candle on the box
Farm Problem	Participant has to divide an L-shaped farm into four parts which have the same size and shape.	Solution:

Hat Rack Problem

Participant has to construct a stable hatrack by using two boards and a Camp.

Solution:

Wedge the two boards between the ceiling and the floor, holding them in place with the clamp, and with the clamp also serving as a hook.

Insightful mathematic problem

Participant has to compute separately the area of the square and that of the parallelogram shown below:

Solution

Restructure the given shape into partially overlapping triangle ABG and ECD. The sum of their area is $2 \times \frac{ab}{2} = ab$

Necklace problem

Necklace Problem

Participant is given four pieces of chain, and each chain is made up of three links, he/she has to join all the pieces by only opening and closing three links.

Solution:

Open all 3 links of one chain, and join the other three chains together.

Radiation Problem A patient has an inoperable tumor in the middle of the body, and there is a ray at a strong intensity that can destroy the tumor, but the ray also harms the healthy tissue that it travels through. At low intensities, the ray will spare the healthy tissue, but will not destroy the tumor. Participant has to think out a way to use the ray to destroy the tumor without damaging healthy tissue.

Solution:
Direct multiple low-intensity rays simultaneously toward the tumor from different directions

Saugstad's "ball Problem" Participant has to transfer steel balls from a drinking glass to a cylinder from a distance of 8 ft by using the following objects: nail, a pair of pliers, a length of string, a pulley, elastic bands, and newspaper. The glass sits on a moveable frame.

Solution: Bend the nail into a hook then attach it to the string. Use it to drag back the frame and remove the balls. Transfer balls into the cylinder using a tube constructed of newspaper and elastic band.

Tree Problem Participant has to plant 10 trees in five rows with four trees in each row

Solution:

(The trees are represented by the dots)

Linguistic Insight Problems

Anagram Participant has to rearrange the scrambled letters to form a new word

Sample Task:

The scrambled letters are "t s l t i n e", and one possible solution is "silent"

RAT	Three words are presented to the participant, he/she has to think of a word that can form associations with each of them.	<p>Sample Task:</p> <p>The three stimulus words are “ blue, cake, cottage”, and one possible solution is the word “cheese”</p>
Rebus	Participant has to figure out the phrase portrayed by the pictogram	<p>Sample task:</p> <p>The pictogram</p> <p>The answer is “first aid”</p>
Riddle	/	<p>Sample Task: A wine bottle is half-filled and corked. How can you drink all of the wine without removing the cork from the bottle?</p> <p>Answer: The cork can be opened by pushing it in.</p>
Word Associates Task	Six words are presented to the participant, he/she has to think out a word that can form an association with each of them.	<p>Sample Task: The six stimulus words are “school, chair, jump, noon, heels, wire”, and one possible solution is the word “high”</p>
Word Fragment Completion	Participant has to complete a word that has various letters missing	<p>Sample task:</p> <p>The stimulus is “OC__N”, and one</p>

Task

possible answer is “OCEAN”

Appendix B

Information Extracted from Each Independent Study

ID	Year	Author	Total N	Nature of the Problem	Misleading Cues	Preparation	Incubation	Incubation Task	Cues	Unbiased Effect Size
1	1967	Gall, M. & Mendelsohn, G.	60	2 (RAT)	0	2	25	2(Non-Verbal Task)	0	-.58
2	1967	Gall, M. & Mendelsohn, G.	60	2 (RAT)	0	2	25	2(Associations Training)	1	-.58
3	1967	Fulgosi, A., & Guilford, J. P.	50	1(Consequence Task)	0	2	10	2 (Number series task)	0	0
4	1967	Fulgosi, A., & Guilford, J. P.	49	1 (Consequence Task)	0	2	20	2 (Number series task)	0	.52
5	1969	Dreistadt R. (Study 1)	20	1 (Farm Problem)	0	5	8	1 (Guess playing card)	0	.34
6	1969	Dreistadt R. (Study 2)	20	1 (Farm Problem)	0	5	8	1 (Guess playing card)	1	.99
7	1969	Dreistadt R. (Study 3)	20	1 (Tree Problem)	0	5	8	1 (Guess playing card)	0	.21
8	1969	Dreistadt R. (Study 4)	20	1 (Tree Problem)	0	5	8	1 (Guess playing card)	1	.48
9	1969	Murry, H. G. & Denny J. P. (Study 1)	36	1 (Saugstad's ball problem, low ability group)	0	5	5	2 (Multiple choice syllogisms, and traced sequences of numbers and letters)	0	.62

Incubation and problem-solving

10								2 (Multiple choice		
	1969	Murry, H. G. & Denny J. P. (Study 2)	36	1 (Saugstad's ball problem, high ability group)	0	5	5	syllogisms, and traced sequences of numbers and letters)	0	-.59
11	1971	Silveria, J. M. (Study 1)	18	1 (Necklace Problem)	0	3	30	1 (read book)	0	.11
12								1 (read book for 30 mins and free activity 3 hrs)	0	.06
	1971	Silveria, J. M. (Study 2)	18	1 (Necklace Problem)	0	3	210			
13	1971	Silveria, J. M. (Study 3)	18	1 (Necklace Problem)	0	13	30	1 (read book)	0	.17
14								1 (read book for 30 mins and free activity 3 hrs)	0	.42
	1971	Silveria, J. M. (Study 4)	18	1 (Necklace Problem)	0	13	210			
15								1 (read book for 30 mins and free activity 3 hrs)	0	.44
	1971	Silveria, J. M. (Study 5)	32	1 (Necklace Problem)	0	13	210			
16	1972	Dominowski, R. L., & Jenrick, R.	27	1(Hat Rack Problem_		5	10	2(Free Association)	0	0
17	1972	Dominowski, R. L., & Jenrick, R.	30	1 (Hat Rack Problem)		3	3	2(Free Association)	0	0
18	1974	Peterson, C.	24	1 (Anagram)		.33	1.8	Anagram	1	.65

Incubation and problem-solving

19	1976	Olton, R. M., & Johnson, D. M. (Study 1)	21	1 (Farm Problem)	0	10	15	0 (Rest)	0	.10
20	1976	Olton, R. M., & Johnson, D. M.(Study 1)	21	1 (Farm Problem)	0	10	15	2 (Stroop test + Counting backward)	0	.11
21	1976	Olton, R. M., & Johnson, D. M (Study 1).	21	1 (Farm Problem)	0	10	15	2 (Review the problem)	1	.10
22	1976	Olton, R. M., & Johnson, D. M. (Study 1)	21	1 (Farm Problem)	0	10	15	2 (Have lecture)	1	0
23	1976	Olton, R. M., & Johnson, D. M. (Study 1)	20	1 (Farm Problem)	0	10	15	0 (Listen music)	0	-.10
24	1976	Olton, R. M., & Johnson, D. M. (Study 1)	20	1 (Farm Problem)	0	10	15	0 (Rest)	2	.10
25	1976	Olton, R. M., & Johnson, D. M. (Study 1)	20	1 (Farm Problem)	0	10	15	2 (Having lecture)	2	-.03
26	1979	Beck, J. (Study 1)	60	0 (verbal divergent thinking task)	0	12	20	0 (relax)	0	2.19
27	1979	Beck, J. (Study 2)	60	0 (verbal divergent thinking task)	0	12	30	2 (write essay)	0	1.07
28	1979	Beck, J. (Study 3)	60	0 (verbal divergent thinking task)	0	12	20	0 (Relax)	0	4.07

Incubation and problem-solving

29	1979	Beck, J. (Study 4)	60	0 (verbal divergent thinking task)	0	12	30	2 (Write essay)	0	4.04
30	1985	Brockett, C. A. (Study 1)	30	0 (Brick)	0	10	20	2 (Questionnaire)	0	.41
31	1985	Brockett, C. A. (Study 2)	30	2 (Remote Associates Task)	0	.33	20	2 (Questionnaire)	0	.36
32	1986	Patrick, A. S.	30	2 (Remote Associate Tasks)	0	2	5	2 (Conversation)	0	0
33	1986	Patrick, A. S.	30	2 (Remote Associate Tasks)	0	2	5	2 (Mental Rotation Task)	0	.66
34	1988	Browne, B. A . & Cruse, D. F. (Study 1)	60	1 (Farm Problem)	0	20	5	0 (Listening music)	0	.47
35	1988	Browne, B. A . & Cruse, D. F. (Study 1)	53	1 (Farm Problem)	0	20	5	1 (Graph drawing)	2	.24
36	1988	Browne, B. A . & Donna F. Cruse (Study 3)	55	1 (Farm Problem)	0	20	5	2 (Memory test)	0	-.17
37	1989	Smith, S. M. & Steven, B. (Study 1)	26	2 (Rebus)	1	0.5	5	0(Rest)	0	.45
38	1989	Smith, S. M. & Steven, B. (Study 1)	26	2 (Rebus)	1	0.5	5	2(Music Perception Task)	0	.48

Incubation and problem-solving

39	1989	Smith, S. M. & Steven, B. (Study 1)	26	2 (Rebus)	1	0.5	15	0(Rest)	0	.49
40	1989	Smith, S. M. & Steven, B. (Study 1)	26	2 (Rebus)	1	0.5	15	2(Music Perception Task)	0	.50
41	1989	Smith, S. M. & Steven, B.(Study 2)	25	2 (Rebus)	1	0.5	5	0(Rest)	1	.05
42	1989	Smith, S. M. & Steven, B.(Study 2)	25	2 (Rebus)	1	0.5	5	2(Music Perception Task)	1	.05
43	1989	Smith, S. M. & Steven, B.(Study 2)	25	2 (Rebus)	1	0.5	15	0(Rest)	1	.41
44	1989	Smith, S. M. & Steven, B.(Study 2)	25	2 (Rebus)	1	0.5	15	2(Music Perception Task)	1	.41
45	1989	Smith, S. M. & Steven, B. (Study 3)	29	2 (Rebus)	1	0.5	10	2(Rebus)	1	.56
46	1989	Smith, S. M. & Steven, B. (Study 3)	29	2 (Rebus)	1	0.5	15	0(Rebus +Rest)	1	.56
47	1989	Smith, S. M. & Steven, B. (Study 3)	29	2 (Rebus)	1	0.5	15	2(Rebus+Maths)	1	.56
48	1989	Smith, S. M. & Steven, B. (Study 3)	29	2 (Rebus)	1	0.5	15	0(Rebus+Music)	1	.56

Incubation and problem-solving

49	1989	Smith, S. M. & Steven, B. (Study 3)	29	2 (Rebus)	1	0.5	15	2(Rebus)	1	.56
50	1989	Smith, S. M. & Steven, B. (Study 4)	49	2 (Rebus)	1	0.5	5	1(Read Story)	0	0
51	1989	Smith, S. M. & Steven, B. (Study 4)	49	2 (Rebus)	1	0.5	5	2(Maths Task)	0	.49
52	1990	Dorfman, J. (Study 1)	15	2 (Word associate task)	0	.49	15	2(Number Series Task)	0	0
53	1990	Dorfman, J. (Study 1)	15	2 (Word associate task)	0	.49	5	2(Number Series Task)	0	1.07
54	1990	Dorfman, J. (Study 2)	27	2 (Word associate task)	0	.49	3	2(Number Series Task)	0	0
55	1990	Dorfman, J. (Study 2)	27	2 (Word associate task)	0	.49	8	2(Number Series Task)	0	0
56	1990	Dorfman, J. (Study 2)	26	2 (Word associate task)	0	.49	13	2(Number Series Task)	0	0
57	1993	Kaplan, C. A. (Study 1)	278	0 (Consequence test)	0	2	30	2 (Psychometric test battery)	0	.06
58	1993	Kaplan, C. A. (Study 2)	64	0 (Consequence test)	0	2	2	2 (Division problem)	0	.71
59	1993	Kaplan, C. A. (Study 3)	36	0 (Consequence test)	0	2	30	2 (Lecture)	0	0
60	1993	Kaplan, C. A. (Study 3)	20	0 (Consequence test)	0	4.57	28.08	2(Division and insight problem)	0	1.08

Incubation and problem-solving

61	1991	Smith, S. M. & Blankenship, S. E.(Study 1)	18	2 (Remote Associates Test)	1	.5	5	1 (Read science fiction)	0	1.70
62	1991	Smith, S. M. & Blankenship, S. E. (Study 1)	21	2 (Remote Associates Test)	0	.5	5	1 (Read science fiction)	0	.64
63	1991	Smith, S. M. & Blankenship, S. E.(Study2	30	2 (Remote Associates Test)	0	1	5	1 (Read science fiction)	0	.37
64	1991	Smith, S. M. & Blankenship, S. E. (Study 2)	30	2 (Remote Associates Test)	1	1	5	1 (Read science fiction)	0	.74
65	1991	Smith, S. M. & Blankenship, S. E. (Study 5)	16	2 (Remote Associates Test)	0	.5	2	2(Free Associations Task)	0	0
66	1991	Smith, S. M. & Blankenship, S. E. (Study 5)	18	2 (Remote Associates Test)	0	.5	.5	2(Free Associations Task)	0	0
67	1991	Smith, S. M. & Blankenship, S. E. (Study 5)	17	2 (Remote Associates Test)	1	.5	.5	2(Free Associations Task)	0	.99
68	1991	Smith, S. M. & Blankenship, S. E. (Study 5)	18	2 (Remote Associates Test)	1	.5	2	2(Free Associations Task)	0	.99
69	1992	Goldman, W. P., Wolters, N. C. W., & Winograd, E. (Study 1)	36	2 (Anagram)	0	.50	20	2 (General Knowledge Questionnaire)	0	.13

Incubation and problem-solving

70	Goldman, W. P., Wolters, N.								
	1992 C. W., & Winograd, E. (Study 1)	36	2 (Anagram)	0	.25	1140	2 (General Knowledge Questionnaire+Free)	0	.66
71	1992 Houtz, J.C. & Frankel	105	1 (Creative Writing)	0	10	10	1 (Anagram)	0	.03
72	1997 Torrance-Perks, J. (Study 1)	15	2 (Word fragment completion)	0	0.5	10	0 (rest+lexical decision test)	0	-.38
73	1997 Torrance-Perks, J. (Study 2)	15	2 (Word fragment completion task)	0	0.5	10	0 (rest+lexical decision test)	1	-.41
74	1997 Torrance-Perks, J. (Study 3)	15	2 (Word fragment completion)	1	0.5	10	0 (rest+lexical decision test)	0	.14
75	1997 Torrance-Perks, J. (Study 4)	15	2 (Word fragment completion)	1	0.5	10	0 (rest+lexical decision test)	1	.07
76	1997 Torrance-Perks, J. (Study 5)	8	1 (Candle Problem)	0	1	8	2 (memory test: cues presented as one of the stimulus)	2	.17
77	1997 Torrance-Perks, J. (Study 6)	8	1 (Candle Problem)	0	1	8	2 (memory test)	0	-.71
78	1997 Torrance-Perks, J. (Study 7)	8	1 (Radiation Problem)	0	5	8	1 (read story : analogy to the problem)	1	.41
79	1997 Torrance-Perks, J. (Study 8)	8	1 (Radiation Problem)	0	5	8	1 (read unrelated story)	0	.00

Incubation and problem-solving

80							2 (analogy : have the		
	1997	Torrance-Perks, J. (Study 9)	8	2 (Remote Associates Tasks)	0	1	8	same solution as the unsolved RAT)	.15
81									
	1997	Torrance-Perks, J. (Study 10)	7	2 (Remote Associates Tasks)	0	1	8	2 (neutral analogy)	.14
82									
	1997	Torrance-Perks, J. (Study 11)	7	1 (Candle Problem)	1	1	8	2 (memory test :cue presented a one of the stimulus)	.56
83									
	1997	Torrance-Perks, J. (Study 12)	7	1 (Candle Problem)	1	1	8	2 (memory test)	-.33
84									
	1997	Torrance-Perks, J. (Study 13)	7	1 (Radiation Problem)	1	5	8	1 (read story :analogy to the problem)	.52
85									
	1997	Torrance-Perks, J. (Study 14)	7	1 (Radiation Problem)	1	5	8	1 (read unrelated story)	.00
86									
	1997	Torrance-Perks, J. (Study 15)	7	2 (Remote Associates Tasks)	1	1	8	2 (analogy: have the same solution as the unsolved RAT)	.17
87									
	1997	Torrance-Perks, J. (Study 16)	7	2 (Remote Associates Task)	1	1	8	2 (neutral analogy)	.28
88									
	1998	Hansberry, M. T.(Study 1)	32	2 (Riddle)	0	1	15	2 (Remote associates task)	.66

Incubation and problem-solving

89	1998	Hansberry, M. T.(Study 2)	20	2 (Remote Associates Task)	0	.5	10	2 (Remote associates task)	0	.04
90	1998	Hansberry, M. T.(Study 2)	20	2 (Remote Associates Task)	1	.5	10	2 (Remote associates task)	0	.28
91	1999	Jamieson, B. A. (Study 1)	52	2 (Remote Associates Task)	0	.33	5	2 (maths problem)	0	.00
92	1999	Jamieson, B. A. (Study 2)	52	2 (Remote Associates Task)	0	.33	5	2 (maths problem)	0	.10
93	1999	Henley, R. J. (Study 3.2)*	48	2 (Anagram)	0	.25	1440	1 (Free activity)	0	0
94	1999	Henley, R. J. (Study 4)*	26	2 (Anagram)	0	.93	1440	1 (Free activity)	0	0
95	2002	Dodds , R., Smith, S. M., & Ward, T. B. (Study 1)	45	2 (Remote Associates Test)	0	10	.5	2 (Insight Problem + Make a word task)	2	.31
96	2002	Dodds , R., Smith, S. M., & Ward, T. B. (Study 2)	45	2 (Remote Associates Test)	0	10	.5	2 (Insight Problem + Make a word task)	1	-.14
97	2002	Dodds , R., Smith, S. M., & Ward, T. B. (Study 3)	42	2 (Remote Associates Test)	0	10	.5	2 (Insight Problem + Make a word task)	0	.10
98	2002	Dodds , R., Smith, S. M., & Ward, T. B. (Study 4)	70	2 (Remote Associates Test)	0	10	.5	2 (Insight Problem + Drawing test)	0	.00
99	2002	Medd, E., Houtz, J. C.	15	1(Creative Writing)	0	10	10	2(Writing)	1	1.05

Incubation and problem-solving

100	2002	Medd, E., Houtz, J. C.	15	1 (Creative Writing)	0	10	10	2(Writing)	0	0
101	2003	Seabrook, R., & Dienes, Z*	60	2 (Anagram)	0	6	7	2 (Word generation)	1	.74
102	2004	Both, L., Needham, D., & Wood, E.	98	2 (Anagram)	0	1.67	6	2 (Search letter and Answer questionnaire)	0	.48
103	2004	Both, L., Needham, D., & Wood, E.	82	2 (Anagram)	0	1.67	6	2 (Search letter and Answer questionnaire)	0	.09
104	2004	Penney, C., Godsell, A., Scott, A., & Balsom, R.	9	2 (Anagram)	0	5.75	30	1 (Free Activity)	1	.74
105	2004	Penney, C., Godsell, A., Scott, A., & Balsom, R	9	2 (Anagram)	0	5.75	120	1 (Free Activity)	1	1.05
106	2004	Penney, C., Godsell, A., Scott, A., & Balsom, R	9	2 (Anagram)	0	5.75	30	1 (Free Activity)	0	.62
107	2004	Penney, C., Godsell, A., Scott, A., & Balsom, R	9	2 (Anagram)	0	5.75	120	1 (Free Activity)	0	.70
108	2004	Segal, E. (Study 1)	20	1 (Insightful mathematic problem)	0	20	4	1 (Read paper)	0	1.17
109	2004	Segal, E. (Study 2)	21	1 (Insightful mathematic problem)	0	20	12	1 (Read paper)	0	1.09
110	2004	Segal, E. (Study 3)	20	1 (Insightful mathematic problem)	0	20	4	2 (Word puzzle)	0	.90

Incubation and problem-solving

111	2004	Segal, E. (Study 4)	23	1 (Insightful mathematic problem)	0	20	12	2 (Word puzzle)	0	.57
112	2007	Vul, E. & Pashler,H. (Study 1)	25	2 (Remote Associates Test)	0	1	5	2 (video game)	0	.86
113	2007	Vul, E. & Pashler,H. (Study 2)	25	2 (Remote Associates Test)	1	1	5	2 (video game)	0	-.14
114	2007	Vul, E. & Pashler,H. (Study 3)	14	2 (Anagram)	0	1	5	2 (video game)	0	-.17
115	2007	Vul, E. & Pashler,H. (Study 4)	14	2 (Anagram)	0	1	5	2 (video game)	0	-.35
116	2007	Vul, E. & Pashler,H. (Study 5)	14	2 (Anagram)	0	1	5	2 (video game)	0	.36
117	2007	Vul, E. & Pashler,H. (Study 6)	14	2 (Anagram)	0	1	5	2 (video game)	0	.25

* Study with a within-subject design: Individual was in both incubation and control conditions, the weighting formula would not apply to it, hence, it would be excluded from the regression analyses.

Appendix C

The regression model describing the interaction effect of Nature of Problem and Incubation Task on the weighted unbiased effect size estimate, included the multiplicative terms between Nature of Problem (In, Li, Di) and Incubation Task (H, L, I) as predictors. The regression model is specified as follows:

$$Y_j = a_j X_j + b_1 * Li_j * H_j + b_2 * Li_j * L_j + b_3 * Li_j * R_j + b_4 * V_j * H_j + b_5 * V_j * L_j + b_6 * V_j * R_j + b_7 * C_j * H_j + b_8 * C_j * L_j + b_9 * C_j * R_j + k_j$$

where Y_j is the weighted unbiased effect size estimation of the study j , and X_j is a vector of other categorical and explanatory variables (Misleading Cues, Cue, and Preparation Period, Incubation Period) of that study. a_j is the corresponding vector of coefficients. b_1 to b_9 are the coefficients of the multiplicative terms, and k_j is the error term. The dummies I_j and Di are eliminated through substituting equations $Li_j + In_j + Di_j = 1$ and $I_j + L_j + H_j = 1$ into the model. The transformed regression model is as follows:

$$Y_j = a_j X_j + (b_3 - b_9) * Li_j + (b_6 - b_9) * V_j + (b_7 - b_9) * H_j + (b_8 - b_9) * L_j + (b_1 - b_7 - b_3 + b_9) * Li_j * H_j + (b_2 - b_8 - b_3 + b_9) * Li_j * L_j + (b_4 - b_7 - b_6 + b_9) * In_j * H_j + (b_5 - b_8 - b_6 + b_9) * V_j * L_j.$$

Regression analysis was carried out to find out the coefficient of the each variable in the transformed model. To examine the interaction effect of Nature of the Problem and Incubation Task, the difference between coefficients' multiplicative terms in the original model was compared. For example, the coefficient of the compound dummies $V_j * R_j$ and $C_j * R_j$ was compared to check if an incubation period filled with low cognitive tasks would improve the performance on Visual problems more than on divergent thinking tasks. A comparison between the original and the transformed model indicated that the coefficient of each variable in the transformed model was actually the combination of the coefficients in the original

model. The coefficients' difference between the compound dummies in the original model could be found by re-interpreting the coefficient in the transformed model. Table C1 represents a list of coefficient differences between compound dummies in the original model and the equivalent combination of the coefficients in the transformed model

Table C1

The Regression Coefficients in the Original and Transformed Regression Models

The coefficient difference between compound dummies in the original regression model	The combination of coefficients in the regression analysis
$Li_j * H_j - Li_j * R_j$	$Li_j * H_j + H_j$
$Li_j * L_j - Li_j * R_j$	$Li_j * L_j + L_j$
$V_j * H_j - V_j * R_j$	$V_j * H_j + H_j$
$V_j * L_j - V_j * R_j$	$V_j * L_j + L_j$
$C_j * H_j - C_j * R_j$	H_j
$C_j * L_j - C_j * R_j$	L_j
$V_j * H_j - C_j * H_j$	$V_j * H_j + V_j$
$Li_j * H_j - C_j * H_j$	$Li_j * H_j + Li_j$
$V_j * L_j - C_j * L_j$	$V_j * L_j + V_j$
$Li_j * L_j - C_j * L_j$	$Li_j * L_j + Li_j$
$V_j * R_j - C_j * R_j$	V_j
$Li_j * R_j - C_j * R_j$	Li_j

Appendix D

The third regression model describing the interaction effect of Nature of Problem and Incubation Task on the weighted unbiased effect size estimate, included the multiplicative terms between Nature of Problem (V, Li) and Incubation Task (H, L, R) as predictors. The regression model is specified as follows:

$$Y_j = a_j X_j + b_1 * Li_j * H_j + b_2 * Li_j * L_j + b_3 * Li_j * R_j + b_4 * V_j * H_j + b_5 * V_j * L_j + b_6 * V_j * R_j + k_j$$

where Y_j is the weighted unbiased effect size estimation of the study j , and X_j is a vector of other categorical and explanatory variables (Misleading Cues, Cue, and Preparation Period, Incubation Period) of that study. a_j is the corresponding vector of coefficients. b_1 to b_9 are the coefficients of the multiplicative terms, and k_j is the error term. The dummy Li and I_j were eliminated through substituting equations $Li_j + V_j = 1$ and $R_j + L_j + H_j = 1$ into the model. The transformed regression model is as follows:

$$Y_j = a_j X_j + (b_6 - b_3) * V_j + (b_1 - b_3) * H_j + (b_2 - b_3) * L_j + (b_3 - b_1 - b_6 + b_4) * V_j * H_j + (b_5 - b_2 - b_6 + b_3) * V_j * L_j.$$

Regression analysis was carried out to find out the coefficient of the each variable in the transformed model. To examine the interaction effect of Nature of the Problem and Incubation Task, the difference between coefficients' multiplicative terms in the original model was compared. For example, the coefficient of the compound dummies $V_j * R_j$ and $Li_j * R_j$ was compared to check if an incubation period filled with low cognitive tasks would improve the performance on Visual problems more than on linguistic insight tasks. A comparison between the original and the transformed model indicated that the coefficient of each variable in the transformed model was actually the combination of the coefficients in the original model. The coefficients' difference between the compound dummies in the original model could be found by

re-interpreting the coefficient in the transformed model. Table D1 represents a list of coefficient differences between compound dummies in the original model and the equivalent combination of the coefficients in the transformed model.

Table D1

The Regression Coefficients in the Original and Transformed Regression Models

The coefficient difference between compound dummies in the original regression model	The combination of coefficients in the regression analysis
$Li_j * H_j - Li_j * R_j$	H_j
$Li_j * L_j - Li_j * R_j$	L_j
$Li_j * L_j - Li_j * H_j$	$L_j - H_j$
$V_j * H_j - V_j * R_j$	$V_j * H_j + H_j$
$V_j * L_j - V_j * R_j$	$V_j * L_j + L_j$
$V_j * L_j - V_j * H_j$	$V_j * L_j + L_j - V_j * H_j - H_j$
$Li_j * H_j - V_j * H_j$	$V_j * H_j + V_j$
$Li_j * L_j - V_j * L_j$	$V_j * L_j + V_j$
$Li_j * R_j - In_j * R_j$	V_j