A A R H U S   U N I V E R S I T E T

SOLIDUM PETIT IN PROFUNDIS · UNIVERSITAS ARHUSIENSIS ·

CREATIVE COGNITION:

ANALOGY AND INCUBATION

Bo T. Christensen

Department of Psychology, University of Aarhus, April 2005

Creative Cognition: Analogy and Incubation

PSYKOLOGISK PH.D.-SKRIFTSERIE – 2005:1
Psykologisk Institut
Aarhus Universitet
Nobelparken
Jens Chr. Skous Vej 4
8000 Århus C.
Tlf.: 89424900

ÅRHUS 2005
1. udgave
Tryk: Fællestrykkeriet for Sundhedsvidenskab, Aarhus Universitet
ISSN 1660-800

A dissertation submitted in partial fulfillment of the requirements for the degree of
Doctor of Philosophy in Psychology, Faculty of Social Sciences, University of Aarhus, Denmark

Supervisor:
Professor, dr. phil., Jens Mammen, University of Aarhus, Denmark

Opponents:
Associate professor, Ph.D., Dorthe Berntsen, University of Aarhus, Denmark
Dr. Linden Ball, Lancaster University, UK
Professor, PhD, Ken Friedman, Norwegian School of Management, Norway

Table of Contents

…Acknowledgements

Thanks are due to the people who helped me in completing this dissertation, by reading and commenting on some of the articles (Jens Mammen, Dorthe Berntsen, and Christian D. Schunn), and assisting me in the data coding process (Kristoffer Riis Pedersen). Also many thanks to the (anonymous) company who graciously allowed me access to their impressive product development department. Equal thanks to Herluf Trolle for putting me in contact with the company.

A special thank you to Christian D. Schunn from LRDC at the University of Pittsburgh, who encouraged me to come to Pittsburgh for an inspiring 8 months, supervised me both during the stay and subsequently upon my return to Denmark, and co-wrote two of the articles in this dissertation. Thank you for teaching me hands-on in vivo research - your ideas and feedback have been positive contributions to the present dissertation.

This dissertation is dedicated to Nete for her loving support, and for not taking me too seriously.

Introduction

Where do creative products come from? We hear of some of them every week in the news and they leave us pondering. Just going through last weeks papers brings us such diverse creations as the discovery that some of the offspring of certain plants with mutated DNA apparently fix their own mutated genes, implying that some organisms may contain a cryptic backup copy of their genome that bypasses the usual mechanisms of heredity. How did they discover this, the scientists at Purdue? We also learn of the Danish carpenter who invented the stand-up floor-hammer, because he was sick of crawling around on his knees, hammering floorboards together. What went through his mind? – and did necessity play it's part? And what about U2 and their new album 'How to Dismantle an Atomic Bomb' – what inspired them?

We hear about such novel and useful creations every week. They change our world and our understanding of it, and make us wonder what muse, inspiration, insight or other mysterious mechanism delivered the basic idea. Creativity researchers have spent the past hundred years looking for those mechanisms. For a large part, researchers have focused on properties of the personality of creative people, such as traits, personality types or intelligence. But other types of explanations have also been sought, by looking at the characteristics of the creative process, the ingredients in creative environments, and by examining the properties of the creative products themselves.

In early creativity research cognitive examinations of mechanisms and processes were somewhat rare in comparison to the above mentioned research areas. But several cornerstones to the cognitive creative landscape were laid early on, such as coining the terms for four stages in the creative process as preparation-incubation-illumination-verification (Wallas, 1926), examinations of fixation (Maier, 1931; Duncker, 1945), synthesis or ‘bisociation’ (Koestler, 1964), and distinguishing between the generative and the evaluative phases of creativity, as in Osborn’s (1963) brainstorming technique. These and other cornerstones in cognitive creativity research were, however, somewhat disjoint research areas. It was not until the 1990’s that Finke, Ward and Smith (1992) coined the term ‘creative cognition’, to delimit an area of study within cognitive science that would focus on creative mechanisms and processes, thereby bringing the above mentioned cornerstones together, along with a lot of other cognitive research. ‘Creative cognition’ studies processes and mechanisms that are important and perhaps somewhat particular to creativity, such as analogical transfer, imagination, incubation, and fixation. But creative cognition also studies general processes and mechanisms that play an important part not only in creativity but also many other endeavors of human activity (such as short and long-term memory, mechanisms of categorization, and spreading activation) – as long as the focus remains on creative activity. The present dissertation concerns several aspects of creative cognition, notably (but not limited to) analogical transfer and incubation effects. It consists of five articles (see table 1). Below I will introduce the analogy and incubation effect concepts and domains, and present some of the most important findings in the five articles, along with their limitations and possibilities for future research. This introduction will provide both introductory and concluding remarks on the 5 articles, and will not be limited to mere summarization since both consequences of and extended argumentation to the articles will be presented. As such this introduction may be read both prior to and subsequent to reading the 5 articles.

Table 1. The five articles

Article 1: ’Problematic assumptions in incubation effect studies, and what to do about them’. A theoretical review of the incubation effect literature.

Article 2: ’Spontaneous access and analogical incubation effects’, co-written with Christian D. Schunn. A lab experiment combining an incubation effect study with analogical transfer, enabling the study of both spontaneous access and analogical mapping.

Article 3: ’A methodology for studying design cognition in the real world’. An article on how in vivo methodology can be applied to study design cognition.

Article 4: ’The relationship of analogical distance to analogical function and pre-inventive structure: The case of engineering design’, co-written with Christian D. Schunn. An in vivo study of analogy in real-world engineering design.

Article 5: ’How can participation in decision making enhance creativity in work groups?’, co-written with Thomas Jønsson. A theoretical synthesis of different explanations of the relationship between creativity and participation in decision making, grounding these explanations in cognitive search theories of problem solving and creativity.

Analogy and Incubation

Analogy involves accessing and transferring elements from familiar categories (often referred to as the ’source’ analogue) to use it in constructing a novel idea (’target’), e.g., in an attempt to solve a problem or explain a concept (e.g., Gentner, 1998). In its most general sense, analogy is the ability to think about relational patterns (Holyoak, Gentner, & Kokinov, 2001). Following Gick & Holyoak’s (1980; 1983) now classic studies of analogy using Duncker’s (1945) ‘radiation’ problem, analogy has grown to be one of the most important research domains in cognitive science. Gick & Holyoak’s version of the radiation problem goes as follows:

> “Suppose you are a doctor faced with a patient who has a malignant tumor in his stomach. It is impossible to operate on the patient, but unless the tumor is destroyed the patient will die. There is a kind of ray that can be used to destroy the tumor. If the rays reach the tumor all at once at a sufficiently high intensity, the tumor will be destroyed. Unfortunately, at this intensity the healthy tissue that the rays pass through on the way to the tumor will also be destroyed. At lower intensities the rays are harmless to healthy tissue, but they will not affect the tumor either. What type of procedure might be used to destroy the tumor with the rays, and at the same time avoid destroying the healthy tissue?” (Gick & Holyoak, 1980, p. 307-8)

To help their subjects solve this problem, Gick & Holyoak would show the subjects an analogous source problem, such as the following:

> “A small country fell under the iron rule of a dictator. The dictator ruled the country from a strong fortress. The fortress was situated in the middle of the country, surrounded by

> farms and villages. Many roads radiated outward from the fortress like spokes on a wheel. A great general arose who raised a large army at the border and vowed to capture the fortress and free the country of the dictator. The general knew that if his entire army could attack the fortress at once it could be captured. His troops were poised at the head of one of the roads leading to the fortress, ready to attack. However, a spy brought the general a disturbing report. The ruthless dictator had planted mines on each of the roads. The mines were set so that small bodies of men could pass over them safely, since the dictator needed to be able to move troops and workers to and from the fortress. However, any large force would detonate the mines. Not only would this blow up the road and render it impassable, but the dictator would then destroy many villages in retaliation. A full-scale direct attack on the fortress therefore appeared impossible. The general, however, was undaunted. He divided his army up into small groups and dispatched each group to the head of a different road. When all was ready he gave the signal, and each group charged down a different road. All of the small groups passed safely over the mines, and the army then attacked the fortress in full strength. In this way, the general was able to capture the fortress and overthrow the dictator." (Gick et al., 1980, p. 351)

The similarity between the radiation problem and the 'general' problem may seem obvious, and the solution to the radiation problem therefore straightforward: The doctor should use multiple low-intensity rays pointing from several different directions, but intersecting on the tumor, to avoid damaging healthy tissue, but destroy the tumor. However, this solution was less straight-forward to Gick & Holyoak's subjects since only few subjects noticed the structural similarity, and spontaneously transferred the solution from the 'general' problem to the radiation problem, unless instructed to attempt to use the 'general' problem. This is somewhat surprising given the fact that the similarity between source and target seems obvious once you have made the connection between the two. But making a spontaneous connection between the 'general' problem and the 'radiation' problem is apparently a relatively rare occurrence. This lack of spontaneous access in analogical problem solving has been repeated many times since (see Anolli, Antonietti, Crisafulli, & Cantoia, 2001 for a brief review). Gick & Holyoak's somewhat surprising finding sparked off a lot of research, trying to explain why subjects had not accessed the source spontaneously. Parts of the answer has been found to concern the importance of instructions, and distinguishing different kinds of similarity (superficial vs. structural similarity) (e.g., Holyoak & Koh, 1987; Gentner, Rattermann, & Forbus, 1993), and article 2 and 4 in this dissertation will go into more detail about these and other possible explanations, and the theories and empirical findings of the analogy literature.

The typical analogical problem solving study asks subjects to first read a story (the 'source') presenting both a problem and its solution, and later (in an apparently unrelated event) asking the subject to solve an analogous problem ('target') to see if the subject then accesses the previous source problem, and transfers solution elements to the target. However, in their original study, Gick & Holyoak (1980) also included a control condition in their experiment V where they postponed the presentation of the 'source' until *after* the subjects had begun working on the radiation problem. This condition was compared to what they called an 'incubation control' condition, making clear the parallels between this target-source condition and studies focusing on incubation effects. This 'reverse' analogical problem solving condition has only been used very few times in the analogy literature, and dominant theories do not predict that the directionality issue (i.e., whether source precedes or follows target) would have any influence on analogical transfer.

But in this dissertation, the ‘reverse’ analogical problem solving direction (with ‘source’ being displayed to the subject *after* having worked on the ‘target’) is important since it shows how analogical problem solving may be one kind of incubation effect. Incubation effects refer to the possibility that setting a problem aside temporarily may help creative problem solving performance (see article 1 for discussion about the definition of incubation effects). Historically it has been related to the idea that unconscious mental processes are working on solutions to creative problems while the problem solver is not consciously thinking of the problem. This was for example the sense in which Wallas (1926) used the term. However, in modern uses of the term, incubation effects primarily refer to the fact that sometimes leaving a problem aside for a while may improve performance, compared to working on the problem continuously. This modern use opens up for a multitude of potential explanations of incubation effects, including benefits from hints or cues received during time away from active problem solving. Unlike analogy, incubation effect research has remained a relatively small research domain, with only a small number of studies conducted. Studies have had problems finding reliable incubation effects (Olton, 1979), although a recent review (Dodds, Ward, & Smith, in press) argued that recent studies show more promise in proving the existence of incubation effects.

The typical experimental setup testing for incubation effects first asks subjects to work on a creative problem for a short while. Then the subject is interrupted and asked to perform a different task. And finally the subject is asked to return to the problem and have another attempt at solving the problem. An incubation effect is said to occur when performance in the condition with time away from problem solving is better following return to problem solving, compared to subjects working continuously (see article 1 for more details). To calculate this, a so-called resolution score is calculated for each group, using the formula:

$$\frac{\text{Total number of problems solved – number of problems solved in pre-incubation period}}{\text{Total number of problems - number of problems solved in pre-incubation period}}$$

The ‘reverse’ analogical problem solving condition (where ‘source’ is presented after initial work on the ‘target’ problem) can be considered a kind of incubation effect study. Article 2 is an experimental study testing for these ‘analogical incubation effects’ on insight problems. This is interesting considering the problems researchers have had finding evidence for both spontaneous access in analogical problem solving, and incubation effects. Further, article 4 was an attempt at showing whether spontaneous analogizing occurs in real-world engineering design, along with the types and functions of such analogies. Both these studies involved and required some methodological improvements over previous studies, and the following section will attempt to clarify some of the methodological problems facing the researcher interested in incubation and analogy, and the solutions that were attempted in the present studies.

Spontaneous Analogizing

Most of the analogical access and transfer studies conducted have been conducted in the psychological lab, using lone novice subjects working on simple tasks requiring little or no beforehand knowledge or expertise. One finding from this kind of research is that subjects are generally better at accessing and transferring from source to target when they are instructed to use (or verbal hints are provided) the source in solving the target problem (Anolli et al. 2001; Gick & Holyoak, 1980; 1983; see also article 2 and 4). However, this finding that subjects who are informed about the link between source and target access the source more frequently than subjects who are uninformed of this link does not say much about the frequency or efficiency of spontaneous access, other than that it is less frequent and efficient than instructing subjects. In so far as the ambition is to generalize these findings to real-world analogizing, it is important to examine spontaneous analogizing in a different way than by simply comparing it to instructing subjects. Rather, spontaneous analogizing could perhaps better be studied by looking at the typical (uncued and uninstructed) problem-solving behavior of subjects, in order to estimate whether the link between source and target is noticed spontaneously more frequently when in fact there is structural similarity between them, as compared to cases where no such structural link exists. Further, it is also important to actually study spontaneous analogizing in real-world settings (i.e., conduct studies with noticeably more ecological validity than the typical analogy study), to see if results obtained in the psychological lab will generalize.

But in order for the analogy researcher to conduct these kinds of studies, several methodological difficulties have to be overcome. For one thing, it is important that the subject is in fact uninformed about the helpful link that could exist between source and target (Ross, Ryan, & Tenpenny, 1989). When measuring spontaneous access in a within-subject design involving several analogous problems it is thus necessary for the researcher to demonstrate that subjects did not become aware of the relevance (i.e., 'catch on', thus becoming informed) after they had accessed some of the relevant sources. This result led Ross et al. (1989) to conclude that in order to avoid having to deal with an unknown mixture of informed and uninformed cases, access issues need to be studied with informed subjects. However, as argued, in some cases the 'uninformed' subjects are more interesting than 'informed' subjects in so far as the purpose of the research is to examine whether subjects spontaneously will pick up analogous cues to unsolved creative problems from the environment, as some interactive incubation effect researchers have hypothesized (Seifert, Meyer, Davidson, Patalano, & Yaniv, 1995). In experimental setups it is thus important that subjects do not 'catch on' to the purpose or structure of the experiment prematurely, thereby suspecting that they are supposed to transfer information from source to target. Further, it is important that in within-subject designs using several tasks, some estimate of whether a subject has ‘caught-on’ (i.e., has become informed) can be made.

In real-world observational studies catching-on is of course less of a worry, since no experimental manipulation of what constitutes an analogous source is included. Here subjects produce or recall their own ‘sources’, and it is this production or recall that is being studied.

In the present dissertation, I conducted two different studies tapping into the issue of spontaneous analogizing, using two very different methodologies. The first study (article 2) was an experimental study of spontaneous access and analogical incubation effects, and the second study (article 4) was a real-world observational study of a team of engineering designers. Below I will outline the methodological progress involved in these studies, along with the findings from these studies that are relevant to the literature on spontaneous analogizing.

An Experimental Study of Spontaneous Analogizing

In order to examine spontaneous access in an experimental setting, we decided to conduct a time-course analysis of problem solving behavior (see article 2). Rather than basing the examination of spontaneous access on performance measures (as is the typical case) where right-wrong answers are taken as proof of usage of source information, we used video recordings of subjects while they were solving insight problems. Subjects were run individually, and first given a cover-story that was constructed to ensure that the subjects did not initially expect a relation between source and target. The instructions read as follows:

The main task:

You are the editor of ‘Puzzle Magazine’, and one of your tasks is to try out all puzzles prior to publishing. Today you have eight puzzles to solve – but only 45 minutes to work on them. Since the puzzles are pretty hard, this means you have to work fast to get them all right, and really try to make the most of your time.

If you get stuck on a puzzle for more than a couple of minutes, you should move on to another puzzle, to make the most use of your time. You can always return later. During the 45 minutes you may attempt to solve the puzzles in any order you like.

During the main task of puzzle solving:

During the 45 minutes you have another task to do as well. Your task is to rate how difficult some other problems are. From time to time your secretary will drop a puzzle (with the answer included) on your desk, and you should immediately read through the puzzle, and it’s answer and make a judgment about how difficult the problem is. This is done on a scale from 1 (very easy) to 5 (very hard). Once you have made this judgment, you hand the puzzle back to the secretary, and continue to work on your main task.

Subjects were then given a booklet with eight pages (one insight problem per page), and started working on the problems. Every 5 minutes the experimenter would place the secondary task on their desk for immediate rating (either an analogous source, or a non-analogous distracter problem; each subject received 4 of each). By placing a video-camera above the subjects, and directing it down on the paper in front of them, it was possible to discern which problem was being worked on, when they moved to other problems (i.e., turned the pages in the problem booklet), and when they attempted to provide an answer for a problem (i.e., wrote an answer in the answer section for each problem). The experimenter would follow the problem solving progress in real-time, and keep track of which problems had been worked on and left. This allowed the experimenter to constantly update a pool of available cues and distracter tasks, one of which was randomly selected every 5 minutes to be handed to the subject. This procedure also ensured that the subjects did not receive source problems for the problem they were currently working on (i.e., they had to actively turn pages in the booklet in order to get to the correct target problem). By ensuring that the subjects had already worked on and left the target problem when the source was given to them, motivation for returning was in place. From the video recordings of the page-turning and problem-solving behavior of the subjects it was possible to code several different constructs (usually measured separately) in the same design. Notably, it was possible to distinguish analogical access from the page-turning behavior of the subject, and at the same time measure analogical transfer using standard performance measures (correct/wrong answer). Further, of relevance to the incubation literature, it was possible to code for impasse (i.e., when a subject left a problem without having produced a correct answer), returns (i.e., when subjects turn back to a problem they had previously attempted to solve), and incubation time (i.e., time away from each individual problem, while working on other problems). Unlike standard incubation effect research, the subjects could work their way through the eight problems without being forced to stop working on a particular problem or return to a particular problem at times not selected by themselves. Rather, subjects controlled which problems they wanted to work on at what time.

From this experimental setup, immediate access was coded when subjects turned the page in the booklet to the target problem as the very next problem following the rating of the relevant source problem. This was compared to a baseline derived from typical page turning behavior following the distracter tasks. Examinations of this typical page turning behavior revealed that two strategies existed: prior to seeing all problems in the booklet, subjects would simply flip to the next unseen page in the booklet. After seeing all problems in the booklet, subjects would primarily return to problems where they had not yet indicated an answer to the problem. This information was used to make a baseline against which the immediate access rates could be estimated. The results indicated that subjects returned immediately significantly more often than their typical page turning behavior predicted.

This result does not, however, control for whether subjects caught-on to the fact that some of the rating tasks were helpful cues during the experiment. In fact, an analysis of the access rates on the 1^{st}, 2^{nd}, 3^{rd}, and 4^{th} cue revealed that a catching-on effect was present. But by eliminating all problems where the subject could possibly have caught-on, it was possible to show that even for this reduced set, access rates still significantly exceeded baseline.

In conclusion, this somewhat complex experimental setup and calculation method showed that the rate of spontaneous access significantly superseded a baseline of the subjects' typical page turning behavior on the distracter tasks. As such, spontaneous access was shown to be present for uninformed subjects. This methodology constitutes a significant improvement over standard measures of spontaneous access using performance measures and the typical use of 'instructed' conditions as comparison groups. The present results indicated that there were both an effect of becoming informed ('catching on'), and a spontaneous access effect in this study.

This methodology and these results could be utilized in future studies were it is necessary to distinguish between access and mapping, or code for self-generated impasse, incubation time or return. For example, future researchers may want to directly compare the directionality issue in analogical problem solving (i.e., do access rates differ between a source-target condition and a target-source condition)? Although dominant theories do not predict any difference between these conditions, the incubation theory of Opportunistic Assimilation does predict such a difference. Further, it is possible to develop this methodology into a computerized version, where all 'page-turning' behavior and timing is automatically registered by the computer and where presentation of cues is also automated. This would significantly reduce the labor-intensiveness of the methodology, reducing the workload of the experimenter, and thereby paving the way for large-scale studies of analogical access and transfer.

Real-world Study of Spontaneous Analogizing

To examine spontaneous analogizing outside the psychological lab, we decided to examine how analogy was used in a product development and inventive domain. Engineering design was chosen as a suitable domain. A company with a history of producing creative and quality design in medical plastics was contacted, and we were allowed to study a particular design project that was about to start. The company, project, team structure, participants, and stage in the design process are all described in some detail in article 3.

Kevin Dunbar's (1995; 1997; 2000; Dunbar & Blanchette, 2001) in vivo methodology was used in collecting data and coding the transcribed data. Basically, in vivo research entails selecting a suitable object of study in the real-world, collecting data from these objects of study using video or audio recordings, transcribing the recorded data, and coding the data according to suitable coding schemes. Dunbar (1995; 1997; 2001) has used this methodology to examine analogy in the domain of science (in particular microbiology), where he found that a suitable object of study for the study of scientific thinking and reasoning was the weekly lab meetings. An analogous object of study in engineering design turned out to be subgroup product development meetings, where a team of designers would engage in a broad cross-section of design activities. These meetings were thus recorded, the data transcribed, segmented and coded to inform the analogy literature. The coding scheme was constructed to examine the amount of spontaneous analogizing, the analogical distance involved (within-domain vs. between-domain), the influence of external representations (in this case sketches and prototypes) on analogical thinking, and the functions that the analogies served (hypotheses and coding scheme is described in detail in article 4). The results indicated that spontaneous analogizing occurs frequently in engineering design. Roughly equal amounts of within domain analogies and between domains analogies occurred. References to prototypes appeared to reduce the amount of between domain analogizing, compared to referencing sketches and ideas without external support. Support was found for the hypothesis that analogy served three functions in the early stages of engineering design; problem identification (primarily within domain), explanation (primarily between domain), and problem solving (a mixture of within and between domain).

Future research using the present methodology and/or data set points in several directions. (1) First, more data should be collected from different design teams, and areas of design, to replicate the present findings on analogy and establish whether the present results will generalize beyond the design team studied. This was, however, beyond the scope of the present dissertation. (2) The engineering design data set can be used to code for other creative cognition phenomena involving thinking and reasoning, such as mental model runs (see article 3 for a preliminary coding), restructuring, aesthethical comments, incubation effects and so on. Some of these coding schemes are currently in preparation. A few remarks on coding for incubation related phenomena in the real world will be made later in this introduction.

Conclusion on Spontaneous Access

Analogy research on spontaneous access conducted in the psychological lab has found spontaneous access to be a relatively rare occurrence. However, the two studies related to spontaneous analogizing in the present dissertation showed evidence of a significant spontaneous access rates in the psychological lab when using typical problem solving behavior as a baseline, and frequent use of analogizing in real-world engineering design. These results indicate that perhaps spontaneous analogizing is not as infrequent as some theorists (e.g., Anolli et al., 2001) have argued, and show that spontaneous access can be found both in the psychological lab and in real-world settings. However, it should also be noted that access rates do increase significantly with instruction to use the source, hints to use the source, or (in experimental settings) when the subjects catch on to the relation between source and target. But the present line of research shows that the existence of this effect of becoming informed should not lead us to assume that spontaneous access does not occur or occur infrequently. Rather, it should merely tell us that spontaneous access rates are somewhat lower than the high levels of access rates with informed subjects. In applied settings requiring creativity, these results may be utilized to exploit the creative potential of analogies.

Incubation Effects

I would like to start this incubation effect section with a short narrative about why I chose to concentrate much of my PhD efforts on incubation effects. This narrative may explain some of the motivation behind my studies, and why the incubation domain initially constituted an irritating element in my creativity landscape, and will help frame conclusions about how my understanding of the incubation domain has changed. My interest in incubation started when I was writing my master's thesis (Christensen, 2002, see also article 5), which argued for an ecological approach to the study of creative cognition. In my masters thesis I attempted to show how much of the creative problem solving literature focused on purely intra-psychical processes and explanations of creativity, and I attempted to direct this research towards studies of interactions (primarily focusing on the concept 'search') between subject and object in the creative process. 'Object' was conceptualized broadly as real-world events and entities, including their possibilities and also their impossibilities. I constructed a framework for the understanding of creativity called 'the creative cycle' modeled after and extending Neisser's (1976) perceptual cycle. A key argument was that rather than viewing creativity as subjective constructive processes (such as 'insights' arriving seemingly from nowhere), creative cognitive studies should instead focus on subject-object interactive processes to be able to follow how something comes to be something else in creative development. This subject-object interaction study could reveal how objects do not come into being ex nihilo (meaning 'out of nothing'), even though that is the dictionary definition as Boden (1990) and Perkins (1988) have pointed out. Rather, the creative process is a search in the possibilities and impossibilities of the real-world. I thought I had a pretty persuasive argument. But I found that the incubation literature seemed to be saying otherwise. First, most of the theories about the cause of incubation effects seemed to point to purely autonomous (intra-psychical) mechanisms (e.g., Wallas, 1926; Simon, 1966; fatigue dissipation - Woodworth & Schlosberg, 1954; Segal, 2004), or to the potential constraining effect of the environment (Smith, 1995). These theories seemed to be saying that either the environment didn't make any difference, or, if it did, the influence was a negative impact on creativity. However, a few theories also pointed to the potential beneficial effect of environmental cues (e.g., Seifert et al., 1995; Langley & Jones, 1988). But far worse than these theoretical controversies was the fact that experimental evidence of incubation effects seemed to suggest that environmental cuing did not produce incubation effects. An early study by Dreistadt (1969) did find evidence for the positive impact of visual analogies during incubation on insight problems, but Olton and Johnson (1976) failed to replicate, and Browne and Cruse (1988) using one of Dreistadt's insight problems found incubation effects for the experimental group receiving cues in only one of two experiments. These and other disappointing incubation effect studies lead Olton (1979) to question the existence of incubation effects.

Similarly disappointing and confusing results have been found on remote associates tasks (RAT), where the cues are presented as associates of the answer. Mednick et al. (1964) found positive effects of associates as cues; Dorfman (1990) found mixed results; and in the largest incubation study to date (a total of 855 subjects in 3 experiments), Dodds et al. (2002) found that associates as cues produced either insignificant results, or in some cases the unrelated cues condition performed significantly *better* than the condition receiving associates of the answer (i.e., in the opposite direction than expected). Furthermore, in another condition subjects received the actual *answer* to the RAT problem during incubation, but in no case did subjects receiving the answer perform better than subjects receiving unrelated cues unless the participants were explicitly instructed that hints or answers would be presented during incubation. Dodds et al. (2002) concluded that no incubation effect was observed when helpful cues were presented during the incubation period. The possible implications of these studies seemed counter intuitive and unbelievable to me: Could it be that incubation effects simply do not occur as a result of receiving cues or even 'answers' from the environment? So I set out to see if I could identify at least one task or domain where reliable incubation effects caused by interaction with the environment could be found.

This narrative sets up the understanding I had of incubation upon going into this PhD project, and in the following sections, I will reference this understanding to illustrate how it has changed in the past three years about incubation and creativity. In the present dissertation, article 1 and 2 deals extensively with incubation effects.

Problematic Assumptions in Incubation Effect Studies

As a precursor to the experimental study in this dissertation, I originally intended to write a review of the experimental incubation effect literature. But as it turned out, such a review was recently written, and still 'in press' (Dodds et al., in press). Instead I decided to conduct a meta-analysis of the experimental incubation effect studies, since the Dodds et al. (in press) review was mainly a qualitative review.

I collected every study claiming it concerned incubation effects that I could find, producing, I believe, a near complete list of studies explicitly concerning incubation effects. As I went through each of these studies in detail, calculating effect sizes, I also made note of important difference in design. Theoretically important differences were noted, so that studies could be classified according to these potential 'moderator variables' following a possible test of heterogeneity. A test of heterogeneity examines whether studies can reasonably be described as sharing a common effect size (Hall, Tickle-Degnen, Rosenthal, & Mosteller, 1994), and a possible rejection of such a test can lead to an examination of whether moderator variables are at work. The number of potential moderator variables quickly grew, starting with task type (open/closed, ill-defined/well-defined, simple/complex), length of pre-incubation time, length of incubation period, length of post-incubation time, task during incubation (relaxation, work on other problems, work on the same type of problem), ability level, were subjects interrupted or allowed to reach an impasse?, were subjects informed that they would be returning to the initial problem following incubation? But the list did not end here as more moderating variables appeared, including did the subjects receive fixating elements in the pre-incubation period?, did they encounter helpful cues (visual analogies, associates of the answer, or the actual answer) during incubation?, if so, were they told that they would be encountering such helpful cues?, and notably what was the dependent variable (resolution score, flexibility, fluency, originality).

As the number of theoretically important moderating variables simply grew and grew, I became increasingly concerned that the incubation effect domain could be violating some of the underlying assumptions in a meta-analysis, notably that the studies should basically be measuring the same thing (note this assumption also holds for qualitative reviews). This may be estimated empirically in tests of homogeneity in meta-analysis, but this must not preclude the researcher from making a theoretical judgment about whether a given set of studies concerns the same thing. Theoretically the important differences in experimental design seemed problematic and made me question whether they were in fact measuring the same thing. Further, a number of other problematic assumptions seemed to dominate the incubation line of research, such as the understanding that there should be a close link between creativity and incubation effects – an understanding that had never been tested, and also seemed theoretically and empirically problematic. In the end, I decided against writing both a qualitative review and meta-analysis, and instead turned to writing a theoretical article on the problematic assumptions underlying incubation research (see article 1).

Some of the problematic assumptions in incubation effect studies are listed below. The consensual operational definition of incubation effects points to a performance test comparing two groups, only one of which gets a break or interruption from problem solving. But this definition can be criticized for being too inclusive. For example, even though I thought I had collected a near complete list of incubation effect studies, it turned out that if the definition is taken seriously, I had merely collected a near complete list of studies *claiming* to be measuring incubation effects. Many more experimental studies/phenomena in cognitive psychology can be claimed to fall under the criteria of the operational definition, including such diverse areas as the tip-of-the-tongue-phenomenon, impasse learning, hypermnesia, reminiscence, and personal maturing. But on the other hand, the operational definition can also be argued to be too exclusive, in that it excludes studies not specifically focusing on performance, such as studies focusing on when subjects reach impasses and return to unsolved problems. This may suggest that the definition of incubation effects may be off the mark in relation to the incubation phenomenon, but this does not make incubation effect studies uninteresting. Incubation effects seemingly violate expectations from research on working memory and forgetting, in that this line of research would predict that interruptions and later resumption of a task would lead to lower levels of performance compared to groups working continuously. Another problematic assumption is that incubation effects are closely related to creativity. The link has never been tried and tested, and seem problematic since the tasks involved in creative and non-creative designs often appear more similar to each other, than do the different tasks said to be measuring creativity. Because of the number and nature of the moderating variables existing between experimental incubation effect studies, it seems problematic to assume that they are all measuring the same thing. For example, it seems problematic to generalize from studies including cues or fixating elements in the environment during incubation to studies not including these key variables. Further, these experimental design issues mirror key theoretically loaded issues in the theories of incubation. Whereas some theories view the incubation effect as an essentially beneficial effect to problem solving, others imagine it to be a temporary detrimental effect. Whereas some theories view incubation effects as autonomous effects, other theories view it an interaction between subject and environment. What this suggests is that we may be dealing with multiple different paths to the same incubation effects, or we may be dealing with different kinds of incubation effects. Or we may be dealing with both. In any case it seems problematic to assume that theories of incubation effects are essentially competing theories.

Analogical Incubation Effects and the Role of Analogical Distance

The last point was an important lesson learned; that there is no need to treat autonomous and interactive incubation effects as competing theories. Quite the contrary seems warranted. This point can be further extended to concern theories arguing that environmental cues are the cause of incubation effects. Whereas my initial interest in incubation effect research was the apparent lack of experimental evidence for the helpful effect of cuing during incubation, it now seems clear that there are multiple ways of cuing and the effect of cuing will depend on a multitude of factors, not least type of cue and task. Further, not all 'environmental cues' will be beneficial to a particular creative task – the environment is not always kind - and some environmental cues will have quite the opposite effect; they will be fixating or detrimental to problem solving. These fixating and beneficial effects need not be competing theories.

However, in the specific case of making analogies based on external cues in creative problem solving, the two studies in the present dissertation may seem to point in two – perhaps contrasting - directions. In article two, it is shown how between-domain analogies are spontaneously accessed and used to help solve creative problems. The experimental design was setup so that subjects were allowed to reach a self-generated impasse (rather than being interrupted as is usually the case), and return at their own choosing (rather than be forced to return at certain time as is usually the case).

The results were very encouraging with relation to proving that at least some forms of external cuing can lead to incubation effects (contrary to the above mentioned previously mixed results in the incubation literature). But in article four, it is shown how in real-world engineering design the accessibility of external exemplars in the form of within domain prototypes may lead to less between-domain analogizing (and thus, perhaps, less original products). Do these two results mean that making analogies from environmental sources in creative problem solving is beneficial or fixating? In a sense, both of these studies document the benefits of making analogies to environmental sources in creative problem solving – in both cases analogies help solve creative problems. However, in the case of making analogies with within domain prototypes available in real world engineering design (article four) it was found that within domain prototypes lead to fewer between domain analogies (suggesting that, perhaps, the creative solutions would be less original than if no prototypes had been present). Thus, the two different effect measures may point in two different directions; whereas analogizing to external sources may always be beneficial to *providing a solution* to a creative problem, analogical distance may modify the *originality* of the solution. In other words, both within domain and between domain analogies may be beneficial to solving creative problems (effect measure: 'solve/no solve'), but between domain analogies may lead to more original solutions than within domain analogies (effect measure: 'degree of originality in solution'). This means that the 'fixating' effect of having within domain sources available (in memory or the environment), is only fixating with respect to degree of originality, but can be counted as beneficial with respect to actually solving creative problems! This is certainly a possible hypothesis to be drawn from the present results. However, the present dissertation did not directly provide a test of this hypothesis. For example, no between-domain external sources were included as a condition in the real-world study of analogizing in engineering design (article 4), and therefore we do not know whether providing between-domain external sources will lead to more between domain analogies – and more original products. Future research should test the hypothesized relationship between the analogical distance of available external exemplars in relation to the two effect measures (‘solve/no-solve’ and ‘degree of originality of solution’) directly.

The Possibility of Studying Incubation in the Real-world

The desirability and possibility of studying incubation in the real-world is explicitly mentioned in article 1 and 2. Such an in vivo study of incubation, using the real-world design data collected (article 3 and 4), is currently in preparation. However, an in vivo study of incubation has to overcome both conceptual and methodological problems. An outline of a coding scheme for this kind of study is described below, along with some of the conceptual and methodological problems that have to be overcome.

In reviewing the problematic assumptions in experimental incubation effect studies (article 1) it was found that the operational definition of incubation effects both captured many diverse studies not explicitly measuring incubation effects, as well as excluded studies not explicitly focusing on performance, when comparing to anecdotes of the incubation phenomenon. Further, it was shown that theories trying to explain incubation effects point to both autonomous and interactive mechanisms, and beneficial and temporary detrimental effects. These and other findings make it a rather open question where to look for incubation in the real world.

To take one example of a conceptual difficulty in studying incubation in the real world, it seems problematic to even operate with the idea that a ‘control group should be working continuously’. In the real-world, no control group exists since a person who has reached an impasse and left an unsolved problem is – by definition – incubating. The person may or may not return to the unsolved problem later, but irregardless the person is still ‘incubating’. Therefore it seems problematic to operate with ‘a control group working continuously’ since all unsolved creative problems by definition ends in an ‘incubation period’. However, since article 1 explicitly criticize the exclusive criterion of performance enhancement in the operational definition used in experimental incubation research, and recommend rejecting the standard operational definition it would be interesting to drop this kind of control group, and instead begin to look at other aspects of the incubation phenomenon.

Rather than focusing on performance differences between receiving and not receiving breaks or interruptions, perhaps real-world incubation research could instead focus on the concepts of self-generated impasses and returns. This could be done by identifying a priori creative and unsolved problems, and then record the creative process as it occurs ‘live’ and ‘online’, using in vivo methodology (see article 3). If each segment of transcribed data is coded for what creative problem(s) are addressed, this would allow the researcher to study when creative problems are left, and returned to. Subsequent analysis of the process and in particular the periods leading up to returns to a problem may reveal if the returns are spontaneous or planned, and perhaps thrown light on what caused the spontaneous returns.

Although this may sound simple, several coding problems have to be resolved. First of all, what exactly constitutes an a priori creative problem? Design problems, for example, can be both implicit and explicit, and coding implicit design problems requires extensive knowledge of the domain. Interviewing members of the team prior to recording the creative process will help bring out explicit design problems, but the implicit ones will still be a problem, and how should they be handled? Further, design problems are not static, but develop, in that existing problems become specified or restructured, and new problems emerge. So, a flexible taxonomy of design problems is required – allowing design problems to evolve over time. The taxonomy also need to be able to handle the fact that design problems are organized in levels (some problems are super ordinate compared to other more specific problems), and that some problems (even at the same level) may be more closely related than others. The reason the relatedness issue is important is that jumps between design problems may be somewhat ‘large’ or ‘small’ depending on the degree of relatedness between the problems (analogous to analogical distance – see article 4). Another problem is whether to code for solutions to problems (thus removing a problem from the taxonomy). In real world design, it remains a fuzzy boundary between having and not having ‘solved’ a problem since multiple possible solutions can be present at a given time, making the solution process as much a matter of ‘selecting’ a solution as actually ‘finding’ it. Yet another problem concerns what grain size to select for segments; in engineering design ‘complete thought’ may be too small since solving design problems appears to be done in episodes, but coding in episodes reduces the number of segments (and thus reducing statistical power), whereby large quantities of data are required, perhaps making the research workload huge.

If these methodological issues can be overcome, then such a coding may tell us something about the reasons and frequency of planned and spontaneous returns to impassed creative problems in the real-world. But it will not tell us much about whether putting a problem aside will help solve the problem, compared to not setting it aside, because this coding would not be focused on performance.

I am currently working on resolving these methodological problems, in the hope that coding for frequency and reasons for planned and spontaneous returns to unsolved creative problems will inform the literature on incubation.

Participation in Decision Making and Team-based Creative Cognition

In a creative cognition topic diverging from a narrow focus on analogy and incubation, article five presents an attempt to theorize about why participation in decision making may enhance creativity in work groups. The article was written for an anthology in preparation on participation in decision making. Participation in decision making is inherently an inter-individual research topic, and therefore the article draws heavily on organizational and social psychology, notably on team-based creativity and innovation.

The article deals with distinct lines of explanations for the possible relationship between participation in decision making and creativity, and attempts to bridge these explanations, and setup a coherent theoretical explanation rooted in key concepts from cognitive search theories of problem solving and creativity psychology, which are more or less in line with current research. The explanation links participation in decision making to creativity through the moderating variables of freedom for creative search, diversity, integration, and commitment. It is hypothesized that these moderating variables each influence specific creative outcome variables. This theoretical explanation, although more or less in line with present research, is somewhat speculative, and needs to be tested directly in experimental studies.

The Five Articles and the Road Ahead

The present introduction has attempted to describe some of the themes and threads of the articles in this dissertation such as spontaneous access in analogical thinking and incubation effects. But notably it has also attempted to show how these five articles make small contributions and incremental steps on the road to better understanding these themes and threads. The empirical results are novel attempts at combining analogy and incubation experimentally, and studying analogy in engineering design in the real-world, but they are perhaps limited by small and restricted samples, and the results should be replicated using more subjects/data, and other design domains to test the generalizability of the present results. But the merits of the present dissertation may be seen as not only contributing some empirical results, but perhaps more importantly contributing methodological and theoretical incremental steps forward for creative cognition studies. In this introduction I have tried to show how the present results, methods, and theoretical developments point towards new research studies.

For one thing, the in vivo methodology championed by Dunbar (e.g., 1995; 2001) should be applied in other real-world creative cognition studies, in design and other creative domains. It holds promise to provide increased ecological validity to creative cognition studies. Two such future codings were explicitly mentioned; studying the relation between information uncertainty and mental model runs (see article 3), and studying frequency and cause of planned and spontaneous returning to unsolved creative problems. Experimentally, it remains important to directly test the hypothesized relationships between participation in decision making and creativity (article 5). The novel time-course analysis methodology used in article two also shows promise of further informing the analogy and incubation literature (especially since it allows for a coding of impasse, returns and performance in the same design). This methodology could potentially be computerized to allow for greater sample sizes.

With this outline of a next few possible steps on the road ahead, it is my hope that the present dissertation will help direct creative cognitive studies along fruitful research paths.

References

Anolli, L., Antonietti, A., Crisafulli, L., & Cantoia, M. (2001). Accessing source information in analogical problem-solving. *Quarterly Journal of Experimental Psychology: Human Experimental Psychology, 54A*, 237-261.

Boden, M. A. (1990). *The creative mind: Myths and mechanisms*. New York: Basic Books.

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

Christensen, B. T. (2002). *The creative process and reality. An analysis of search and cognition in the creative process and a call for an ecological cognitive framework for creativity.* (vols. 5, no. 3 - Psykologisk Studieskriftserie) Aarhus, Denmark: Department of Psychology, University of Aarhus.

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

Dodds, R. A., Ward, T. B., & Smith, S. M. (in press). A review of the experimental literature on incubation in problem solving and creativity. In M.A.Runco (Ed.), *Creativity research handbook* (3rd ed.), Cresskill, NJ: Hampton Press.

Dorfman, J. (1990). *Metacognitions and incubation effects in insight problem solving.* Unpublished doctoral dissertation. University of California, San Diego.

Dreistadt, R. (1969). The use of analogies and incubation in obtaining insights in creative problem solving. *Journal of Psychology, 71,* 159-175.

Dunbar, K. (1995). How scientists really reason: Scientific reasoning in real-world laboratories. In R.J.Sternberg & J. E. Davidson (Eds.), *The nature of insight.* (pp. 365-395). Cambridge, MA, US: The MIT Press.

Dunbar, K. (1997). How scientists think: On-line creativity and conceptual change in science. In T.B.Ward, S. M. Smith, & J. Vaid (Eds.), *Creative thought: An investigation of conceptual structures and processes.* (pp. 461-493). Washington, DC, US: American Psychological Association.

Dunbar, K. (2000). How scientists think in the real world: Implications for science education. *Journal of Applied Developmental Psychology, 21,* 49-58.

Dunbar, K. (2001). The analogical paradox: Why analogy is so easy in naturalistic settings yet so difficult in the psychological laboratory. In D.Gentner, K. J. Holyoak, & B. N. Kokinov (Eds.), *The analogical mind: Perspectives from cognitive science.* (pp. 313-334). Cambridge, MA: The MIT Press.

Dunbar, K. & Blanchette, I. (2001). The invivo/invitro approach to cognition: the case of analogy. *Trends in Cognitive Sciences, 5,* 334-339.

Duncker, K. (1945). *On Problem-solving.* Westport, CT, US: Greenwood Press, Publ.

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

Gentner, D. (1998). Analogy. In W.Bechtel & G. Graham (Eds.), *A companion to cognitive science* (pp. 107-113). Malden, MA, USA: Blackwell Publ.

Gentner, D., Rattermann, M. J., & Forbus, K. D. (1993). The roles of similarity in transfer: Separating retrievability from inferential soundness. *Cognitive Psychology, 25:* 524-575.

Gick, M. L. & Holyoak, K. J. (1980). Analogical problem solving. *Cognitive Psychology, 12,* 306-355.

Gick, M. L. & Holyoak, K. J. (1983). Schema induction and analogical transfer. *Cognitive Psychology, 15,* 1-38.

Hall, J. A., Tickle-Degnen, L., Rosenthal, R., & Mosteller, F. (1994). Hypotheses and problems in research synthesis. In H.Cooper & L. V. Hedges (Eds.), *The handbook of research synthesis* (pp. 17-28). New York: Russel Sage Foundation.

Holyoak, K. J., Gentner, D., & Kokinov, B. N. (2001). The place of analogy in cognition. In D.Gentner, K. J. Holyoak, & B. N. Kokinov (Eds.), *The analogical mind: Perspectives from cognitive science* (pp. 1-20). Cambridge, MA, USA: MIT press.

Holyoak, K. J. & Koh, K. (1987). Surface and structural similarity in analogical transfer. *Memory & Cognition, 15,* 332-340.

Koestler, A. (1964). *The act of creation*. London: Hutchinson.

Langley, P. & Jones, R. (1988). A computational model of scientific insight. In R.J.Sternberg (Ed.), *The nature of creativity: Contemporary psychological perspectives.* (pp. 177-201). New York: Cambridge University Press.

Maier, N. R. F. (1931). Reasoning in humans II, The solution of a problem and its appearance in consciousness. *The Journal of Comparative Psychology, 8,* 181-194.

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

Neisser, U. (1976). *Cognition and reality*. San Francisco, CA, US: W. H. Freeman and Co.

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

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

Osborn, A. F. (1963). *Applied imagination*. (3rd revised ed.) New York: Charles Scribner's Sons.

Perkins, D. N. (1988). The possibility of invention. In R.J.Sternberg (Ed.), *The nature of creativity: Contemporary psychological perspectives.* (pp. 362-385). New York, NY, US: Cambridge University Press.

Ross, B. H., Ryan, W. J., & Tenpenny, P. L. (1989). The access of relevant information for solving problems. *Memory & Cognition, 17,* 639-651.

Segal, E. (2004). Incubation in insight problem solving. *Creativity Research Journal, 16,* 141-148.

Seifert, C. M., Meyer, D. E., Davidson, N., Patalano, A. L., & Yaniv, I. (1995). Demystification of cognitive insight: Opportunistic assimilation and the prepared-mind perspective. In R.J.Sternberg & J. E. Davidson (Eds.), *The nature of insight.* (pp. 65-124). Cambridge, MA: The MIT Press.

Simon, H. A. (1966). Scientific discovery and the psychology of problem solving. In R.G.Colodny (Ed.), *Mind and Cosmos* (pp. 22-41). Pittsburgh, PA: 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. 135-156). Cambridge, MA: The MIT Press.

Wallas, G. (1926). *The art of thought*. London, England: Jonathan Cape.

Woodworth, R. S. & Schlosberg, H. (1954). *Experimental psychology*. (Revised ed.) New York: Holt, Rinehart and Winston.

Article 1

Problematic Assumptions in Incubation Effect Studies and What to do about Them

Christensen, B. T. (submitted)

Abstract

The standard incubation effect research paradigm has been operating under a number of problematic implicit assumptions. It is argued that the operational definition of incubation effects is off the mark when attempting to generalize to the incubation phenomenon as seen in anecdotal evidence, since it includes studies not typically associated with incubation and excludes issues not dealing with performance. Incubation effects may not have a single common cause, but rather multiple, leaving non-competing theories. Incubation effects may have no special relationship to creativity, but concern problem solving in general. Despite these problems, standard incubation effect studies may still be interesting in that they violate expectations from working memory and forgetting research. It is recommended that future incubation researchers reject the standard operational definition and instead pursue direct tests of the theories of incubation, perhaps focusing on ecologically valid studies.

Acknowledgements

I have greatly appreciated the comments and suggestions of Christian D. Schunn and Dorthe Berntsen on an earlier version of this article.

Introduction

In the about 68 years since the first published study of incubation effects (Patrick, 1937), a relatively small number of incubation effect studies have been published. A recent review counted 39 such studies, and concluded that the global impression was that incubation effects represents a real phenomenon worthy of considerable additional study (Dodds, Ward, & Smith, in press). There are, however, a number of problematic assumptions underlying experimental studies of the incubation effect and the incubation construct, which seem to dominate the incubation line of creativity research. The present article will examine a number of these problematic assumptions, and make recommendations as to how they should be handled in future research on incubation. The present article takes the perspective of the integrative reviewer hoping to guide future research on incubation in fruitful directions.

Although similar concepts had been used previously in the literature (e.g., von Helmholtz, 1896), the concept of incubation was popularized by Wallas (1926) who made a landmark contribution to the understanding of the creative process by coining terms for four stages in the process: preparation, incubation, illumination, verification. He conceptualized incubation as a stage in the creative process that follows hard initial preparatory work, but preceding insight. Wallas suggested that active unconscious processing of the problem took place during the incubation stage. The term ‘incubation’ is a biological metaphor suggesting that these unconscious processes ‘mature’ the solution, suggesting it evolves gradually. Much like the incubation of an egg ends with hatching a chicken, incubation of ideas ends in the ‘hatching’ of the solution in a sudden and surprising insight. The extraordinary leap in problem solving, from seeming conscious inactivity to sudden and surprising insight, is so impressive that it seems to suggest that the act of setting the problem aside somehow facilitated the problem solving (Olton, 1979). And it is this alleged performance enhancing feature of setting the problem aside that incubation researchers have focussed on.

Anecdotal Accounts

Wallas (1926) derived his four stages of the creative process from anecdotes by highly creative scientists, artists, and inventors producing creative contributions to their domain. Anecdotal accounts of grand creators making breakthrough inventions, science, and art after spending time away from the problem are frequently quoted in the creativity literature. The face validity of these phenomenological descriptions, where long periods of intellectual drought sometimes ends in breakthroughs, may be what led many a creativity and cognitive researcher to pursue rigid experimental tests of incubation.

One of the most frequently cited anecdotal accounts are from Henri Poincaré:

> “Just at this time, I left Caen, where I was living, to go on a geologic excursion under the auspices of the School of Mines. The incidents of the travel made me forget my mathematical work. Having reached Coutances, we entered an omnibus to go some place or other. At the moment when I put my foot on the step, the idea came to me, without anything in my former thoughts seeming to have paved the way for it, that the transformations I had used to define the Fuchsian functions were identical with those of non-Euclidian geometry. I did not verify the idea; I should not have had the time, as, upon taking my seat in the omnibus, I went on with a conversation already commenced, but I felt a perfect certainty. On my return to Caen, for conscience’ sake, I verified the result at my leisure.” (Poincaré quoted in Hadamard, 1945, p. 13)

This story has often been used to highlight that sometimes the creator is unable to identify ‘anything in their former thoughts’ that generated the sudden shift from incubation to insight. Other anecdotes illustrate that sometimes an external (e.g., Archimedies’ discovery of displacement while sinking into the bathtub, noting the water rise) or internal (e.g., Kekulé’s discovery of the molecular structure of Benzen, while dozing in front of the fireplace) event is identified as promoting insight. However, these anecdotes do not tell us much about the mechanisms behind incubation. As Boden (1990, p. 15) wrote; what these creators had been able to tell us of the mechanisms behind incubation and insight could best be summed up as ‘the bath, the bed, the bus’ (that is, they were sitting in the bath, laying in bed, or sitting on a bus when all of a sudden…). Based on the fact that there seem to be a rather small number of anecdotes in the incubation literature, Ohlsson (1992) has questioned whether we are in fact dealing with a large number of separate anecdotal accounts, or rather a small number of accounts repeated endlessly in the literature. It is true that the above three anecdotes, along with about a handful more (e.g., Darwin, Hadamard, Mozart, Gutenberg, Goethe and Wolfgang Köhlers primate ‘Sultan’) are standard references in the field. But many more examples of anecdotes and descriptions of characteristics of typical creative processes by the creators themselves can be found in works such as Rossman (1931), Ghiselin (1954), Platt and Baker (1931), Shepard (1978), Whyte (1973), and Gratzer (2002).

The existence of such anecdotal accounts should not, however, lead us to accept the existence of the incubation phenomenon. After-the-fact introspective reports of mental processes like these are notoriously unreliable (e.g., Perkins, 1981; Weisberg, 1993). The stories are excellent at illustrating the phenomenon you seek to explain, but they do not prove the existence of such a phenomenon, nor do they prove that particular theories of incubation are true or false. One of the reasons for their extended use in the literature seem to be that each story suggests links between the incubation phenomenon and specific theories thereof (for example, the Poincaré story suggests that unconscious mental processes are at work during time away from the problem). But when testing incubation, it is important to keep the phenomenon under scrutiny separate from theoretical accounts thereof (e.g., Woodworth & Schlosberg, 1954; Perkins, 1981; Kaplan, 1989; Olton, 1979).

Definitions of Incubation Effects

Experimental researchers trying to find evidence for the incubation phenomenon in the psychological laboratory have encountered the daunting task of having to operationally define incubation, in order that its existence could be tested along with theories of why it occurred. The anecdotal accounts share a number of elements and the operational definition needs to identify the necessary and sufficient criteria to include and exclude in incubation studies. Is it for example necessary for the subject to reach an impasse on the problem before setting the problem aside, or do interruptions work the same way? Is it necessary for the subject to return on his own (rather than be forced to return), and should this happen spontaneously and unexpectedly? Can the subject be consciously aware of what triggered a solution? What kind of problem should be used?, etc. These are all questions of possible necessary criteria for inclusion and exclusion in the operational definition of the incubation phenomenon.

Historically, the first published experimental study testing incubation defined it as “the spontaneous recurrence from time to time of an idea with more or less modification, while the subject is thinking of other topics” (Patrick, 1938, p. 81). This heavy emphasis on the spontaneous popping into mind was shown “if an idea occurred early in the report, recurred one or more times, the subject meanwhile talking of other things, and at last appeared as the chief topic of the experiment”(Patrick, 1938, p. 60). The operational definition of incubation later changed to “laying a problem aside as a step toward solution” (Woodworth et al., 1954, p. 838) which stresses the performance improving aspect of incubation. Later operational definitions on the incubation phenomenon have practically unanimously agreed that what needs to be explained is mainly (some seems to argue only) this performance improvement aspect of laying the problem aside, and it is now referred to as the ‘incubation effect’. Some later examples of this consensus can be seen in table 1.

Table 1: Examples of definitions of incubation effects.

Incubation of a creative performance may be defined as a stage of the process during which no active attempts at a solution are being made but which nevertheless results in improvements of performance. (Mednick et al., 1964; Bennett, 1975).

…that a person who has worked on a problem without success will be aided by engaging in an interpolated activity for a period of time, then returning to the problem (Dominowski et al., 1972).

Incubation refers to an increase in the likelihood of successfully solving a problem that results from a delay between the period of intense work which initiates the problem solution and another period of conscious effort which finalizes the solution (Posner, 1973; Dorfman, 1990; Davidson, 1995).

…a period of inactivity between active work on the problem and achievement of resolution during which progress is somehow made nonetheless (Perkins, 1995; Jamieson, 1999)

…incubation will be defined as any positive effect of an interruption on problem solving performance. (Kaplan, 1989)

It is a facilitation of thinking (not simply of recall) that is evident after an interval during which no conscious work was done on the task, assuming an earlier period of substantial conscious work. (Olton, 1979).

Various constraints have been linked to the operational definition, including that the task at hand could not merely be mental recall such as the Tip-Of-the-Tongue phenomenon (but rather had to involve some kind of novelty) (Olton, 1979); that incubation leads to illumination or insight (Smith & Blankenship, 1989); or that the subject needs to reach an impasse prior to incubation rather than being interrupted (Butler & Thomas, 1999). Many definitions include explicitly that the activity during incubation should be either inactivity, or at least no conscious work on the problem (e.g., Mednick, Mednick, & Mednick, 1964; Silveira, 1971; Olton, 1979; Kirkwood, 1984; Houtz & Frankel, 1992; Torrance-Perks, 1997; Noble, 1998; Butler et al., 1999; Jamieson, 1999), but there is not universal agreement on this point, since it is omitted from other definitions (e.g., Dominowski & Jenrick, 1972; Posner, 1973; Kaplan, 1989; Smith et al., 1989; Dorfman, 1990; Smith & Blankenship, 1991; Davidson, 1995; Hansberry, 1998; Dodds, 1999; Seabrook & Dienes, 2003). Olton (1979) distinguished incubation from what he called creative worrying, which is basically further conscious work done from time to time during the incubation period which is later forgotten. Only few definitions explicitly mention the criterion that the task has to be a creative one (e.g., Houtz et al., 1992; Mednick et al., 1964; Smith & Dodds, 1999), while the rest mainly refers to the ‘problem’ or problem solving. Some authors explicitly include in their operational definition that a control group should work continuously (Ohlsson, 1992).

Although the incubation metaphor implies that a gradual change is taking place, there is no mention of the necessity of such a gradual change in operational definitions of incubation effects.

Theories of Incubation Effects

As mentioned, the term ’incubation’ was derived from the theory that *unconscious processes* are operating during time away from the problem, gradually bringing a solution closer to completion, which is then ’delivered’ to consciousness in an insight or AHA!-experience.

Variations of this theory have been put forth many times in the literature, whether the mechanism is referred to as the unconscious (Carpenter, 1876; Wallas, 1926; Hadamard, 1945), the pre-conscious (Kris, 1952; Kubie, 1958), the non-conscious or off-conscious (Rugg, 1963), or the endocept (Arieti, 1976). For example, Wallas wrote:

> ”The Incubation stage covers two different things, of which the first is the negative fact that during Incubation we do not voluntarily or consciously think on a particular problem, and the second is the positive fact that a series of unconscious and involuntary (or foreconscious and forevoluntary) mental events may take place during the period.” (Wallas, 1926, p. 86)

Theories of unconscious processing that do not detail the mechanisms behind this unconscious processing are extremely hard – or maybe even impossible - to test experimentally, which is the reason why some authors have suggested that this theory should be put on hold until other more testable theories have been examined (Woodworth et al., 1954; Perkins, 1981). Campbell (1960) tried to explain the workings of this unconscious by analogy to the theory of natural selection, in stating that creativity was basically a process of blind variation and selective retention.

There are, however, other types of theories about incubation. The *fatigue dissipation* theory simply holds that the pre-impasse active problem solving is very demanding work, and spending time away allows you to become rested, and refreshed so that you will do better the next time you engage the problem (Woodworth et al., 1954; Helmholtz, 1896). In addition to merely relieving fatigue, *sleep* has been hypothesized to restructure new memory representations (Wagner, Gais, Haider, Verleger, & Born, 2004). Other embodied theories of incubation have hypothesized the importance of *arousal* in incubation (Ochse, 1990; Martindale, 1990). Martindale (1990) hypothesized that creativity requires remote associative thought (see e.g., Mednick, 1962), which is best accomplished by defocused attention and a low state of cortical arousal, since this enables a large number of mental representations to be activated simultaneously. Thus, a high state of cortical arousal (e.g., when actively trying to solve the problem; Luria, 1973) may hinder problem solving because remote associations are weeded out by a narrow focus of attention (see Olton, 1979 for similar ideas of defocused attention).

Several other theories in cognitive science have proposed causal explanations for incubation effects.

The idea that *facilitating cues* or information from the environment is the cause of incubation effects has been put forth a number of times in the literature. It is usually accompanied by a theory of cognitive preparedness to notice these cues, thus epitomizing Louis Pasteur's notion that 'chance favors the prepared mind' (quoted in Posner, 1973). Olton (1979) argued that a basic way of demonstrating the facilitating nature of cues could simply consist of telling a problem solver the answer to the problem during incubation! However, he argued, more interesting and realistic results could be obtained by varying the salience or 'obviousness' of the cues given in a series of research studies. As it happens, this is exactly the course of development that the analogical problem solving paradigm seem to have taken. The paradigm started off with difficulties in demonstrating spontaneous transfer on distant analogies sharing mainly structural similarity (e.g., Gick & Holyoak, 1980; Gick & Holyoak, 1983), but later found a higher degree of spontaneous transfer when more obvious superficial similarity was shared between source and target (e.g., Holyoak & Koh, 1987). Analogical transfer (Stagner & Karwoski, 1952; Dreistadt, 1969; Olton, 1979; Gick et al., 1980; Langley & Jones, 1988; Seifert, Meyer, Davidson, Patalano, & Yaniv, 1995) or environmental cues triggering retrieval of previously unretrieved relevant information or schemas (e.g., Yaniv & Meyer, 1987; Langley et al., 1988) have been proposed as a cause of incubation effects several times.

These prepared-mind theories have proposed various standard or non-standard cognitive mechanisms to explain why the environmental cues are noticed so that spontaneous transfer or retrieval can occur. These explanations include non-focal attention mechanisms (Olton, 1979), encoding of so-called –'failure-indices' in memory (Patalano & Seifert, 1994; Seifert et al., 1995), indexing and differential trace strength (Langley et al., 1988), and initial sub-threshold activation is raised above threshold by external cues (Yaniv et al., 1987).

Another important theory of incubation explains incubation effects with forgetting of fixating elements that prevented the subject from reaching a solution during initial solution attempts (Woodworth et al., 1954; Anderson, 1975; Smith, 1995a; Smith, 1995b; Segal, 2004). Smith (1995b) has argued that incubation effects are caused by competing responses to a problem which blocks (through interference patterns; e.g., Mensink & Raaijmakers, 1988) the solution from being reached. These blocks lead to fixation, or 'mental ruts'. Active work in trying to reach the solution only activates the blocker further. Following impasse, there is a shifting accessibility of fixating elements and target (solution), with the target information increasing in relative accessibility to the fixating elements, thereby increasing the probability that the solution is reached upon return to the problem. In other words, the fixating elements are forgotten upon returning to the problem. Smith (1995a) has also proposed that if the fixating elements were related to the context, contextual shifting (in addition to forgetting) during incubation may improve performance.

More theories are also related to forgetting: A simple theory is that incubation effects are due to more conscious work (so-called *conscious recurrence*) done from time to time on the problem during incubation, which is later forgotten (e.g., Woodworth et al., 1954). Simon (1966) hypothesized that incubation effects were due to *selective forgetting.* Using problem space theory, he hypothesized that during problem solving a goal tree is constantly being constructed and updated in short-term memory, at the same time as information about the problem environment is being familiarized and stored in long-term memory. If the problem solver leaves the problem for a while (following impasse), the goal-tree is quickly forgotten. At a possible later return to the problem the problem solver will start off with only a higher level goal (the finer details about the goal tree being forgotten), but with better problem environment familiarization. Using this information, the new problem solving session may thus take new routes in the problem space and potentially quickly solve the problem.

Experimental Studies of Incubation effects

A number of experimenters have attempted to either test these theories of incubation effects, or tried to establish that incubation effects exist at all. Since the first published experimental studies of incubation effects in creative problem solving (Patrick, 1937; 1938) only a small number of studies have been conducted. The latest review (Dodds et al., in press) counted 39 studies that explicitly tested incubation effects. In addition to these studies, some recent publications should be mentioned (Dodds, Smith, & Ward, 2002; Moss, 2002; Segal, 2004), along with recent unpublished doctoral dissertations (Davidson, 1995; Torrance-Perks, 1997; Hansberry, 1998; Noble, 1998; Jamieson, 1999), and conference proceedings (Seabrook et al., 2003).

The typical experimental lab design paradigm applied in such studies compares performance on creative problem solving between two groups, only one of which gets an ‘incubation’ break away from active problem solving, and then comparing their performance in the post-incubation period (calculated using a resolution score: (total number of problems solved – number of problems solved pre-incubation)/(total number of problems- number of problems solved pre-incubation). Insofar as the experimental group performs significantly better after the break than the control group, an incubation effect is said to occur. But, null results aside, a comparison of performance between groups may go in two directions, with either group having the higher mean score. One can find either an incubation effect (if the experimental group has a higher mean) or what could be called a ‘continuation effect’ (if the control group has a higher mean). The break may entail a pause from all active problem solving, but since this potentially allows the subject time to work consciously on the problem during the break (creating a bias in the experiment), frequently the subjects are asked to work on other problems during the break. These ‘other’ problems may be unrelated to the experimental materials, but in some repeated measures experiments the ‘incubation’ break is considered a break away from the *individual* experimental problem. In the latter case, the subject gets a break away from the individual experimental problem (by interleaving the problems in a problem solving sequence, forcing the subject to attempt problem solving and return to particular problems at certain times), but they do not receive a break from mental work per se, nor from active problem solving on the experimental problems in general. Such paradigms would seem to follow more or less directly from the operational definition of incubation effects. Depending on the focus of the particular study a large number of variations over this theme exist in the experimental incubation literature.

An early review of these experimental incubation studies by Olton (1979) questioned the existence of incubation effects based on the literature of that day, noting that ”…[it] is important to establish that our customary advice to ’use’ incubation as a way of facilitating thinking is not simply an invitation to share an illusion” (Olton, 1979, p. 20-21).

But although mixed results have been found in the literature, Dodds et al. (in press) concluded that recent progress warranted optimism for the existence of the incubation effect. Of the 39 studies included in the review, they counted 26 successful demonstrations of incubation effects, 10 failures to find an effect, and 3 studies finding incubation effects for only certain groups of subjects. They concluded that the global impression was that incubation effects represent a real phenomenon worthy of considerable additional study. Furthermore, they concluded that several variables appear to affect the incubation process, in particular length of incubation, preparatory activity, and clues given during incubation. Less influential variables included problem type; activities during the incubation period; and ability and gender.

In addition to the Dodds et al. (in press) list, several more theoretically important variables could be added that have implications for the experimental research. For example, an important variable seems to be whether the subjects were informed of the fact that they would be returning to the experimental problem (this variable has been particularly important in the related 'spontaneous analogical transfer' literature, e.g., Gick et al., 1980). It also seems theoretically important to distinguish between problem types, such as open vs. closed problems (e.g., Wakefield, 1989), and ill-defined vs. well-defined problems (Reitman, 1965). A further addition could be the theoretically important distinction between interrupting subjects vs. allowing them to reach a self-initiated impasse (Seifert et al., 1995; see also Anderson, 1975 for an unpublished study by Donald W. Taylor on this issue). Most incubation effect studies interrupt subjects (i.e., forcing them to perform a different task at a time chosen by the experimenter) rather than allowing them to reach a self-initiated impasse where they themselves have given up on reaching a solution at this time (as is the case in the anecdotal accounts). Jett & George (2003) have argued that four different kinds of interruptions exist: intrusions, breaks, distractions and discrepancies, each having potential negative and positive consequences for the person being interrupted in the workplace. It could easily be argued that incubation effect studies have examined '*intrusions*' (i.e., an unexpected encounter initiated by another person that interrupts the flow and continuity of an individual's work and brings that work to a temporary halt). Contrary to intrusions, anecdotal accounts of incubation describe 'breaks' (i.e., planned or spontaneous recesses from work on a task that interrupts the task's flow and continuity, but which entails anticipated or self-initiated time away from performing work). As such, it could be argued that incubation effect studies are examining the wrong kind of interruptions, if they wish to generalize to the incubation phenomenon. There are, thus, a rather large number of potentially important moderating variables that have been varied in the incubation effects literature.

Given the large number of potential moderating variables, an important question for a reviewer becomes whether it can be safely assumed that the sample of incubation effect studies is in fact representative of the universe of potential incubation effect studies to which one wishes to generalize (as delimited by the operational definition). If it is not representative, then the review will have a sampling problem, much like an experimental study can have sampling problems, potentially skewing results (Hedges, 1994). For a number of reasons it seems unlikely that the present sample is representative of the universe of potential incubation effect studies. First, only a tiny proportion (forty something) of the thousands of cells (counted as 2 (hints vs. no hint) x 2 (fixation vs. no fixation) x 2 (interruption vs. impasse) x …) of potential incubation effect studies delimited by the potential moderating variables listed above have received experimental study. Second, some of these cells are clearly over-represented, such as studies using RAT tasks (used in at least 11 studies). Third, as we will see later, certain types of studies have simply never been conducted despite the fact that they clearly fall under the operational definition. This calls for caution as to the conclusion that can be derived from any review or meta-analysis of the present sample of incubation effect studies.

It should be with this caution in mind that one reads the conclusions made by the Dodds et al. (in press) review, which stated that it would appear that on balance there is support for the existence of the incubation effect. But besides the sampling issue, one must carefully take a look at other potential validity problems that could threaten such a conclusion.

These could be validity problems with the individual incubation effect studies, such as the *garbage in – garbage out* problem (the quality of the individual experimental studies determines the quality of the review/meta analysis conclusions, e.g., that some incubation studies do not include a control group, or perform biased comparisons between groups). It could also be problems with the present sample of incubation effect studies; that is validity problems stemming from trying to generalize on the basis of the present group of studies. Examples are the *file drawer* problem (studies failing to find a significant effect in the expected direction are less likely to become published), the *apples and oranges* problem (different studies claiming to be measuring incubation effects may not be measuring the same thing), and reliability issues (such as the repeated failures to replicate studies finding incubation effects). And it could also be validity problems with the relations between incubation effect studies and the phenomenon they purport to be examining (e.g., does the experimental design and operational definition warrant a generalization to the incubation phenomenon, or are there important differences making this difficult or impossible?).

All these potential problems (which are not specific to incubation effects) should be examined before jumping to conclusions about the existence of incubation effects, based on the experimental literature. The Dodds et al. (in press) review did a good job of summarizing the 39 studies they had located in the literature, but they focussed less on discussing such validity problems. In the following, I will deal with five basic assumptions (involving validity problems) in the experimental incubation effect literature. The hope for the outcome is to develop theory, and to assist new and existing incubation effect researchers in asking the right kinds of questions concerning incubation and incubation effects.

Assumption 1: The Operational Definition Captures the Incubation Phenomenon

In outlining what needs to be explained in the anecdotal accounts of incubation, researchers have faced no easy task. Nevertheless, as we have seen, a consensual operational definition of the ’incubation effect’ arose early in the incubation literature, and has remained virtually unchanged since that time. However, there is a long way from the alleged complex phenomenon coining the incubation term, to this operational definition. What is left in the definition is simply a performance test comparing groups doing problem solving – one of which gets a break. Important constraints that seemed to delimit the original phenomenon have been dropped, including the reaching of an impasse, the insight that follows, and even the criterion validity issue that the task at hand should be a creative one! Olton (1979) also mentions constraints such as individual motivation, the length of the incubation periods, and notably the expertise of the subjects as potentially relevant criteria, which separates experimental studies of incubation from anecdotes thereof. Somewhere on the road from the real-world into the psychological lab it evolved into a definition somewhat far removed from the original observations. One can question whether this definition adequately succeeds in capturing the right kinds of studies - i.e., studies that examine incubation - while excluding studies that do not.

This is a question of the adequacy of the operational definition used in experimental studies, and consequently also a question of the ecological validity of incubation effect research. The definition succeeds in setting up clear parameters for experimental studies of incubation, but does this definition succeed in capturing all that is important about incubation, and does it succeed in excluding all that does not concern incubation?

One could argue that the definition is too inclusive with its broad focus on problem solving performance (thereby including studies about problem solving that does not concern creativity). But one could also argue that the definition is too exclusive, in that it excludes examination of characteristics of the alleged incubation phenomenon that does not concern performance (e.g., impasse, self-generated returns to the problem, attentional shifts to and from the problem, etc.).

If we take seriously the incubation effect definition and look at which studies would be included under such a definition, it becomes clear that studies labeling themselves ‘incubation effect studies’ are not the only ones studying incubation effects. Widely different research areas all display the characteristics sought in the incubation effect definition. Below I will mention 6 such areas.

These include (1) the Tip-of-the-Tongue (TOT) state (e.g., Schwartz, 2001; Schwartz & Smith, 1997; Jones & Langford, 1987), where subjects are temporarily unable to produce a word or name that they have a strong meta-cognitive feeling of knowing. The subjective feeling is that the word is right on the tip of the tongue. The state can frequently be resolved by leaving the problem for a while and returning to it later. Like incubation effects, the TOT phenomenon displays the temporal sequence of reaching an impasse that is resolved after a break away from the problem.

(2) Another research area which display the characteristics of incubation effects is impasse-driven learning events (VanLehn, 1988; 1989; 1990; 1991). Learning events are brief switches of attention away from executing familiar knowledge during problem solving to reasoning about the domain knowledge itself (e.g., learning a new principle or other piece of knowledge). Siegler & Jenkins (1989) coined the term for the brief episodes where children change strategies for counting on their fingers. In some situations, most learning events are triggered by impasses (characterized by long pauses or verbal signs of confusion) (VanLehn, 1991). Following the learning event, performance is improved (e.g., by utilizing another strategy). Again we see important characteristics are shared between impasse-driven learning events and incubation effects (impasse, pause, return, performance improvement).

(3) A memory phenomenon displaying incubation type effects is hypermnesia. The term hypermnesia is used to refer to improvements in net recall levels associated with increasing retention intervals (Payne, 1987). Subjects are presented with a set of to-be-remembered materials (such as pictures, words, poetry) and then, after various retention intervals, administered several recall tests. If subject performance improves across repeated tests, then hypermnesia is said to occur. According to a review of the literature, Payne (1987) concluded that hypermnesia is a reliable memory phenomenon, although it is more readily obtained with pictorial materials than

with verbal materials. It depends critically on the subject repeatedly attempting to retrieve the previously studied information. The experimental design used in hypermnesia studies resemble the standard incubation effect study design in many ways, with the comparison of performance in two production or problem solving sessions spaced in time being the essential component. As such, hypermnesia could be said to be an incubation effect. The similarity between incubation and hypermnesia has been noted in the literature (Mandler, 1994; Shaw, 1987).

(4) A precursor to present day hypermnesia research is the phenomenon of ’reminiscence’. Although originally defined by Ballard as ”the remembering again of the forgotten without re-learning” (1913, quoted in Payne, 1987, p. 5), it was broadened by some learning theorists to concern all conditioned reactions. Clark L. Hull (1943; see also Dorsel, 1979) described ‘reminiscence’ as follows.

> “In case a simple conditioned reaction is set up to an appreciable strength by massed practice and the final reinforcement is followed by a no-practice period several times as long as the interval between reinforcements, after which the stimulus is again delivered, motivation remaining constant, the reaction-evocation potentiality of this stimulus will be greater than it was at the termination of the original reinforcement sequence” (Hull, 1943, p296).

Reminiscence in this sense is basically a temporary depressing effect on performance following massed practice, which disappears with a long rest, as observed in human (e.g., rote learning) and animal behavior. Again we are dealing with an effect that displays the characteristics sought in the incubation effect definition.

(5) Some evidence from research on categorization suggests that asking subjects to list examples of categories will produce incubation effects. In this case, the type of incubation effect to be expected is fluency in production of examples to categories (analogous to open-ended creative problems). Although no direct test has been conducted to test for this incubation effect, two established lines of

research seem to converge on this suggestion. First, when subjects are asked to name examples rapidly to questions such as ’Name all the cities in the U.S. you can’, the first responses occur with relative ease, but subsequently the rate of production slows down exponentially, and eventually the subject shows evidence of effort in producing them (Bousfield & Sedgewick, 1944; Johnson, Johnson, & Mark, 1951). Second, categories display what is called graded structure (members of a category vary in how good an example/how typical they are of their category). This is the case for every category observed (Barsalou, 1987), including ad hoc categories (Barsalou, 1983). Graded structure makes it possible to predict the frequency with which people generate members of categories, with typical exemplars being generated more often than atypical exemplars. However, research has shown that the within-subject reliability of these typicality evaluations shows substantial instability over time (Barsalou, 1987), with reliability dropping to .80 after only a week. Notably, this also seem to be the case on exemplar production (within subject reliability .69 after a week; Bellezza, 1984) and free association (50% different responses after 60 days; Fox, 1970) for common taxonomic categories (such as ’lamp’). Further, different factors determine the graded structure of the same category in different contexts (Barsalou, 1985; Roth & Shoben, 1983). Thus, graded structure appears to be flexible, unstable and context dependent (Barsalou, 1987). When these two lines of research are combined, and seen in the light of incubation effects, we would expect that on open-ended production tasks using fluency as the effect measure the rate of production would decrease exponentially until the problem is left (impasse). Following an incubation period (perhaps involving context change), we would expect a changed graded structure (even for the same production task) which would elevate the rate of new exemplars produced (fluency) compared to a control group that had worked continuously. This would constitute an incubation effect.

(6) A final research area that displays incubation effect characteristics is what could broadly be

termed personal maturing (i.e., a developmental approach). In developmental psychology, when a child or adult fails to perform a task (whether it be a motor task, cognitive task or otherwise) at one age, but masters this task at a later age, development or maturation can be said to have occurred. Murphy (1947) had somewhat similar ideas about incubation, suggesting that accumulation of knowledge and skill over years or decades could be central. As it were, such maturation displays the characteristics sought in incubation effects; that initial failure followed by time away from a task leads to improvement in performance of the task.

When such widely disparate areas are grouped together by a definition, then you have an indication that the definition is too inclusive, and does not cover a single or coherent phenomenon, but rather groups a range of phenomena, where care should be taken before generalization across the group is made.

On the other hand, if the ambition is to scrutinize the alleged incubation phenomenon described in anecdotal accounts, then the incubation effect definition may be considered too exclusive in that it leaves out studies examining aspects not related to performance. Why, when and how do people reach an impasse on creative problems? Why, when and how do they return to unsolved problems? Are we dealing with a single return, or multiple? Is the return automatic or deliberate? These are all important questions that are currently not being asked in the incubation literature due to the heavy emphasis on performance which is closely linked to a particular experimental design.

It is true that anecdotal evidence would have us believe that incubated problems result in groundbreaking insights (i.e., excellent performance), but on the other hand such anecdotes only survive if they are interesting, and they would probably not be very interesting if the problem had never been solved. Therefore we should be careful about conclusions that incubation enhances performance based on anecdotes. Following this line of argumentation to a more radical conclusion, it could be argued that incubation has got nothing to do with performance increases at all! What if, for example, incubation is a mere stage in the creative process, but the presence (or length) of the stage does not predict performance increases (see figure 1)? This could be the case if a creative task type (such as ill-defined problems) predicted that a person would temporarily be unable to solve a problem, perhaps due to lack of information. In this case, the task type would predict the presence of a stage in problem solving where the subject took time away from the problem, but the presence of the incubation stage would not predict performance (i.e., the subject may never return to the problem, or upon return may not solve the problem). Arguing in this manner would make it just as viable to define incubation effects as 'an increase in the probability of setting a problem aside on creative problems' (task type is the independent variable; setting the problem aside is the dependent variable). For example, incubation effects could arise from problem types that lure the problem solver into traps, thus predicting impasses. Or one could define incubation effects as 'an increase in the probability of returning to a creative problem following a break' (e.g., task type or impasse as the independent variable; returning to a problem after setting it aside is the dependent variable). For example, the incubation effect could simply be the effect of returning more frequently to 'irritating problems'.

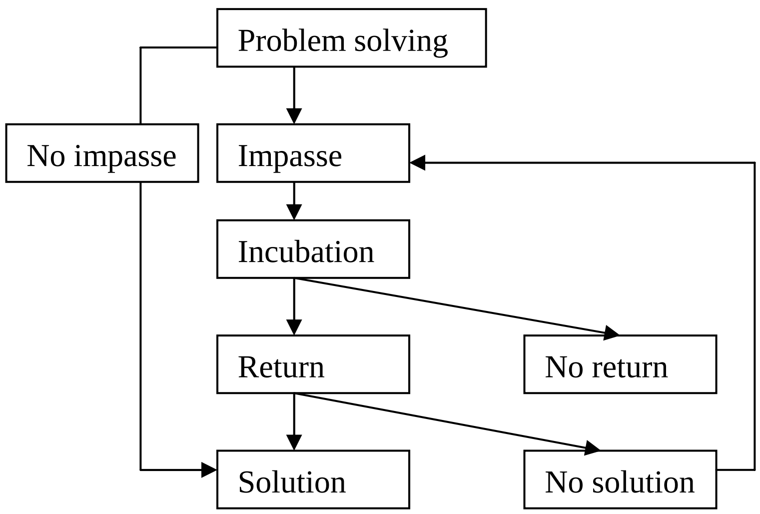


Figure 1. Some possible important incubation effect variables.

Thus, there appears to be a kind of conceptual gap between the alleged incubation phenomenon, and the operational definition of incubation effects. If incubation effects (defined consistent with the conventional operationalization) are supposed to capture the incubation phenomenon (as described by case studies and anecdotes), one could argue that the operational definition is both too inclusive, and too exclusive – or even that it may be off the mark, in so far as it focuses exclusively on performance. The conclusion I will draw here is that incubation researchers should not merely assume that incubation effect studies would tell us all there is to know about the alleged incubation phenomenon.

Assumption 2: Incubation Effects have a Single Cause

An important question concerning theories of incubation effects is whether such theories are in fact competing theories, or rather parallel or commensurable theories. The answer to this question has implications for both the research questions to pursue, and for the generalizations to be made from incubation effect studies.

Theories are competing if they differ in mutually exclusive ways on what causes incubation effects – this implies that the incubation effect phenomenon has a single cause, and that theories disagree on the nature of this cause. While some researchers have argued that theories of incubation effects need not be competing (e.g., Anderson, 1975; Mandler, 1994), at least some experimental research seem to be operating under the assumption that at least some theories are competing, by contrasting them, and interpreting support for one theory as weakened support for other theories.

In the following I will argue that theories of incubation effects are not competing, and that when these different theories have been tested in incubation effect studies, the experimental design has been varied in important ways depending on which theory was being tested. In fact, I will argue that these design differences have meant that different studies claiming to be measuring incubation effects, are in fact measuring a number of different phenomena, rather than a single phenomenon.

Studies of incubation effects have included a multitude of independent variables in their designs - there is no tradition in the incubation effect literature that time away from problem solving should be the only independent variable. Task type, length of incubation time, task during incubation, length of preparation time, interruption vs. impasse, informed vs. uninformed subjects, ability level, gender, and effect measures include but a few of the variations that exist between incubation effect studies. All these variations are captured by the standard operational definition of incubation effects, but the problem is that vastly different designs are included, where both dependent and independent variables differ significantly based on what appears to be differences in the theoretical grounds.

Take for example the fixation-forgetting theory which hypothesizes that fixating elements in the problem environment are what is keeping the problem solver from reaching a solution. Testing such a theory implies including fixating elements in the experimental design (e.g., Smith et al., 1991; Smith et al., 1989; Torrance-Perks, 1997; Hansberry, 1998; Jamieson, 1999). In so far as such a design supports the notion that fixating elements can generate incubation effects, the fixation-forgetting theory is supported. But this result cannot easily be generalized to other incubation effect studies NOT using fixating elements in the design (at least not unless you make the explicit assumption that such fixating elements are present regardless of whether they have been explicitly manipulated by the experimenter). Similarly with the facilitating cue theories; in so far as presenting cues during incubation benefits performance (as tested in studies by Dreistadt, 1969; Mednick et al., 1964; Gick et al., 1980; Browne & Cruse, 1988; Dodds et al., 2002; Dorfman, 1990; Olton & Johnson, 1976; Dominowski et al., 1972; Christensen & Schunn, in press), such theories are supported. But again, this does not generalize easily to other designs not explicitly including such a variable, unless bold assumptions are made, that such analogous cues are present nonetheless. Thus, it would appear that experimental designs constructed to test different theories of incubation effects include variables that make it hard to generalize to other studies which also claim to be measuring incubation effects. Such theories are not competing because their do not generalize to other incubation effect studies that do not manipulate key variables. In other words, since some theories of incubation effects necessitate the presence of key variables in the experimental design, such theories should not claim to be explanations of incubation effect studies that do not include these key variables.

But it need not be only differences in the independent variables that differ between studies. Another example of a theory driven difference between studies making it hard to generalize concerns the incubation effect performance measure itself. The most common way to measure performance is to use so-called closed tasks (i.e., tasks with one or a few ‘correct’ answers); this is for example used in both the fixating and cuing paradigms mentioned above. But a number of studies have also applied open-ended tasks (e.g., Kaplan, 1989; Fulgosi & Guilford, 1968; Silveira, 1971, exp 3; Butler et al., 1999; Houtz et al., 1992; Beck, 1979; Brockett, 1985; Nikolajsen, 1988). Open-ended tasks have no predetermined correct answer, but instead the quality – or creativity - of the answer is determined in other ways, frequently by fluency (Butler et al., 1999; Kaplan, 1989; Silveira, 1971), originality (Beck, 1979), flexibility, or a combination of these measures. A typical example of such a task is the consequences task (used by e.g., Fulgosi et al., 1968; Kaplan, 1989; Silveira, 1971, exp3), which asks subjects to write down as many answers as they can think of to a question such as ‘What would be the result if everybody suddenly lost ability to read and write’. The question is whether it seems theoretically justified to group studies using such different performance measures as open and closed tasks under the common heading ‘incubation effect studies’? If we have reason to believe that results from two effect measures - one of which measures performance by counting correct answers to creative problems, the other which measures performance (in some cases) solely by the *sheer number of answers* given (regardless of other quality estimates) – are in fact not measuring the same thing, then it seems reasonable to assume that we are dealing with different underlying phenomena. In the meta-analysis and review literature this validity issue has been dubbed the ‘apples and oranges’ problem. The problem arises if experimental design varies in important ways between studies making it problematic to assume that the studies are all measuring the same thing – which in turn makes it problematic to make inferences to the whole population of incubation effect studies. Again, if we are dealing with different phenomena, there is no need to assume that theories are competing.

Several more arguments can be put forth that supports the notions that we may be dealing with either more than one cause of incubation effects, or more than one type of incubation effects, or both.

For one thing several of the theories of incubation effects have close relationships with parallel – or analogous – domains outside the incubation literature. If we again look at fixation-forgetting, analogous domains include theories dealing with fixation (e.g., Duncker, 1945), problem solving set/einstellung (e.g., Luchins, 1942), and the Tip-Of-the-Tongue phenomenon (Schwartz, 2001; Schwartz et al., 1997), in that each of these domains hypothesize the importance of blocking elements that has to be overcome to produce a solution. Bowers (Bowers, Farvolden, & Mermigis, 1995; Bowers, Regehr, Balthazard, & Parker, 1990) has further shown that like TOT states, Remote Associates Test items, which are often used in incubation studies, also display a subjective feeling of knowing the answer. Similarly facilitating cue theories have close relationships with theories dealing with analogical problem solving (e.g., Gick et al., 1980; Gick et al., 1983; Keane, 1985) and impasse learning (VanLehn, 1991; VanLehn, 1989). The study of fluency production in open-ended problems is in many ways like some of exemplar production studies used to study concepts (e.g., Barsalou, 1983; Fox, 1970). It could be argued that since these analogous domains are usually considered distinct and unrelated, then we have little reason to believe that they should be related when dealing with incubation effects, just because they are all potentially able to explain one. This would suggest that perhaps several distinct phenomena can generate incubation effects – leaving non-competing theories that each deal with distinct phenomena. Currently it seems unjustified to assume that disparate domains dealing with different cognitive functions and different levels of cognition should really be about the same underlying phenomenon.

This can be further substantiated by what could be called an extreme-case argument. As already mentioned Olton (1979) argued that an – albeit rather uninteresting – test of the facilitating cue theory would be to simply tell the subject the answer to the problem during incubation. In this case an incubation effect would surely be found. Similar arguments could be put forth for fixation-forgetting. If deceitful information is added to the pre-incubation problem solving session making it practically impossible to solve the problem, and this information is removed for the post-incubation problem solving session, an incubation effect will occur. These two extreme cases are obviously unrelated, since they work for different reasons. Whereas the facilitating cue case works because it gives a large benefit to the experimental group, the fixation-forgetting case works because it produces a large temporary detrimental effect on the experimental group. Note that this difference is perfectly testable (see figure 2.).

There are probably more theory-loaded differences between studies making it problematic to generalize, and making it unreasonable to assume that we are dealing with competing theories.

For these reasons I think it would be theoretically unreasonable to assume that theories of incubation effects are competing theories. We may be dealing with different theories identifying multiple non-competing paths to the same phenomenon (i.e., multiple causes of incubation effects).

	Pre-incubation	Post-incubation	
Experimental group receiving *fixating elements* during pre-incubation	<	=	Control group receiving no fixation, and no cues
Experimental group receiving *cues* during incubation	=	>	Control group receiving no fixation, and no cues

Figure 2. Performance level pre and post incubation for experimental groups receiving facilitating cues, fixating elements, or control groups receiving neither

Or we may be dealing with multiple different phenomena (i.e., multiple different incubation effects). Or we may be dealing with both. It has not been my intention to try to prove that there is a fixed number of incubation effects, or causes thereof, substantiated by data - only that there are probably more than one. Rather I only try to show that 1) theories of incubation effects are not competing theories, and 2) due to design differences (some of them theory driven) studies are probably measuring several different phenomena, rather than a single phenomenon. Of course this is not to say that some grand theory synthesizing all this information may not be created some day. But for now, it seems justified to be cautious about treating theories of incubation effects as competing theories. Therefore it is theoretically unreasonable to simply group all studies claiming to measure incubation effects.

The conclusion to be drawn from all this is perhaps that the research question pursued by incubation effect studies so far has been slightly off mark. In the past the research question has been ‘Do incubation effects exist?’ Rather, from the arguments above, it would seem that a better research question to pursue is ‘under what conditions do different incubation effects occur reliably?' This will place focus on the fact that there may be multiple different incubation effects and/or causes thereof – and will set off a search for various conditions (such as different kinds of facilitating cues, fixating elements, task types, etc.) producing reliable incubation effects. Thus, we should not be surprised at finding incubation effects in data, but what should surprise us is if such incubation effects are a consistent finding across *all possible* studies labeled 'incubation effect studies'.

Assumption 3: Incubation Effects are Creativity Constructs

Incubation (along with insight and problem finding) is one of the most prototypical of the phenomena in the creativity domain. Wallas (1926) first described it precisely to characterize the creative process, and most textbooks would have us place the phenomenon under the creativity headline. Thus, it would seem that incubation and creativity go hand in hand.

Nonetheless, there appears to be ambivalence in the domain as to the precise relation between incubation effects and creativity. On the one hand the ambition is definitely to try to generalize and make inferences from incubation effect studies to the incubation phenomenon and creativity. But on the other hand, quite a few studies do not include 'creativity' as a necessary criterion for the study to be valid. As Kaplan (1989) noted, studies testing for incubation effects have consistently used 'non-standard' problems in their experiments, but that is not to say that they all concern creativity. The reason for this cautious treatment of the 'creativity' construct in some of these studies may well be the fear that they may be criticized for their choice of task when stating that the task is a creative one. Creativity tests have a rather poor discriminant and predictive, along with criterion, validity (Wallach, 1976; Weisberg, 1993). Many different tasks and tests have been used, e.g., remote associates tests, insight problems, anagrams, the hat rack problem, word fragment completion, and open-ended problems such as the consequences task, 'uses of a brick', and smoking roommate. But it is safe to say that there is no single best way or agreed upon task that ensures criterion validity in creativity tests; each task has problems.

No study has to my knowledge ever tested whether ‘creative’ and ‘non-creative / standard’ tasks differ with respect to incubation effects. So, the assumption held by most theories and studies of incubation effects, that they are dealing with a creativity construct *par excellence*, has never been tried and tested. Rather, the reverse seems to be true; there are reasons to believe that incubation effects are not limited to creativity alone. One such reason is the fact that most of the tasks used in incubation effect studies appear to have more in common with extra-creativity-domain tasks that are not considered related to creativity, than they do with each other (see table 2 for a selection).

Table 2: The relation between ‘creative’ and ‘non-creative’ tasks

’Creative’	’Non-creative’
Math insight problem (Metcalfe & Wiebe, 1987) Water lilies double in area every 24 hours. At the beginning of summer there is one water lily on the lake. It takes 60 days for the lake to be completely covered with water lilies. On which day is the lake half covered?	**Impasse learning problem** (VanLehn, 1990) 8001 - 43
RAT (Mednick et al., 1964) (Instructions: Find the common associate) surprise line birthday (Answer= party)	**TOT** (Schwartz, 2001) “For which country is the rupee the monetary unit” (Answer= India)
Fluency production (Kaplan, 1989) (Instructions: write down as many answers as you can to the following question) ‘What would be the result if everybody suddenly lost ability to read and write’	**Exemplar generation** (Barsalou, 1983) (Instructions: Generate the first four exemplars that come to mind when you see this category) ‘Things that can attack something’

As it were, some of these non-creative tasks also display incubation effects! For example, the very essence of the TOT phenomenon is a kind of incubation effect, where time away helps performance, compared to continuous work.

Thus, incubation effect studies appear to be suffering under the implicit assumption of a close relationship with creativity. Several implicit assumptions appear to be present in the operational definition, such as (1) the task type should be 'non-standard'; (2) there is a difference between 'standard' and 'non-standard' task types when dealing with incubation effects; (3) 'Standard' problem solving will show reliable continuation effects. Stating the problem in this manner makes it interesting to contrast 'standard' with 'non-standard' problem solving experimentally when doing incubation studies. For some reason, no one seems to have conducted such an experiment, although such task type contrasts have been used elsewhere in creativity research (e.g., Metcalfe & Wiebe, 1987; Schooler, Ohlsson, & Brooks, 1993; Jausovec, 1997). Conducting this task type contrast seems extremely important, since it could potentially explain why incubation studies have had problems finding large reliable incubation effects. Perhaps the incubation phenomenon only increases performance on 'non-standard' problems *relative to* other 'standard problems'. If this is the case, then very small reliable incubation effects (or perhaps even small continuation effects) on 'non-standard' problems will be interesting, if 'standard' problems show reliable and large continuation effects. We may have reason to believe that 'standard' problems will produce such large continuation effects, as we will see in the next assumption.

In summary, some good advice for the incubation researcher is that a tight link between incubation effects and creativity should not be assumed. The link has never been empirically tested, and seems problematic. In fact, the reverse seems to be warranted - that incubation effects can occur on some 'non-creative' or 'standard' tasks. This also means that it should not be assumed that the only studies measuring incubation effects are the ones that *claim* they measure incubation effects.

For example, TOT studies, some impasse learning studies and maturation studies can all be said to be measuring incubation effects, even though they do not list themselves under that heading. Finally care should be taken in selecting tasks when conducting incubation effects studies.

Assumption 4: Incubation Effect Studies are Interesting because they concern the Incubation Phenomenon

When we question the link between the incubation phenomenon and the incubation effect studies, then it seems reasonable to ask if the present incubation effect studies are completely uninteresting in their own right. They need not be, and the reason is memory research in cognitive science, not specifically concerning creativity. Well-known cognitive effects in working memory, including rapid forgetting (Peterson & Peterson, 1959) due to capacity limitations (Miller, 1956) and fragility to distraction and interference (e.g., Nairne, 1996) would have us believe that in so far as problem solving draws heavily on working memory (as is expected in most problem solving theories, (including heuristic space theory; Newell & Simon, 1972)) it is surprising that putting the problem aside should improve performance. Quite the contrary is to be expected; on standard problem solving problems (such as arithmetic problems), a continuation effect should be the norm due to rapid forgetting processes when the problem is set aside, in part since re-familiarization and re-orientation (‘now, where was I?’) with the problem should be required upon returning to the problem (also known as resumption lag, Altmann & Trafton, 2004). Support for this claim can be found in organizational psychology and human-computer interaction studies examining the effect of interruptions on performance. A growing literature provides evidence that, at least on complex tasks, interruptions (to perform other tasks) have a negative effect on performance in terms of how quickly people perform the post-interruption task (Gillie & Broadbent, 1989; Field & Spence, 1994; Rogers & Monsell, 1995; Altmann et al., 2004; Bailey, Konstan, & Carlis, 2001; Burmistrov &

Leonova, 2003; Cutrell, Czerwinski, & Horvitz, 2001), and the number of errors made (Gillie et al., 1989; Speier, Vessey, & Valacich, 2003; Edwards & Gronlund, 1998; Rogers et al., 1995). This literature implies that interrupting complex tasks will produce some level of disruption. This is contrary to what incubation effect studies are trying to find evidence of, and any reliable incubation effect would therefore be interesting to memory research. As already highlighted, existing studies on incubation effects usually employ 'interruptions', rather than allowing subjects to reach a self-generated impasse. So a large and reliable incubation effect in problem solving using the standard definition and design should be extremely surprising, due to the existence of other well established findings in cognitive science domains such as forgetting and working memory. It could still be argued that perhaps creative problem solving takes place in long term memory (unconsciously) instead of in working memory, but no mainstream theory of long-term memory hypothesizes that higher order cognitive problem solving occurs in this kind of storage. In summary, although existing laboratory incubation effect studies using the standard operational definition and experimental design may have a questionable relation to the alleged incubation phenomenon, that does not make them uninteresting. Quite the contrary; well-established cognitive phenomena would have us predict continuation effects in such experimental designs as are usually employed in incubation effect studies, and so any incubation effect would be interesting. The interest would thus not arise from the relation to the incubation phenomenon, but rather from the relation to memory research.

Assumption 5: Incubation Effects Arise from Benefits to Performance, and Show Positive Correlations between Length of Time away and Performance

The assumption exists in some of the incubation literature that results from studies using the operational definition of incubation effects basically shows how length of time away from active problem solving correlates positively with problem solving performance. The assumption seems straightforward (at least as seen from the perspective of unconscious idea generation theories), but given close scrutiny, it does not hold up, or at least should be qualified. There are two problems with this assumption. First, incubation study designs do not frequently include variations of length of time away from problems solving, although a few studies have done this (e.g., Fulgosi et al., 1968; Silveira, 1971; Beck, 1979; Dorfman, 1990). The most frequent paradigm simply compares an experimental group receiving some time away from problem solving, with a control group that does not. Such a paradigm does not tell you that the *length of* time away matters, only that *some* time away matters. The relatively few studies that have manipulated length of time away have examined a rather narrow timeframe (usually less than one hour) and have found somewhat mixed results, although it does appear that within this narrow timeframe, about 30 minutes seems to improve performance slightly more compared to shorter conditions (Dodds et al., in press). One problem with studying *length* of incubation time is that it is entirely unclear what the timeframe should be. It remains a completely open question whether incubation effects are a frequent occurrence (e.g., several times an hour or day) or a rare exception (on a yearly basis or perhaps a few times in a lifetime). The gap between anecdotal accounts and experimental studies becomes apparent in this regard, with the former being obviously a far more rare occurrence than is assumed in the experimental paradigm.

The distinction between *length of time* away and the binary *some time vs. no time* away is theoretically important. The prior assumes that performance is related to the amount of time away, while the latter makes no such assumption. This makes the distinction somewhat related to what Dorfman, Shames & Kihlstrom (1996) called autonomous versus interactive incubation effects. Autonomous incubation effects refer to effects where the cause is assumed to be some (cognitive) phenomena which is associated with length of incubation time (typically as a positive correlate or curvilinear relationship). Interactive incubation effects, on the other hand, assume that the cause is some interaction between subject and environment during incubation (e.g., picking up new information, contextual shifting, fixating elements in the environment). Here focus is on variables correlating with the binary ‘spending *some* time away (as opposed to *no* time)’ from problem solving, such as perceiving analogous information to the as yet unsolved problem. This distinction between autonomous and interactive incubation effects effectively divides theories of incubation into groups, and, as argued, carries implications for the experimental study as well.

Another assumption in the above statement that does not stand up to close scrutiny, is the assumption that the cause behind incubation effects is necessarily a *beneficial* contribution to the performance in the experimental group, that is, a phenomenon which increases performance. Of course it may be beneficial, as hypothesized by the facilitating cues and unconscious idea generation theories, but it need not be. As already mentioned in this paper, some incubation effect theories hypothesize that the cause of incubation effects is in essence a *temporary detrimental* effect on performance in the experimental group. This is for example the case in the fixation-forgetting hypothesis, where the group receiving fixating elements will perform worse than the control group in the initial problem solving period (prior to incubation). I would like to propose that the distinction between beneficial and temporary detrimental incubation effects is a potentially important one, which suggests different types of causal mechanisms.

These two distinctions (between autonomous/interactive and beneficial/temporary detrimental) can be seen as two orthogonal dimensions in which theories of incubation effects can be placed (see table 3).

Table 3: Theories of incubation effects divided by whether they hypothesize autonomous vs. interactive mechanisms, and beneficial vs. temporary detrimental effects

	Autonomous	Interactive
Beneficial	Unconscious idea generation (e.g., Wallas, 1926)	Facilitating cues (e.g., Seifert et al., 1995; Langley et al., 1988)
Temporary detrimental	Fixation-forgetting (e.g., the returning-act hypothesis, Segal, 2004); fatigue dissipation (e.g., Woodworth et al., 1954)	Fixation-forgetting w/contextual shifting (e.g., Smith, 1995a)

In conclusion, it should not be assumed that incubation effects are necessarily related to the length of time away from problem solving (although that may be the case). It may just as well be that the incubation effect performance increase is caused by some correlate of the binary spending *some* time away from problem solving as opposed to spending no time away. Thus, the *amount* of time spent away from problem solving may be somewhat irrelevant to incubation effects. Furthermore, it should not be assumed that the cause of incubation effects implies a beneficial contribution to performance (although that may be the case). Incubation effects can also be caused by a temporary detrimental effect on performance, which is then somehow removed during time away.

Conclusion: The Future of Incubation Studies

The past about 68 years of research on incubation have seen incubation studies develop from a few early fragmented studies into a somewhat established field of research with seeming agreement on both operational definitions and experimental design.

But the establishment of definitional and design issues also implies some assumptions about incubation and incubation effects – assumptions that are widely held in the incubation field of research, but some of which are problematic, and should be treated with care.

For one thing, there appears to have developed a kind of conceptual gap between the operational definition of incubation effects used in experimental studies, and the incubation phenomenon they purport to be studying, as described in anecdotes. The operational definition is ‘any positive effect from having a break away from active problem solving on performance’, but this definition includes non-creative studies along with creative ones, while excluding any studies not dealing explicitly with performance issues. In so far as the incubation phenomenon is restricted to creative tasks or in so far as the incubation phenomenon is not solely a performance issue, the definition may be off mark. This, however, does not mean that incubation effect studies are irrelevant, but it does mean that the reason why they are interesting is somewhat different from what is usually assumed in the literature. They may not be interesting because of their direct relation to the alleged incubation phenomenon, but may instead be interesting because they seemingly violate expectations from other well-known memory phenomena in cognitive science, such as working memory capacity and forgetting mechanisms. Thus, it is problematic to assume that incubation effect studies are directly related to the incubation phenomenon, and care should be taken before generalizing from the one to the other.

A second problematic assumption deals with theories and studies of incubation effects. Different theories of incubation effects postulate different key variables that explain the cause of the effect. But these variables have consequences for the experimental designs used, making it problematic to assume that designs testing one theory will generalize to another design testing another theory. In fact, the use of significantly different designs makes it probable that incubation effects are either more than one effect or caused by more than one phenomenon, or both. This is further supported by the fact that different experimental studies using creative or non-standard materials seemingly have more in common with research domains outside creativity, than they do with each other. The implication is that the research question pursued should not concern merely testing if incubation effects exist at all, but rather attempt to localize (disparate) conditions where they occur reliably. Further, care should be taken before generalizing across all incubation effect studies, since they may not be measuring the same effect. One should not assume that incubation effects are tightly linked to creativity – rather the creativity-incubation effect link should be tried and tested in experimental task type contrasts before making such assertions.

Finally, different theories have helped highlight that incubation effects can be caused both by a phenomenon operating in the cognitive system without reference to environmental stimuli (so-called autonomous processes) and from phenomena operating from an interaction between the cognitive system and the environment (so-called interactive processes). Incubation effects need not be a correlate with the length of the incubation period – but may simply be a correlate of spending *some* time away (regardless of length) from problem solving. Incubation effects can arise from both a positive contribution to the problem solving process (i.e., be beneficial to problem solving), but it may also arise from a temporary negative effect on problem solving, where the detrimental effect is removed during incubation.

The conclusion based on the treatment of the problematic underlying assumptions behind the typical incubation effect study, thus seems to be that the typical incubation effect study need not tell us much, if anything, about the incubation phenomenon as described in anecdotes. Although the typical incubation effect study could still potentially be a (perhaps limited) window into the workings of the incubation phenomenon, the present article has tried to present arguments against this assertion. But on the other hand, that does not make the typical incubation effect studies irrelevant or pointless, as I have also tried to show, due to research on working memory and forgetting.

In pointing to this seeming distinction between the operational definition of incubation effects and the incubation phenomenon, I have attempted to show the need for redirecting incubation research. As shown in the above analysis, the standard operational definition of incubation effects leads to research that is off the mark in so far as the ambition is to generalize to the incubation phenomenon. It has not proven fruitful to pursue the research question of whether incubation effects exist at all by conducting and comparing studies using the standard operational definition. A more fruitful approach would be to divide studies into theoretically relevant categories (e.g., by dividing studies according to the particular theory of incubation they purport to be studying, or along whether the study tests theories hypothesizing beneficial vs. temporary detrimental effects and autonomous vs. interactive effects) and then examine what evidence exist for the different (kinds of) theories of incubation effects. Such an approach would open up for the possibility that different (kinds of) theories have different conceptualizations of what incubation is, and therefore may be pointing towards different kinds of effects, where the theoretical differences make generalizations across all incubation effect studies difficult or impossible.

The problematic consequences of using the standard operational definition of incubation effects also suggests that it is time to study incubation in other ways than by using the standard operational definition. Rather than starting every incubation study by designing an experiment in accordance with the standard operational definition, it would seem to be a better approach to develop experiments that directly test predictions from the different incubation effect theories. Rejecting the standard operational definition of incubation effects, and focusing in stead on the different theories of incubation will force the incubation researcher back to the experimental design drawing board. New approaches to the study of the different theories of incubation may take on a variety of forms, but since previous incubation effect studies have been lacking ecological validity (Olton, 1979; Kaplan, 1989), one interesting research path would be to attempt to study the incubation phenomenon in the real-world, by studying creative scientists, artists, and inventors in their everyday creative process. Such studies should preferably be prospective in nature, avoiding the pitfall of the retrospective anecdotes used so frequently in the past. This is not a recommendation for eliminating experimental studies of incubation, but rather an attempt to tie the experimental studies much closer in with real-world studies of incubation. Inspiration for such an approach may come from Dunbar’s (1995; 1997; 2000; 2001; 1999; Blanchette & Dunbar, 2001) in vivo-in vitro approach, which is precisely an attempt show how the same phenomenon can be studied both prospectively in the real-world, and in the psychological laboratory within the same design. The methodology has been used to study analogy and other higher-order cognitive functions using variations of think-aloud protocols, and so may be useful in the study of incubation as well.

Even though I have argued for the rejection of the standard operational definition of incubation effects, the working memory researcher may still find this definition interesting, and may want to pursue the traditional experimental incubation effect studies. But in stead of having the ambition to generalize to the incubation phenomenon, the ambition should be to inform theories of memory, in particular working memory and forgetting. Part of such a line of research could consist of conducting task type contrasts (in particular between ‘creative’ and ‘non-creative’ tasks) which may (or may not) reveal that working memory attributes such as capacity and forgetting mechanisms may be somewhat task specific.

In closing, the hope is that some of the confusing assumptions underlying incubation studies have been removed (although some may complain that a bit of the incubation mystique went along with it). My hope is that nonetheless this paper has helped incubation researchers to ask better questions, starting new research approaches, while avoiding some of the pitfalls. And with this I will put incubation aside in the hope that many future creative scientists will return to the problem with fresh insights.

References

Altmann, E. M. & Trafton, J. G. (2004). Task interruption: Resumption lag and the role of cues. In *Proceedings of the 26th annual conference of the Cognitive Science Society*.

Anderson, B. F. (1975). *Cognitive psychology. The study of knowing, learning, and thinking*. London: Academic Press.

Arieti, S. (1976). *Creativity: The magic synthesis*. Oxford, England: Basic.

Bailey, B. P., Konstan, J. A., & Carlis, J. V. (2001). The effects of interruptions on task performance, annoyance, and anxiety in the user interface. In M. Hirose (Ed.), *Human-Computer Interaction – INTERACT 2001 Conference Proceedings. IOS Press, IFIP*, 593-601.

Barsalou, L. W. (1983). Ad hoc categories. *Memory and Cognition, 11,* 211-227.

Barsalou, L. W. (1985). Ideals, central tendency, and frequency of instantiation as determinants of graded structure in categories. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 11,* 629-654.

Barsalou, L. W. (1987). The instability of graded structure: Implications for the nature of concepts. In U.Neisser (Ed.), *Concepts and conceptual development: Ecological and intellectual factors in categorization* (pp. 101-140). Cambridge: Cambridge University Press.

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

Bellezza, F. S. (1984). Reliability of retrieval from semantic memory: Common categories. *Bulletin of the Psychonomic Society, 22,* 324-326.

Bennett, S. M. (1975). *Effect of incubation on the associative process of creativity.* Unpublished doctoral dissertation Nortre Dame.

Blanchette, I. & Dunbar, K. (2001). Analogy use in naturalistic settings: The influence of audience, emotion, and goals. *Memory and Cognition, 29,* 730-735.

Boden, M. A. (1990). *The creative mind: Myths and mechanisms*. New York: Basic Books.

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

Bowers, K. S., Farvolden, P., & Mermigis, L. (1995). Intuitive antecedents of insight. In S.M.Smith, T. B. Ward, & R. A. Finke (Eds.), *The creative cognition approach.* (pp. 27-51). Cambridge, MA, US: The MIT Press.

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

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

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

Burmistrov, I. & Leonova, A. (2003). Do interrupted users work faster or slower? The micro-analysis of computerized text editing task. In J. Jacko & C. Stephanidis (Eds.), *Human-Computer Interaction: Theory and Practice (Part I) – Proceedings of HCI International 2003, Vol. 1*. Mahwah: Lawrence Erlbaum Associates, 621-625.

Butler, D. L. & Thomas, K. M. (1999). Preliminary study of the effectiveness of some heuristics used to generate solutions to ill-structured problems. *Psychological Reports, 84,* 817-827.

Campbell, D. T. (1960). Blind variation and selective retentions in creative thought as in other knowledge processes. *Psychological Review, 67,* 380-400.

Carpenter, W. B. (1876). *Principles of mental physiology*. London: Routledge.

Christensen, B. T. & Schunn, C. D. Spontaneous access and analogical incubation effects. *Creativity Research Journal,* (in press).

Cutrell, E., Czerwinski, M., & Horvitz, E. (2001). Notification, disruption, and memory: Effects of messaging interruptions on memory and performance. In M. Hirose (Ed.), (pp. 263-269). Amsterdam: IOS Press.

Davidson, N. S. (1995). *The role of retrieval failures in memory incubation.* Unpublished doctoral dissertation. University of Michigan.

Dodds, R. A. (1999). *The use of environmental clues during incubation.* Unpublished doctoral dissertation. Texas A&M University.

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

Dodds, R. A., Ward, T. B., & Smith, S. M. (in press). A review of the experimental literature on incubation in problem solving and creativity. In M.A.Runco (Ed.), *Creativity research handbook* (3rd ed.), Cresskill, NJ: Hampton Press.

Dominowski, R. L. & Jenrick, R. (1972). Effects of hints and interpolated activity on solution of an insight problem. *Psychonomic Science, 26(6),* 335-338.

Dorfman, J. (1990). *Metacognitions and incubation effects in insight problem solving.* Unpublished doctoral dissertation. University of California, San Diego.

Dorfman, J., Shames, V. A., & Kihlstrom, J. F. (1996). Intuition, incubation, and insight: Implicit cognition in problem solving. In G.D.M.Underwood (Ed.), *Implicit cognition* (pp. 257-296). New York: Oxford University Press.

Dorsel, T. N. (1979). Creativity: incubation as a special case of reminiscence. *Journal of*

Creative Behavior, 13, 53-55.

Dreistadt, R. (1969). The use of analogies and incubation in obtaining insights in creative problem solving. *Journal of Psychology, 71,* 159-175.

Dunbar, K. (1995). How scientists really reason: Scientific reasoning in real-world laboratories. In R.J.Sternberg & J. E. Davidson (Eds.), *The nature of insight.* (pp. 365-395). Cambridge, MA, US: The MIT Press.

Dunbar, K. (1997). How scientists think: On-line creativity and conceptual change in science. In T.B.Ward, S. M. Smith, & J. Vaid (Eds.), *Creative thought: An investigation of conceptual structures and processes.* (pp. 461-493). Washington, DC: American Psychological Association.

Dunbar, K. (1999). How scientists build models: InVivo science as a window on the scientific mind. In L.Magnani, N. Nersessian, & P. Thagard (Eds.), *Model-based reasoning in scientific discovery.* (pp. 89-98). NY, US: Plenum Press.

Dunbar, K. (2000). How scientists think in the real world: Implications for science education. *Journal of Applied Developmental Psychology, 21,* 49-58.

Dunbar, K. (2001). The analogical paradox: Why analogy is so easy in naturalistic settings yet so difficult in the psychological laboratory. In D.Gentner, K. J. Holyoak, & B. N. Kokinov (Eds.), *The analogical mind: Perspectives from cognitive science.* (pp. 313-334). Cambridge, MA: The MIT Press.

Duncker, K. (1945). *On Problem-solving.* Westport, CT, US: Greenwood Press, Publ.

Edwards, M. B. & Gronlund, S. D. (1998). Task interruption and its effect on memory. *Memory, 6,* 665-687.

Field, G. & Spence, R. (1994). Now, where was I? *New Zealand Journal of Computing, 5,* 35-43.

Fox, P. W. (1970). Patterns of stability and change in behaviors of free association. *Journal of Verbal Learning and Verbal Behavior, 9,* 30-36.

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

Ghiselin, B. (1954). *The creative process: A symposium*. Berkeley: University of California Press.

Gick, M. L. & Holyoak, K. J. (1980). Analogical problem solving. *Cognitive Psychology, 12,* 306-355.

Gick, M. L. & Holyoak, K. J. (1983). Schema induction and analogical transfer. *Cognitive Psychology, 15,* 1-38.

Gillie, T. & Broadbent, D. (1989). What makes interruptions disruptive? A study of length, similarity, and complexity. *Psychological research, 50,* 243-250.

Gratzer, W. (2002). *Eurekas and euphorias. The Oxford book of scientific anecdotes*. New York: Oxford University Press.

Hadamard, J. (1945). *The psychology of invention in the mathematical field*. Princeton: Princeton University Press.

Hansberry, M. T. (1998). *Fixation and incubation effects in problem-solving.* Unpublished doctoral dissertation University of New Hampshire.

Hedges, L. V. (1994). Statistical considerations. In H.Cooper & L. V. Hedges (Eds.), *The handbook of research synthesis* (pp. 29-40). New York: Russel Sage Foundation.

Helmholtz, H. v. (1896). *Vorträge und Reden*. Braunsweig, Germany: Friedrich Vieweg und Sohn.

Holyoak, K. J. & Koh, K. (1987). Surface and structural similarity in analogical transfer. *Memory & Cognition, 15,* 332-340.

Houtz, J. C. & Frankel, A. D. (1992). Effects of incubation and imagery training on creativity. *Creativity Research Journal, 5,* 183-189.

Hull, C. L. (1943). *Principles of behavior. An introduction to behavior theory*. (Seventh printing) New York: Appleton-Century-Crofts.

Jamieson, B. A. (1999). *Incubation and aging: The nature of processing underlying insight.* Unpublished doctoral dissertation University of Georgia.

Jausovec, N. (1997). Differences in EEG activity during the solution of closed and open problems. *Creativity Research Journal, 10:* 317-324.

Jett, Q. R. & George, J. M. (2003). Work interrupted: A closer look at the role of interruptions in organizational life. *Academy of Management Review,* 494-507.

Johnson, D. M., Johnson, R. C., & Mark, A. L. (1951). A mathematical analysis of verbal fluency. *Journal of General Psychology, 44,* 121-128.

Jones, G. V. & Langford, S. (1987). Phonological blocking in the tip of the tongue state. *Cognition, 26:* 115-122.

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

Keane, M. (1985). On drawing analogies when solving problems: A theory and test of solution generation in an analogical problem-solving task. *British Journal of Psychology, 76,* 449-458.

Kirkwood, W. G. (1984). Effects of incubation sequences on communication and problem solving in small groups. *Journal of Creative Behavior, 18,* 45-61.

Kris, E. (1952). *Psychoanalytic explorations in art*. New York: International Universities Press.

Kubie, L. S. (1958). *Neurotic distortion of the creative process*. Toronto: The Noonday Press.

Langley, P. & Jones, R. (1988). A computational model of scientific insight. In R.J.Sternberg (Ed.), *The nature of creativity: Contemporary psychological perspectives.* (pp. 177-201). New York: Cambridge University Press.

Luchins, A. S. (1942). Mechanisation in problem solving. The effect of Einstellung. *Psychological Monographs, 54,* no. 248.

Luria, A. R. (1973). *The working brain*. New York: Penguin.

Mandler, G. (1994). Hypermnesia, incubation and mind popping: On remembering without really trying. In C.Umiltà & M. Moscovitch (Eds.), *Conscious and nonconscious information processing* (pp. 3-33). Cambridge, MA: The MIT Press.

Martindale, C. (1990). Creative imagination and neural activity. In R.G.Kunzendorf & A. A. Sheikh (Eds.), *The psychophysiology of mental imagery*. (pp. 89-108). Amityville, NY, US: Baywood.

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

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

Mensink, G. J. & Raaijmakers, J. G. (1988). A model for interference and forgetting. *Psychological Review, 95,* 434-455.

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

Miller, G. A. (1956). The magic number seven, plus or minus two: Some limits on our capacity for processing information. *Psychological Review, 27,* 338-352.

Moss, S. A. (2002). The impact of environmental clues in problem solving and incubation: The moderating effect of ability. *Creativity Research Journal, 14,* 207-211.

Murphy, G. (1947). *Personality: A biosocial approach to origins and structure*. New York: Basic Books Inc., Publishers.

Nairne, J. S. (1996). Short-term/working memory. In E.L.Bjork & R. A. Bjork (Eds.), *Memory* (2nd ed., pp. 102-130). San Diego: Academic press.

Newell, A. & Simon, H. A. (1972). *Human problem solving*. Englewood Cliffs, NJ: Prentice-Hall.

Nikolajsen, H. R. (1988). *Relaxation as the form of incubation in the creative thinking process.* Unpublished doctoral dissertation Texas A&M University.

Noble, J. (1998). *A search for incubation in the ecology of learning.* Unpublished doctoral dissertation The Fielding Institute.

Ochse, R. (1990). *Before the gates of excellence*. Cambridge: Cambridge University Press.

Ohlsson, S. (1992). Information-processing explanations of insight and related phenomena. In M.T.Keane & K. J. Gilhooly (Eds.), *Advances in the psychology of thinking.* (pp. 1-40). London, England: Harvester Wheatsheaf.

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

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

Patalano, A. L. & Seifert, C. M. (1994). Memory for impasses during problem solving. *Memory and Cognition, 22,* 234-242.

Patrick, C. (1937). Creative thought in artists. *Journal of Psychology, 4,* 35-73.

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

Payne, D. G. (1987). Hypermnesia and reminiscence in recall: A historical and empirical review. *Psychological Bulletin, 101,* 5-27.

Perkins, D. N. (1981). *The mind's best work.* Cambridge, MA: Harvard University Press.

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, US: The MIT Press.

Peterson, L. R. & Peterson, M. J. (1959). Short-term retention of individual verbal items. *Journal of Experimental Psychology, 58,* 193-198.

Platt, W. & Baker, R. A. (1931). The relation of the scientific 'hunch' to research. *Journal of Chemical Education, 8,* 1969-2002.

Posner, M. I. (1973). *Cognition: An introduction*. Glenview, Il: Scott, Foresman & Co.

Reitman, W. R. (1965). *Cognition and thought: An Information-Processing approach*. New York: John Wiley & Sons.

Rogers, R. D. & Monsell, S. (1995). Costs of a predictable switch between simple cognitive tasks. *Journal of Experimental Psychology: General, 124,* 207-231.

Rossman, J. (1931). *The psychology of the inventor*. Washington: The Inventors Publishing Co.

Roth, E. M. & Shoben, E. J. (1983). The effect of context on the structure of categories.

Cognitive Psychology, 15, 346-378.

Rugg, H. (1963). *Imagination*. New York: Harper & Row, Publ.

Schooler, J. W., Ohlsson, S., & Brooks, K. (1993). Thoughts beyond words: When language overshadows insight. *Journal of Experimental Psychology: General, 122,* 166-183.

Schwartz, B. L. (2001). The relation of tip-of-the-tongue states and retrieval time. *Memory & Cognition, 29,* 117-126.

Schwartz, B. L. & Smith, S. M. (1997). The retrieval of related information influences tip-of-the-tongue states. *Journal of Memory & Language, 36,* 68-86.

Seabrook, R. & Dienes, Z. (2003). Incubation in problem solving as a context effect. In R. Alterman & D. Kirsch (eds.). *Proceedings from the 25th Annual Meeting of the Cognitive S*cience Society. Boston: Cognitive Science Society.

Segal, E. (2004). Incubation in insight problem solving. *Creativity Research Journal, 16,* 141-148.

Seifert, C. M., Meyer, D. E., Davidson, N., Patalano, A. L., & Yaniv, I. (1995). Demystification of cognitive insight: Opportunistic assimilation and the prepared-mind perspective. In R.J.Sternberg & J. E. Davidson (Eds.), *The nature of insight.* (pp. 65-124). Cambridge, MA: The MIT Press.

Shaw, G. A. (1987). Creativity and hypermnesia for words and pictures. *Journal of General Psychology, 114,* 167-178.

Shepard, R. N. (1978). Externalization of mental images and the act of creation. In B.S.Randawa & W. E. Cofman (Eds.), *Visual learning, thinking, and communication* (pp. 133-189). New York: Academic Press.

Siegler, R. S. & Jenkins, E. (1989). *How children discover new strategies*. Hillsdale, NJ: Erlbaum.

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

Simon, H. A. (1966). Scientific discovery and the psychology of problem solving. In R.G.Colodny (Ed.), *Mind and Cosmos* (pp. 22-41). Pittsburgh, PA: University of Pittsburgh Press.

Smith, S. M. (1995a). 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. 135-156). Cambridge, MA: The MIT Press.

Smith, S. M. (1995b). Getting into and out of mental ruts: A theory of fixation, incubation, and insight. In R.J.Sternberg & J. E. Davidson (Eds.), *The nature of insight.* (pp. 229-251). Cambridge, MA: The MIT Press.

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

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

Smith, S. M. & Dodds, R. A. (1999). Incubation. In M.A.Runco & S. R. Pritzker (Eds.), *Encyclopedia of creativity* (pp. 39-43). San Diego, CA: Academic Press.

Speier, C., Vessey, I., & Valacich, J. S. (2003). The effects of interruptions, task complexity, and information presentation on computer-supported decision-making performance. *Decision Sciences, 34,* 771-797.

Stagner, R. & Karwoski, T. F. (1952). *Psychology*. New York: McGraw-Hill Book Company, Inc.

Torrance-Perks, J. (1997). *The incubation effect: Implications for underlying mechanisms.* Unpublished doctoral dissertation University of Waterloo.

VanLehn, K. (1988). Toward a theory of impasse-driven learning. In H.Mandl & A. Lesgold (Eds.), *Learning issues for intelligent tutoring systems.* New York: Springer-Verlag.

VanLehn, K. (1989). Problem solving and cognitive skill acquisition. In M.I.Posner (Ed.), *Foundations of cognitive science* (pp. 527-579). Cambridge, MA, US: The MIT Press.

VanLehn, K. (1990). *Mind bugs: The origins of procedural misconceptions*. Cambridge, Massachusetts: The MIT Press.

VanLehn, K. (1991). Rule acquisition events in the discovery of problem-solving strategies. *Cognitive Science, 15,* 1-47.

Wagner, U., Gais, S., Haider, H., Verleger, R., & Born, J. (2004). Sleep inspires insight. *Nature, 417,* 352-355.

Wakefield, J. F. (1989). Creativity and cognition: Some implications for arts education. *Creativity Research Journal, 2,* 51-63.

Wallach, M. A. (1976). Tests tell us little about talent. *American scientist, Jan-Feb,* 57-63.

Wallas, G. (1926). *The art of thought*. London, England: Jonathan Cape.

Weisberg, R. W. (1993). *Creativity. Beyond the myth of genius.* New York: W. H. Freeman and company.

Whyte, L. L. (1973). Where do those bright ideas come from? In G.Tate (Ed.), *From discovery to style: A reader* (pp. 34-41). Cambridge, MA: Wintrop Publishers, Inc.

Woodworth, R. S. & Schlosberg, H. (1954). *Experimental psychology*. (Revised ed.) New York: Holt, Rinehart and Winston.

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

Article 2

Spontaneous Access and Analogical Incubation Effects

Christensen, B. T. & Schunn, C. D. (2005),

Creativity Research Journal, 17(2), 207-220

Abstract

Incubation often plays a role in creative problem solving. Theories of analogical problem solving, and Opportunistic Assimilation (OA) theory (Seifert, Meyer, Davidson, Patalano, & Yaniv, 1995) of incubation in creative problem solving, were tested. OA theory predicts that a previously encoded unsolved problem will be spontaneously accessed and mapped by a later encounter with an analogous problem-with-solution. The present study tested the OA predictions on insight problems, and found spontaneous access effects for uninformed subjects. Spontaneous access was compared to a baseline derived from the same subject's typical problem solving behavior following distracter tasks. Following access of the analogous cue, subjects increased their performance, demonstrating incubation effects. These results support the theory that at least some incubation effects are caused by previously unsolved problems being solved upon later chance encounters with relevant information in the environment.

Introduction

Since Gick & Holyoak (1980; 1983) conducted their classic studies of Duncker's (1945) radiation problem, the standard way of studying analogical problem solving has consisted of first letting subjects encode a problem-with-solution (source), and then afterwards presenting the subjects with an analogous problem to solve (target). This task structure is mirrored in prevailing models of reasoning by analogy and similarity such as MAC/FAC (Forbus, Gentner, & Law, 1994), LISA (Hummel & Holyoak, 1997) and ARCS (Thagard, Holyoak, Nelson, & Gochfeld, 1990). The canonical model divides analogical problem-solving into at least three phases: first, the source is encoded; second, a target induces access of a source relevant to the solution of the target; and third, the source is mapped unto the target to draw a solution (e.g., Anolli, Antonietti, Crisafulli, & Cantoia, 2001).

However, experiment V of Gick and Holyoak (1980, experiment V) also included a variation of this analogical problem solving ordering. The experimental variation consisted of changing the order of the source and target, by first allowing the subjects time to work on the radiation problem (target), followed by reading and recall of an analogous story (source - the Attack-Dispersion story), again followed by work on the radiation problem. This 'target-source' condition was compared to the standard way of doing analogical problems solving (i.e., source before target), and to a control receiving no analogous story, but with incubation time away from the problem. There was no indication that the target-source condition increased performance over the source-target condition, but receiving the analogous story did improve performance over the incubation control condition. This variation of the standard analogical problem solving paradigm has only been used a few times in the analogical problem solving literature.

Another study using this 'reverse' (i.e., target-source) analogical problem solving design was by Keane (1985). That study tested whether presenting merely the solutions to analogous problems was enough to improve performance over an incubation control group, again using the radiation problem, but this time using distracter tasks. Subjects were instructed that the solutions might be useful. This hypothesis was supported.

Although the dominant theories of analogical access would not predict any difference for these two different directions of analogical retrieval, it is especially interesting for studies of creativity to examine this 'reverse' direction further. For creative problems, where the solution can be surprising and may restructure the entire problem representation (Dominowski & Dallob, 1995; Kaplan & Simon, 1990), this ordering may be important since a restructured problem space may analogically trigger access to different source analogs. The present study was a careful time-course analysis of 'reverse' analogical problem solving using video recordings of the problem solving process. By using this non-standard 'reverse' analogical problem solving design we were able to code for impasse, incubation, spontaneous access and mapping in the same design.

In the sections that follow, we will briefly review the relevant pieces of the analogical transfer literature and then the incubation literature.

Analogical Transfer: Access, retrieval and noticing

A number of factors have been found to influence spontaneous analogical access. Access is improved when the source is presented as a problem rather than as declarative knowledge (Adams, Kasserman, Yearwood, & Perfetto, 1988; Lockhart, Lamon, & Gick, 1988), if explicitness of the cross-domain relational similarity is increased (Clement, Mawby, & Giles, 1994), or if experimental conditions stress a rich encoding and structural information at retrieval (Dunbar, 2001). Furthermore, it has been found that the presence of salient superficial similarity increases spontaneous access (e.g., Holyoak & Koh, 1987; Gentner, Ratterman & Forbus, 1993).

A large number of studies have found that providing subjects with an analogous source prior to showing the target rarely leads to significant access (or mapping) of the source, unless explicit instructions are provided to make use of previous information (i.e., subjects are informed of the potential relevance of previous problems during the instructions), or if the source is very similar to the target (see e.g., Anolli et al., 2001 for a brief review). Anolli et al. (2001) argued these findings suggested that analogical access is not a spontaneous process. However, as Ross, Ryan & Tenpenny (1989) noted, studies have not shown that people never spontaneously access relevant information, but only that in cases where they are expected to do so, they often do not. Trying to measure spontaneous access on uninformed subjects is a methodological problem that previous studies have struggled with. Typically either informed subjects are used in a paradigm where the subjects are asked to write down any previous problems they access (e.g., Gentner, Rattermann, & Forbus, 1993), or performance measures are used (i.e., correct or incorrect solution), which is a less optimal measure of access since some subjects may have accessed the source without being able to map between source and target (e.g., Gick & Holyoak, 1980). An exception was Ross (1987) who used correct/incorrect performance to measure mapping, but used a different score (whether the subject *appeared to be making use* of a source formula on the answer sheet) to measure access. This measure enables a separation of access and mapping in the same design, but still runs the risk of underestimating access. For example, subjects may access a source without making use of it on the answer sheet, because they do not consider it relevant.

As Ross et al. (1989) argued, an important aspect of measuring spontaneous access in a within-subject design is that the subject should not be aware of the relevance of the source analogues (i.e. be informed). When measuring spontaneous access in a within-subject design it is thus necessary for the researcher to demonstrate that subjects did not become aware of the relevance (i.e., catch on, thus becoming informed) after they had accessed relevant sources. This result led Ross et al. (1989) to conclude that in order to avoid having to deal with an unknown mixture of informed and uninformed cases, access issues need to be studied with informed subjects.

Efforts were made in the present study to design a novel way of measuring spontaneous access on uninformed subjects, while avoiding the above pitfalls, by separating access and mapping measures in a within-subject design while controlling for whether the subjects had caught on.

A note on terminology: Ross (1989) has made a distinction between access, noticing and retrieval of a source. Access involves both noticing and retrieving a source. Where the noticing involves recognizing or selecting a particular earlier example, retrieval concerns the actual remembrance of this earlier example. Further, the analogical access literature have used terms like ‘spontaneous’ (e.g., Holyoak & Koh, 1987) and ‘automatic’ (e.g., Anolli et al., 2001) for experimental conditions where the subject is uninformed about relevance of the source information. In the present study, the term ‘immediate’ access refers to access of the relevant unsolved target problem immediately following presentation of an analogous source problem. The term ‘spontaneous immediate’ access is used to refer to immediate access where the subjects are also uninformed about the relevance of the source information for solving the problem, and where they have not yet caught on to this fact. The term ‘automatic’ is not used in order to avoid suggesting a constant law or high frequency of access because this suggestion seems unwarranted given the hard time researchers have had finding evidence for this phenomenon.

Incubation as Reversed Analogical Transfer

The ‘reversed’ analogical problem solving design (i.e., target-source) is identical with incubation effect study designs, and thus allows for testing of incubation effect hypotheses. The incubation effect refers to the phenomenon that spending time away from a problem may be beneficial to performance, compared to continuous problem solving. Incubation starts off with an individual, having reached an impasse on a problem, sets the problem aside. During incubation, no conscious problem solving appears to be going on, and individuals go about their other unrelated business. At some point the individual will return to problem solving, often described as occurring suddenly and surprisingly, in a flash of ‘insight’, where the solution is readily evident. Experimental studies of this phenomenon have received mixed results. An early review (Olton, 1979) questioned the existence of the incubation phenomenon, but a more recent review (Dodds, Ward, & Smith, in press) argues that recent studies show more promise.

Incubation theories can be divided into two groups: autonomous incubation refers to theories hypothesizing that the effect is due to some phenomenon relying only on the passage of time, while interactive incubation theories require that new relevant information must be available in the environment (Dorfman, Shames, & Kihlstrom, 1996). Autonomous theories of incubation have for example attributed the phenomenon to unconscious idea generation (e.g., Campbell, 1960), selective forgetting (Simon, 1966), and forgetting of fixating elements (Smith, 1995; Smith & Blankenship, 1991). However, the interactive theory of Opportunistic Assimilation (OA) (Seifert et al., 1995; Patalano & Seifert, 1994) tries to explain incubation effects with chance (e.g., analogical) cuing from the environment. The theory states that reaching an impasse on a creative problem will encode 'failure indices' in memory, which will be activated through automatic spreading of activation by standard perception and comprehension processes if the subject later encounters an analogous solution or solution element in the environment, thus bringing the previously unsolved problem back into awareness for new solution attempts. These 'failure indices' constitutes a form of predictive encoding which allows the individual to recognize later opportunities to achieve pending goals (Patalano & Seifert, 1997; Seifert, 1994), much like it occurs in everyday planning. In this way, the theory ties in with the Zeigarnik (1927) effect in that it hypothesizes that memory for unsolved problems is better than for solved problems, and with Louis Pasteur's notion that 'chance favors the prepared mind' (e.g., Posner, 1973). OA thus predicts that seeing relevant information in the environment will trigger spontaneous access to previously unsolved analogous problems. Furthermore, this access is predicted to increase performance on the previously unsolved problem. Both of these hypotheses can be tested using the 'reverse' analogical problem solving paradigm.

On testing the importance of impasses in OA theory, Seifert et al. (1995) found that only when subjects had previously reached an impasse in answering difficult factual questions, and when they were later followed by incidental relevant information, did subjects improve performance. Patalano

& Seifert (1994) investigated the memorability of solved and unsolved problems following impasse, and found this effect depended on set size. Only when impasses are infrequent (i.e., when set size is small), are unsolved problems more available in memory than solved problems.

Few studies have tested for interactive incubation effects, and with mixed results. An early study by Dreistadt (1969) showed that the presentation of visual analogical cues during incubation on insight problems had a significant effect on performance when compared to subjects working continuously on the problem. However, Olton & Johnson (1976) failed to replicate this result, and Browne & Cruse (1988) using one of Dreistadt's insight problems found incubation effects for the experimental group receiving cues in only one of their two experiments. Dominowski & Jenrick (1972) found incubation effects for providing verbal cues on the hat-rack problem. Remote Associates Test (RAT) items (Mednick, 1962) have been used in several studies; Mednick et al. (1964, exp II) found that subjects performed better on items where they were primed with cues during incubation than when they were not. Using a variation of RAT items, Dorfman (1990) found mixed results depending on whether the incubation task consisted of working on other similar problems (yielding an interactive incubation effect) or working on an unrelated arithmetic task (yielding an autonomous incubation effect). In a large study, Dodds, Smith and Ward (2002) tested whether the presentation of associates of the answer would automatically lead to performance increases following initial solution attempts on RAT items, but no significant effect was found. It should be noted that, in all RAT studies, associates were used as cues, which is not the same as analogues.

In conclusion, studies testing for interactive incubation effects have yielded mixed results when using visual analogies for insight problems and RAT items. However, as noted above, two studies have found incubation effects when providing analogous stories (Gick et al., 1980, exp V) or solutions to analogous problems (Keane, 1985) during incubation. But both these studies suffer from shortcomings. Gick and Holyoak (1980, exp V) did not provide distracter tasks (only the Attack-Dispersion story was read during incubation). Adding distracter tasks could potentially have influenced spontaneous transfer, as indicated by their experiment IV, and possibly reduced spontaneous transfer to insignificant amounts given the small sample size used (10 subjects in the control group, and 20 subjects in the incubation condition). Keane (1985) did provide distracter tasks, but he also instructed subjects that the stories they read might be relevant in solving the problem, thereby informing subjects of the relevance of the cue. In summary, it has yet to be established what (if any) kinds of creative problems and cues will show reliable interactive incubation effects for uninformed subjects. Moreover, none of these reverse analogy studies (nor in fact the majority of traditional analogy studies) have looked at the process of spontaneous access and success of analogical transfer in the same design. By using video in a careful time-course analysis the current study measured impasse, incubation, analogical access, and analogical mapping separately in the same design. The study tested whether encountering analogous solutions to previously unsolved problems would lead to spontaneous immediate access of the unsolved problem, and performance increase through analogical mapping on insight problems (see Figure 1). Further, it was tested whether initial encoding time of the problem or incubation time predicted immediate retrieval of the unsolved problem upon seeing the relevant analogue. Finally, categories that are typically not measurable in standard analogical problem solving designs were examined; in particular we examined problems that are immediately accessed, but where no performance increase is found, and problems that are not immediately accessed, but where performance is nonetheless increased once the subject returns to the problem. These categories are typically not measurable since they require a separation of access and mapping measures in the same design.

In sum, we examined whether analogous cues do produce immediate access to related unsolved problems, and then whether access produce successful analogical mappings (thereby accounting for at least one source of incubation effects). It is possible that analogous cues do not produce immediate access, but instead produces analogical mappings only when people happen to return to the unsolved problem later. For cases with low surface similarity, this result could be expected from the frequent failures to find spontaneous analogical access with low surface similarity. It is also possible that analogous cues do produce immediate access of related unsolved problems, but that analogical mappings are not successfully made with high frequency (in which case OA is a poor account of incubation effects).

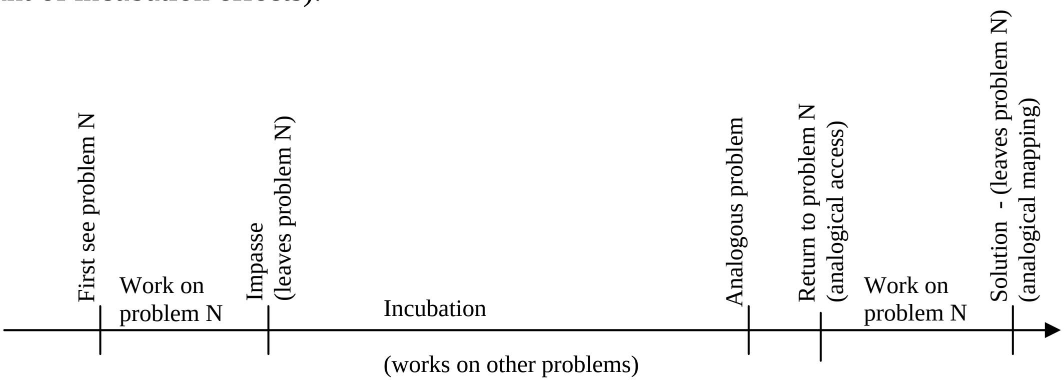


Fig. 1. Problem solving timeline on individual problems

Methods

Subjects

Forty undergraduates (12 female) from the University of Pittsburgh subject pool participated for course credits.

Materials

Subjects worked on 8 insight problems adapted from the literature (see appendix A). Each problem was presented on a separate page. Eight analogous insight problems were constructed as cues (analogous cues), along with four unrelated insight problems (distracter tasks). These were presented as rating tasks, where the subjects had to rate the problem for difficulty on a scale from 1 to 5. These rating tasks were constructed so the problem was presented on the front, and the solution to the problem and the rating scale on the back of a sheet of paper. Rating tasks and problem order were varied across subjects. Each subject received four analogous and four distracter rating tasks.

The present experimental materials were constructed to minimize superficial similarity between source and target as much as possible. However, in designing the materials we found that the distinction between superficial similarity and structural similarity is somewhat subjective, and may be a continuum rather than a dichotomy. We found that it is always possible for a motivated person to find some superficial similarity between problems that are supposed to contain 'only' structural similarity. For example, in the present materials, the analogue we constructed to the Tennis Tournament problem was a Viking battle where Vikings fight each other to go to Valhalla (see appendix A). This analogue was constructed to minimize story line similarity, leaving mainly structural similarity. However, some of the concepts used are somewhat interchangeable, although different words are applied; for example the words 'match' as in 'tennis match', and the word 'fight' in the Viking story basically points to the same underlying construct even though different words are applied. In some cases, these words can be used interchangeably, as in 'boxing match' and 'boxing fight'. Does this correspondence mean that the presence of 'match' in the tennis story and the word 'fight' in the Viking story constitutes structural or superficial similarity? We suggest it has elements of both; the match/fight construct serves to secure a structurally relevant mapping, but in so far as the match/fight concepts are interchangeable (in a given problem context), they can also be characterized as somewhat superficially similar. To take another example from the Viking story, the words 'tournament' and 'battle' are again used as analogous concepts. In this case, however, the tournament (in the context of a tennis match) has additional constraints attached to it, compared to the 'battle' concept. Tournaments involve a strict hierarchical structure, where the winner from one match moves to the next ‘level’ in the tournament. This strict hierarchy is not an essential part of the ‘battle’ concept, and thus this constraint had to be added to the Viking story to make the analogy a tight one. Thus, in this case, it has to be inferred by the problem solver, that ‘battle’ is analogous to ‘tournament’ if this constraint is added. In this case, we argue the similarity between ‘battle and ‘tournament’ is more structural than superficial.

Our materials appear to sit somewhere in the middle of the continuum between superficial and structural similarity. However, even though some amount of superficial similarity may be present in our materials, this does not make our results uninteresting to the ‘analogical access’ literature. As Hammond, Seifert, & Gray (1991) argued, memory serves a function, and theories and experiments on memory should include (rather than exclude) those sets of useful features that are already part of the tasks that memory serves. Superficial similarity is one such feature. Furthermore, as briefly reviewed in the introduction, interactive incubation studies have had problems finding reliable results, which would make any positive incubation effect interesting – regardless of whether it is triggered in part by superficial similarity. Finally, problem-cue pairs were designed to minimize superficial similarity, and our stimuli were more typical of the “structural-only” similarity found in source-target conditions that fail to find spontaneous analogical transfer than typical of the “superficial” similarity conditions that find spontaneous analogical transfer.

Procedure

Subjects read instructions with a cover story about being the editor of ‘Puzzle Magazine’ to prevent subjects from expecting a relationship between source and target problems. They were informed that they would be working on 8 puzzles, but that they would have to rate other puzzles for difficulty (on a scale from 1 to 5) as well. Their secretary (the experimenter) would from time to time put a puzzle (i.e., either an analogous cue or distracter cue) on their desk, and they should immediately read through it (including the answer on the back), rate it for difficulty and hand it back. It was stressed that they should work fast, and that they could return to previous puzzles at any time.

The experiment was carried out on one subject at a time. Subjects were given 45 minutes to solve all 8 problems. During this time they could freely move between problems. Progress on the problems was videotaped. Every 5 minutes the subjects were handed an analogous or distracter rating task. These tasks were handed to the subjects at a time when they had already worked on the relevant problem, and moved on (i.e. after reaching a self-generated impasse). Through a computer linked to the video camera, the experimenter kept track of which problems had been worked on and left (i.e., where the subject had reached an impasse on the problem), and which problem the subject was currently working on. This procedure made it possible to constantly update a pool of available tasks, wherefrom a task was randomly selected (using random number tables) and handed to the subject every 5 minutes. When the 45 minutes had elapsed, the subject was asked to fill out questionnaires about memory for the problems and awareness of the rating task/problem relationship. Impasse was coded as the time when a subject left a problem after having first worked on it, without a correct solution. Access was coded based on which page was turned to next in the booklet. A correct access was defined as returning to the relevant previously unsolved problem as the very next problem following the analogous rating task. Mapping was coded by performance (i.e., correct or incorrect solution). Time spent working on individual problems, both before and after reaching an impasse, was also measured. See Figure 2 for an example of typical problem solving behavior.

Time Problem	Cue	Subject behavior	Provides solution?	Correct?	Immediate retrieval?
0:00		Opens booklet and turns page to Problem #1	No		
2:47		Turns page to Problem #2	No		
4:39		Turns page to Problem #3	Yes	No	
	5:10	Is handed irrelevant cue			
	6:03	Hands cue back			
6:54		Turns page to Problem #4	No		
8:08		Turns page to Problem #5	Yes	Yes (excluded)	
8:36		Turns page to Problem #6	Yes	Yes (excluded)	
	10:12	Is handed relevant cue for Problem #3			No
	10:41	Hands cue back			
10:56		Turns page to Problem #7	Yes	No	
14:48		Turns page to Problem #8	Yes	Yes (excluded)	
	15:12	Is handed irrelevant cue			
	16:21	Hands cue back			
		ALL PROBLEMS SEEN			
19:23		Turns page to Problem #3	Yes	Yes	
	20:15	Is handed irrelevant cue			
	20:28	Hands cue back			
21:20		Turns page to Problem #4	No		
22:50		Turns page to Problem #2	Yes	Yes	
24:55		Turns page to Problem #1	No		
	25:10	Is handed relevant cue for Problem #5 (already solved)			(excluded)
	25:44	Hands cue back			
29:18		Turns page to Problem #4			
	30:18	Is handed relevant cue for Problem #1			Yes
	31:17	Hands cue back			
31:18		Turns page to Problem #1	Yes	Yes	
34:40		Turns page to Problem #4	No		
	35:21	Is handed irrelevant cue			
	35:34	Hands cue back			
	40:14	Is handed relevant cue for Problem #7			Yes
	40:40	Hands cue back			
40:50		Turns page to Problem #7	No		
41:13		Turns page to Problem #4	Yes	No	
45:01		Time elapsed – experimenter removes booklet			

Fig. 2. Example of problem solving behavior: Subject 35

Results

Successful Analogical Mapping

Subjects solved 44% of the problems. Of the solved problems, 39% were solved prior to impasse and the remaining 61% were solved after returning to the problem. Problems solved prior to impasse were excluded from further analysis. A resolution score (e.g., Smith et al., 1991 - see appendix B) was calculated for each subject for the analogous-cued problems and the distracter-cued problems respectively. The resolution score calculates the proportion of impassed problems that were eventually solved. Mean per subject resolution scores were .53 and .17 for the analogous-cued and distractor-cued problems respectively (standard errors were .05 and .04 respectively). Analogous cues produced noticeably higher resolution scores ($t(39)=5.62$, $p<.001$), showing that the cues were highly effective and produced a large effect over a control of the distracter-cued problems that were initially unsolved and left. It could be argued that this comparison does not directly measure a traditional incubation effect as such because no control group working continuously was included; perhaps the distracter-cued problems would also have an incubation effect (of the autonomous rather than interactive kind) had they been compared to a control group working continuously. What it does demonstrate, however, is that an incubation effect based on new relevant information from the environment is significantly larger than any potential incubation effects based on autonomous cognitive processes alone (.36 vs. .17 at most). Our apparent interactive incubation effect could still potentially be explained away if the participants spent more time on the analogous-cued problems after returning to the problems. Bowden (1985) had similar objections to the original Gick & Holyoak (1980) study. To test for this possibility, we measured how much time subjects spent working on distracter-cued and analogous-cued problems after having first reached an impasse on the problem (i.e., after leaving the problem when having first worked on it). Subjects spent almost 40% longer (not shorter) after impasse on distracter-cued problems than analogous-cued problems ($t(39)=2.17$, $p<.04$), thus ruling out the possibility that subjects had simply worked longer on the analogous-cued problems after impasse.

Analogical Access

Immediate analogical access was measured by examining (using video) which problem the subject would turn to in the booklet as the next problem following the presentation of an analogous cue. On average, subjects returned immediately 50% of the time to the relevant problem. It could be argued that this 50% should be compared to a theoretical 1/7 chance of going to any particular problem next. However, we used a more accurate baseline based upon typical page-turning behavior following the *distracter* tasks. Two strategies were identified in subjects page-turning. Before the subjects had seen all problems, a dominant strategy was simply to turn to the next unseen problem in the booklet. This strategy occurred 96% of the time. The time until subjects had seen all problems in the booklet and moved on to other problems was on average 21 minutes and 31 seconds. After the subjects had seen all the problems, they appeared to be employing a strategy of locating a previously unanswered problem (i.e., a problem where they had not yet indicated any solution in the answer section of the page) to work on. Following any given distracter task after seeing all problems, subjects would on average have 33% unanswered problems in the booklet – but they would return to these 68% of the time. This information was used to test whether the subjects were returning to the relevant previously unsolved problems significantly more often than their usual behavior on the distracter tasks would predict.

For each subject, an immediate access rate and baseline was calculated (see appendix B) BEFORE and AFTER all problems had been seen (see table 1). In both cases, access rates were significantly above baseline.

Table 1

Mean immediate access rates and baseline calculations before and after seeing all problems

	N	Mean	S.E.	t	*p*
Access rate BEFORE seeing all problems	25	.26	.08	3.04	<.007
Baseline BEFORE seeing all problems	25	.01	.00		
Access rate AFTER seeing all problems	39	.64	.06	7.56	<.001
Baseline AFTER seeing all problems	39	.17	.01		

This immediate access measure does not take into account that the subjects may become informed (i.e., catch on to the relevance of the analogous cues during problem solving). Being informed should be taken into account when trying to estimate *spontaneous* immediate access. Immediate access rates were higher after seeing all problems than before (t(23)=3.49, *p*<.003), which could suggest that the initial dominant strategy of turning to the next unseen problem depressed the access rates. But it may also indicate that participants caught on during problem solving. To examine this issue further, access rates were divided into a 2x2 depending on whether subjects were expected to have become informed, and whether they had seen all problems (see table 2). Subjects were scored as ’informed’ if they had returned immediately to a relevant problem at least once before. Problems that were solved before impasse were once again excluded from this analysis. A small number of problems where the subject had not yet immediately accessed a relevant problem, but where it was indeterminable whether the subjects had caught on, because they had received analogous cues to problems they had already solved, where excluded from this analysis.

Table 2

Access rates by being informed and page-turning strategy (whether subject had seen all problems in the booklet)

	Not seen all	Seen all
Not caught on	.24 (n=34)	.46 (n=41)
Caught on	.60 (n=5)	.78 (n=36)

Note. In the figure, n refers to number of problems in each category across subjects. Excluded are problems where it was indeterminable whether the subject had caught on, and problems solved pre-impasse

Comparing rows in table 2 illustrates that there is an effect of being informed ($\chi^2(1)=16.64$, $p<.001$). Comparing columns in Table 2 shows that there is an effect of page-turning strategy (lower retrieval rates prior to seeing all problems) ($\chi^2(1)=13.62$, $p<.001$).

To estimate spontaneous immediate retrieval, we excluded all problems where the subject was expected to have caught on, and calculated a mean access rate and baseline per subject following the same procedure as for immediate access measures. The results were still highly significant for both before ($t(24)=3.10$, $p<.006$) and after ($t(23)=5.78$, $p<.001$) seeing all problems in the booklet. Means before seeing all problems were .28 and .01 for the baseline; means after seeing all problems were .66 and .16 for the baseline. Note that these means are within subject, and therefore differs from the means presented in Table 3 which is across subjects. These results indicate that when spontaneous immediate access to analogous unsolved problems is measured against a control of the same person's problem solving behavior following distracter tasks, a large effect is found, and when controlling for subjects who become informed during the experiment.

The results show that when subjects were given relevant cues they deviated from their typical page-turning strategies in order to return to the relevant problem in order to try to utilize the new information to solve that problem. They did not randomly look through the booklet in order to try to find a match for the cue, but rather went straight to the relevant problem, suggesting that becoming informed did not change behavior in the absence of access. Using the same subjects' typical page-turning behavior following distracter tasks as a control measure, and controlling for subjects who become informed, thus revealed an immediate access effect. It was also found, as it has been many times in the literature, that the immediate access rate was larger for informed subjects than for uninformed subjects.

What Factors Predict Immediate Retrieval?

Analogous cues did not lead to 100% immediate retrieval. Examining what factors predicted immediate retrievals can provide further information on the mechanisms of analogical retrieval in insight problems. From a simple memory perspective, one would predict that encoding time on the impassed problems would predict how easily they are retrieved upon cuing. To test whether encoding time predicted immediate access, we measured time spent working on each problem for each subject (i.e., how much time was spent working on each individual page in the booklet) prior to receiving the analogous cue. Then we conducted a within-subject t-test comparing encoding time for problems that were immediately accessed, with those that were not. The result was significant ($t(26)=3.81$, $p<.002$), showing that encoding time did predict immediate retrieval. Means for immediate access and non-immediate access were 3 min 49 sec and 2 min 12 sec respectively.

From the perspective of autonomous incubation effects, time away from the impassed problem might predict degree of immediate access. We tested whether time away from the problem (incubation time) predicted immediate access, by measuring the time from last leaving the problem to presentation of the analogous cue. Similar to the encoding time analysis, we then compared incubation time for problems that were immediately accessed, with those that were not. This comparison was not significant (t(26)=-.57, *p*>.57). Means for immediate access and non-immediate access were 5 min 49 sec and 6 min 35 sec respectively.

Is Immediate Analogical Mapping Different from Later Returns?

Since both access and mapping were measured separately in the same design, it was possible to calculate whether resolution scores for the immediate and non-immediate returns differed, and differed from resolution scores for the problems receiving only distracter tasks (see Table 3).

Table 3

Mean per category resolution scores by whether subject did or did not immediately return to the relevant problem upon seeing an analogous cue.

	Analogous-cue		Distractor task
	Immediate return	Non-immediate return	
Resolution score	.67 (n=64)	.32 (n=62)	.16 (n=135)

Note. n refers to number of problems in each category. All means differed significantly pairwise (.001<*p*<.012)

For problems where the subjects accessed the relevant problem immediately there was a higher resolution score than for either the non-immediate returns, or the problems receiving only distracter tasks. But on problems where the subject did not immediately access the relevant problem, they still had a significant effect of relevant cuing on performance compared to the problems receiving distracter task. This result could potentially be accounted for if the non-immediate returns benefited performance by a process similar to semantic priming (see e.g., Schunn & Dunbar, 1996 for another analogical transfer study finding a priming effect). In semantic priming subjects are unaware that their performance is increased by relevant information. Because we measured awareness of each problem-cue relationship after the experiment, we were able to test for this possibility. In the non-immediate return category subjects reported being aware of the relationship between the problem and the cue for 17 of the 20 (i.e., 85%) problems that were solved. This result gives a clear indication that once mapping took place, subjects were explicitly aware of the problem/cue relationship. So semantic priming without explicit awareness appeared not to be the cause.

Discussion

The novel design developed in the present experiment showed it was possible to separate a number of constructs for each problem in a within-subject design using video. These constructs, which are typically only manipulated across experimental conditions, include encoding time, impasse, incubation time, analogical access, and analogical transfer, along with measures of becoming informed and problem solving strategies. The careful time-course analysis used here thus has the potential for informing a number of theories in the analogical problem solving and incubation/insight literature

The present experiment was able to show that analogical cuing lead to immediate access, when controlling for becoming informed, by using the subjects' typical problem solving behavior following distracter problems as a control.

The results suggest that perhaps the failure of much of the previous analogical transfer literature to find such a spontaneous transfer effect may in part be attributable to their use of comparison groups. Previous studies have often showed that when comparing informed and uninformed subjects, the informed subjects accessed and transferred more frequently than the uninformed. But perhaps this comparison is unfair when trying to establish whether and how frequently spontaneous immediate access occurs. A large effect of analogous cuing on spontaneous immediate access was thus found when using the same subjects' typical problem solving behavior following distracter tasks as a baseline. The present experiment also showed, as has often been noted, that when subjects became informed (i.e., caught on) they accessed more frequently. Thus, there appears to be a spontaneous immediate access effect of receiving analogous cues, but an even larger effect of being informed. This latter effect of being informed should not, however, lead one to assume that no spontaneous access effect is present even though it may appear small in comparison to the effect of being informed (as e.g., Anolli et al., 2001 seems to do).

Encoding time predicted immediate access, thus supporting Dunbar's (2001) hypothesis that encoding time of the analogue to be accessed is extremely important, and may help explain why some researchers have failed to find transfer effects. Incubation time, however, did not predict immediate access suggesting that autonomous incubation effects played little role in this study.

The Opportunistic Assimilation (Seifert et al., 1995) theory of incubation was also supported by this spontaneous immediate access effect, and OA was further supported by the result that receiving an analogous problem during incubation (after impasse) increased performance through analogical mapping. These results lend support to the OA theory that - following impasse on creative problems – a chance encounter with analogous relevant new information in the environment will (at least in some cases) spontaneously trigger access and mapping to the previously unsolved problem. Thus, presenting analogous problems during incubation is at least one reliable way of demonstrating incubation effects, unlike visual analogies and associates in RAT problems, which, as noted in the introduction, have shown mixed results. On at least some analogical problems, chance does appear to favor the prepared mind.

Several more experiments should be conducted to further test OA theory. For example, the theory would predict that spontaneous immediate access rates should increase after impasse. This prediction could potentially be tested by presenting cues after impasse (as it is done in this experiment) and comparing them with other problems where the cue is presented before impasse (possibly before seeing the target as it is done in standard analogical problem solving). This design is essentially the same variation that was done in the original Gick & Holyoak (1980, exp V) study. They found no significant difference, but this result should be retested using a variation of the present design in a larger study.

From the retrospective reports, subjects were in most cases explicitly aware of the problem-cue relationship. This result means that a limitation of the present study is that it cannot purport to explain incubation phenomena where subjects are unaware of how they solved the problem. Much anecdotal evidence concerns individuals who later reported that an idea just simply occurred to them when they for example were in the bath or sitting on a bus. Unless a mechanism can be identified that would explain why people later often do not report that they used cues from the environment in their solution of the problem, OA may not be able to explain these cases. However, it could be accounted for by the standard memory mechanism of differential rehearsal or retrieval practice (e.g., Bjork, 1988). Perhaps when a creative discovery or invention is rehearsed over the years through repeated retelling of the problem and solution by its creator, memory of how the problem was originally solved is not rehearsed as frequently. In this case, standard forgetting processes in addition to OA may account for the phenomenon.

Finally, the present experiment may seem to lack ecological validity, and may appear somewhat contrived and artificial. As other studies on analogy have shown (e.g., Dunbar, 2001; Blanchette &

Dunbar, 2001) this difference between experimental settings and real-world context is sometimes very important and may lead to different results. Therefore it remains to be seen whether more naturalistic research will be able to find evidence of these spontaneous access and incubation effects. We are presently collecting data on design problem solving using Dunbar's (1995; 1997) in vivo methodology to answer this question.

Acknowledgements

The authors would like to thank Kwangsu Cho, Anthony Harrison, Xiaohui Kong, and Mark McGregor for helpful comments on an earlier draft. This work was supported by Grant N00014-01-1-0321 from the U.S. Office of Naval Research to the second author.

References

Adams, L. T., Kasserman, J. E., Yearwood, A. A., & Perfetto, G. A. (1988). Memory access: The effects of fact-oriented versus problem-oriented acquisition. *Memory and Cognition, 16*, 167-175.

Anolli, L., Antonietti, A., Crisafulli, L., & Cantoia, M. (2001). Accessing source information in analogical problem-solving. *Quarterly Journal of Experimental Psychology: Human Experimental Psychology, 54A*, 237-261.

Bjork, R. A. (1988). Retrieval practice and the maintenance of knowledge. In M.M.Gruneberg, P. E. Morris, & R. N. Sykes (Eds.), *Practical aspects of memory: Current research and issues* (pp. 396-401). Chichester, UK: John Wiley & Sons.

Blanchette, I. & Dunbar, K. (2001). Analogy use in naturalistic settings: The influence of audience, emotion, and goals. *Memory and Cognition, 29,* 730-735.

Bowden, E. M. (1985). Accessing relevant information during problem solving: Time constraints on search in the problem space. *Memory & Cognition*, *13,* 280-286.

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

Campbell, D. T. (1960). Blind variation and selective retentions in creative thought as in other knowledge processes. *Psychological Review, 67,* 380-400.

Chen, Z. (1995). Analogical transfer: From schematic pictures to problem solving. *Memory and Cognition, 23:* 255-269.

Clement, C. A., Mawby, R., & Giles, D. E. (1994). The effects of manifest relational similarity on analog retrieval. *Journal of Memory and Language*, *33*, 396-420.

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

Dodds, R. A., Ward, T. B., & Smith, S. M. (in press). A review of the experimental literature on incubation in problem solving and creativity. In M.A.Runco (Ed.), *Creativity research handbook* (3rd ed.). Cresskill, NJ: Hampton Press.

Dominowski, R. L. & Dallob, P. (1995). Insight and problem solving. In R.J.Sternberg & J. E. Davidson (Eds.), *The nature of insight.* (pp. 33-62). Cambridge, MA: The MIT Press.

Dominowski, R. L. & Jenrick, R. (1972). Effects of hints and interpolated activity on solution of an insight problem. *Psychonomic Science,* Mar-338.

Dorfman, J. (1990). *Metacognitions and incubation effects in insight problem solving.* Unpublished dissertation for the degree of Doctor of Philosophy University of California, San Diego.

Dorfman, J., Shames, V. A., & Kihlstrom, J. F. (1996). Intuition, incubation, and insight: Implicit cognition in problem solving. In G.D.M.Underwood (Ed.), *Implicit cognition* (pp. 257-296). New York,NY,US: Oxford University Press.

Dreistadt, R. (1969). The use of analogies and incubation in obtaining insights in creative problem solving. *Journal of Psychology, 71,* 159-175.

Dunbar, K. (1995). How scientists really reason: Scientific reasoning in real-world laboratories. In R.J.Sternberg & J. E. Davidson (Eds.), *The nature of insight.* (pp. 365-395). Cambridge, MA, US: The MIT Press.

Dunbar, K. (1997). How scientists think: On-line creativity and conceptual change in science. In T.B.Ward, S. M. Smith, & J. Vaid (Eds.), *Creative thought: An investigation of conceptual structures and processes.* (pp. 461-493). Washington, DC, US: American Psychological Association.

Dunbar, K. (2001). The analogical paradox: Why analogy is so easy in naturalistic settings yet so difficult in the psychological laboratory. In D.Gentner, K. J. Holyoak, & B. N. Kokinov (Eds.), *The analogical mind: Perspectives from cognitive science.* (pp. 313-334). Cambridge, MA: The MIT Press.

Duncker, K. (1945). *On Problem-solving.* Westport, CT, US: Greenwood Press, Publ.

Forbus, K. D., Gentner, D., & Law, K. (1994). MAC/FAC: A model of similarity-based retrieval. *Cognitive Science, 19,* 141-205.

Gentner, D., Rattermann, M. J., & Forbus, K. D. (1993). The roles of similarity in transfer: Separating retrievability from inferential soundness. *Cognitive Psychology, 25:* 524-575.

Gick, M. L. & Holyoak, K. J. (1980). Analogical problem solving. *Cognitive Psychology, 12,* 306-355.

Gick, M. L. & Holyoak, K. J. (1983). Schema induction and analogical transfer. *Cognitive Psychology, 15,* 1-38.

Hammond, K. J., Seifert, C. M., & Gray, K. C. (1991). Functionality in analogical transfer: A hard match is good to find. *Journal of the Learning Sciences, 1:* 111-152.

Holyoak, K. J. & Koh, K. (1987). Surface and structural similarity in analogical transfer. *Memory & Cognition, 15,* 332-340.

Hummel, J. E. & Holyoak, K. J. (1997). Distributed representations of structure: A theory of analogical access and mapping. *Psychological Review, 104,* 427-466.

Kaplan, C. A. & Simon, H. A. (1990). In search of insight. *Cognitive Psychology, 22:* 374-419.

Keane, M. (1985). On drawing analogies when solving problems: A theory and test of solution generation in an analogical problem-solving task. *British Journal of Psychology, 76,* 449-458.

Lockhart, R. S., Lamon, M., & Gick, M. L. (1988). Conceptual transfer in simple insight problems. *Memory and Cognition, 16,* 36-44.

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

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

Ohlsson, S. (1984). Restructuring revisited: I. Summary and critique of the Gestalt theory of problem solving. *Scandinavian Journal of Psychology, 25:* 65-78.

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

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

Patalano, A. L. & Seifert, C. M. (1994). Memory for impasses during problem solving. *Memory and Cognition, 22:* 234-242.

Patalano, A. L. & Seifert, C. M. (1997). Opportunistic planning: Being reminded of pending goals. *Cognitive Psychology, 34:* 1-36.

Perkins, D. N. (2000). *Archimedes' bathtub The art and logic of Breakthrough Thinking.* New York, NY: WW Norton & Company.

Posner, M. I. (1973). *Cognition: An introduction.* Glenview, Il: Scott, Foresman & Co.

Ross, B. H. (1987). This is like that: The use of earlier problems and the separation of similarity effects. *Journal of Experimental Psychology: Learning, Memory, & Cognition, 13,* 629-639.

Ross, B. H. (1989). Distinguishing types of superficial similarities: Different effects on the access and use of earlier problems. *Journal of Experimental Psychology: Learning, Memory, & Cognition, 15,* 456-468.

Ross, B. H., Ryan, W. J., & Tenpenny, P. L. (1989). The access of relevant information for solving problems. *Memory & Cognition, 17,* 639-651.

Schunn, C. D. & Dunbar, K. (1996). Priming, Analogy, & Awareness in complex reasoning. *Memory and Cognition, 24,* 271-284.

Seifert, C. (1994). The role of goals in retrieving analogical cases. In J.A.Barnden & K. J. Holyoak (Eds.), *Analogy, metaphor, and reminding.* (pp. 95-168). Norwood, NJ, US: Ablex Publ. Coorporation.

Seifert, C. M., Meyer, D. E., Davidson, N., Patalano, A. L., & Yaniv, I. (1995). Demystification of cognitive insight: Opportunistic assimilation and the prepared-mind perspective.

In R.J.Sternberg & J. E. Davidson (Eds.), *The nature of insight.* (pp. 65-124). Cambridge, MA: The MIT Press.

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

Simon, H. A. (1966). Scientific discovery and the psychology of problem solving. In R.G.Colodny (Ed.), *Mind and Cosmos* (pp. 22-41). Pittsburgh, PA: 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. 135-156). Cambridge, MA: The MIT Press.

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

Thagard, P., Holyoak, K. J., Nelson, G., & Gochfeld, D. (1990). Analogical retrieval by constraint satisfaction. *Artificial Intelligence, 46,* 259-310.

Zeigarnik, B. (1927). Das Behalten von erledigten und unerledigten Handlungen. *Psychologie Forschung, 9*, 1-85.

Appendix A. Problems used.

The insight problems used included slightly changed versions of the elephant problem (Chen, 1995), the ten trees problem (Dreistadt, 1969), the tennis tournament problem (Ohlsson, 1984), the sports car racers, 10 matches, shot, and fly problems (all Perkins, 2000), the bracelet problem (Silveira, 1971), as well as other analogical insight puzzles that were generated with elements taken from puzzle books and puzzle sites on the internet. Two detailed examples of a problem and cue are presented below.

PROBLEM: TENNIS TOURNAMENT

A singles tennis tournament is being planned that will involve 88 players. The tournament will be single elimination, so any player who loses a match is out of the tournament and the winner will be the person who doesn't lose at all. The organizers of the tournament need to know how many matches will be played so they can budget for court time and balls. How many matches will be played in all during the tournament?

CUE: THE VIKINGS

> The Vikings believed that by dying on the battlefield they would be allowed a place in Viking heaven (Valhalla). Forty-four Vikings are assembled to battle each other until only one is left standing. However, even Vikings have rules; the Vikings fight one-on-one at a time, until one of them is dead. Once the battle begins all Vikings will start fighting with the nearest other Viking. The Viking who is still alive then moves on to find a new Viking to fight, while the loser goes to Valhalla. And so on. What is the number of one-on-one Viking fights there will be in this battle?
>
> ***Answer:*** The number of one-on-one Viking fights will necessarily equal the number of dead Vikings (i.e., Vikings that go to Valhalla). Therefore there will be 43 one-on-one Viking fights before the battle is over and only one is left standing.

PROBLEM: MARY AND ELIZA

Mary and Eliza were ardent rivals in sports car racing, socializing, and the game of life in general. Mary’s husband got sick of their competitiveness and decided to teach them a lesson. He asked them to meet him at a deserted racetrack one day with their sports cars. He announced, "The winner of the race will get a brand new sports car. But this is a race with a difference. The one whose car crosses the finish line last wins. Marry and Eliza hopped into the cars and roared off around the track as fast as they could go. Why?

CUE: THE ESTATE

On his deathbed, a rancher came up with a scheme to pass on his estate. He called in his two sons, and told them to each take their horse and race to the nearest town and back. The horse that came second would win the estate for its owner. The young cowboys set out but were obviously trying to go as slow as possible in order in finish second. They were stopped by a knowledgeable villager and asked what they were doing. The sons explained the situation to the villager. The villager thought for a minute then came up with an idea that had the sons hurtling off at breakneck speed. What did the villager propose?

Answer: "Take the other’s horse"

Appendix B. Calculation methods

Calculation methods for resolution scores

Resolution scores by subject were calculated by applying the following equation:

$$\frac{X_i - Y_i}{4 - Y_i}$$

In the equation, X refers to number of problems solved ultimately, Y refers to number of problems solved before impasse (i.e., before leaving problem the first time), and i refers to i^{th} subject. This was done for both the analogous cues, and the distracter cues.

Calculation methods for immediate access and spontaneous immediate access rates, before and after seeing all problems

Access rates were calculated by applying the following equations:

$$\frac{\sum_{j=1}^{p} X_{ij}}{p_i}$$

In the equation, X refers to whether subject did (1) or did not (0) immediately turn to the relevant page in the booklet following cue; p refers to number of analogous-cued problems that were unsolved prior to cuing; i refers to i^{th} subject, and j to j^{th} cued problem. The problem sample size of p_i varied depending on condition: before/after seeing all problems in the booklet; and before/after catching on.

Calculation methods for the baseline for immediate access and spontaneous immediate access rates, before and after seeing all problems.

Baselines for access rates were calculated by examining typical problem solving behavior following the distracter cues. Following distracter cues and before seeing all problems, subjects would simply turn to the next unseen problem 96% of the time. After seeing all problems, subjects would return to previously unanswered problems 68% of the time, and to answered problems 32% of the time. Baselines were thus calculated by applying the equations:

<u>Baseline before seeing all problems</u>

$$\frac{\sum_{j=1}^{p}\left(\frac{0.04}{k_{ij}}\right)}{p_i}$$

<u>Baseline after seeing all problems</u>

$$\frac{\sum_{j=1}^{p}\left(\frac{m_{ij}}{.68u_{ij}+.32(7-u_{ij})}\right)}{p_i}$$

In the equations, k refers to number of seen problems at cue time; u to number of unanswered problems at cue time (excluding the problem the subject is currently working on), the value of m is .32 if the cued problem has already been answered whereas it is .68 if it has not, p refers to number of distracter-cued problems that were unsolved prior to cuing, i refers to i^{th} subject, and j to j^{th} cued problem. The problem sample size of p_i varied depending on condition: before/after seeing all problems in the booklet; and before/after catching on).

Article 3

A Methodology for Studying Design Cognition in the Real-World

Christensen, B. T. (2005), in proceedings from the First Nordic Design Research Conference,

Copenhagen, Denmark, 29-31 May 2005

Abstract

The in vivo research methodology holds promise to improve some of the limitations of typical design cognition methodologies. Whereas typical design cognition methodology use protocol analysis (utilizing special 'think-aloud' instructions and/or artificial settings) or retrospective analyses, in vivo research attempts to study design thinking and reasoning 'live' or 'online' as it takes place in the real world. No special instructions are used since the method relies on natural dialogue taking place between designers. By recording verbalizations at product development meetings (or other suitable objects of study), transcribing, and coding the data, it is possible to test hypotheses about design cognition in the real-world. This promises to improve the ecological validity over typical design cognition studies. Problems with the methodology include labor-intensiveness leading to small samples (possible sampling errors). To deal with this problem, it is recommended to supplement in vivo research with traditional larger sample laboratory studies.

Introduction: Design Research Method Limitations

Design activity includes cognitive processes such as problem solving and creativity (Cross, Christiaans, & Dorst, 1996), making the design domain an obvious choice (along with science and art) when cognitive scientists want to explore higher cognitive functions. Although it has been proposed that design problem solving may differ in some respects from other kinds of problem solving, the distinctions are not always sharp enough to warrant a domain-independent theory of design problem solving (Zimring & Craig, 2001). This led Zimring & Craig (2001) to argue for a design research á la carte, treating design problem solving as problem solving in general, and to focus on the reasoning processes involved (such as analogy, mental simulation, argumentation, decision making, synthesis) as these processes help construct novel and useful artifacts. Such research could potentially contribute to theoretical development in cognitive science, and facilitate the development of problem solving and creativity research that can cross narrow design disciplinary boundaries.

Studies attempting to examine design cognition usually employ methodologies such as protocol analysis, questionnaires, and interviews. Retrospective or anecdotal evidence (such as historical analyses, interviews, dairy studies or questionnaires about design processes) from designers has been used to try to pinpoint the cognitive mechanisms behind design cognition.

These retrospective methods are however very unreliable when dealing with cognitive mechanisms where the subject cannot be expected to have accurate memory of – or perhaps even conscious access to - what exactly is going on in the process (Perkins, 1981). For example, research on cognition in science has shown that conscious reconstruction of the steps that led to a discovery did not include significant elements and mechanisms that were recorded by a present observer (Dunbar, 1997). Subjects' poor memory of the steps and mechanisms involved in creative processes, as well as their inability to accurately reconstruct the events, should be taken into account in the methodology used to study such phenomena. Further, retrospective studies often provide a highly filtered view of the subject's cognitive processes, making them problematic in studies of the processes and mechanisms in design cognition. Therefore, it has been recommended that design cognition be studied using 'live' or 'on-line' research methods.

In fact, design cognition has used a particular 'on-line' methodology (protocol analysis) in the past 30 years (Craig, 2001), where subjects are instructed to 'think-aloud' while solving design problems. The use of protocol analysis seems to have increased in recent years. Ericsson & Simon (1980; 1999) developed think-aloud protocols, and argued that they did not significantly interfere with, and could accurately report the content of short term memory, and thereby reveal the processes going on in regular problem solving.

Eastman (1970) studying architecture was the first to conduct a protocol analysis in design, and since that time protocol analysis has been used to study for example, goal analysis, co-evolution of problem and solution, fixation and attachment to concepts, the role of sketching, opportunism, and modal shifts (see Cross, 2001 for a review). In 1994 the second Delft Workshop was entitled 'Research in Design Thinking II – Analysing Design Activity' (Cross et al., 1996) and focused precisely on protocol analysis. Here a number of design researchers were asked to analyze the same verbal protocols derived from experimental studies of designers. The outcome of this workshop appears to have given protocol analysis a boost in the design literature.

But even the 'online' methodology of protocol analysis is problematic as a methodology to study design cognition. A major part of protocol analysis studies focus on single subjects verbalizing concurrently while performing a given task. In this type of study, the subjects are given special 'think-aloud' instructions to verbalize all that is currently going through their head while performing the task. These instructions force the subjects to verbalize, and if they grow quite for short periods of time, the experimenter will remind them to 'please, think aloud' or 'keep talking'. Recent research has shown that forcing subjects to verbalize during problem solving can interfere with performance or change cognitive behavior (Davies, 1995; Lloyd, Lawson, & Scott, 1995). Schooler et al. showed that not only may forced 'think-aloud' protocols be inaccurate in reporting what is going on in creative cognition by interfering with non-verbal modalities (Schooler & Melcher, 1995), but they are also detrimental to the very creative process they seek to study

(Schooler, Ohlsson, & Brooks, 1993). In a number of experiments, Schooler et al. (1993) showed that think-aloud protocols apparently interfered with subjects' abilities to solve insight problems. The results could not be explained merely with respect to the conscious effort necessary to perform verbal 'online' self-reports of cognitive processes. Somehow forced think-aloud protocols interfered with ('overshadowed') the creative processes going on. Thus it seems that forced verbalizations are problematic in the study of at least some types of cognition.

Further, the typical protocol analysis study employs an experimental laboratory setup using relatively simple and artificially constructed design tasks (Cross, 2001) with a very short time span (typically less than 2 hours) using subjects (sometimes non-experts) working on their own (Ball & Ormerod, 2000b). This obviously contrasts with real-world design where the typical design task is highly complex and may span months or years. In real world design the contextual setting is typically social and team-based, but most protocol analysis studies use individuals working on their own, and even protocol analysis studies using team-based interactions often utilize teams of strangers, depriving the designers of their persistent social network and normal interaction partners. In the real world, the individual expert designers work in a personally tuned environment (e.g., their own office) with personalized tools etc (Craig, 2001), unlike the laboratory where they are asked to function without such tools. Since experts rely on external aids such as drawings and notes (Norman, 1988), it becomes increasingly important to incorporate such aids and tools in the study of design cognition, rather than focusing on verbalizations alone (Chi, 1997). Further, in experimental settings the experimenter is frequently used as 'the client', but interaction between designer and 'client' is restricted to scripted and prefabricated responses to anticipated design questions thus prohibiting more natural conversations and a meaningful image of the client (Craig, 2001). These experimental settings employed in the typical protocol analysis study have been found to have a heavy influence on the protocol data (Cross et al., 1996). In contrast, several theorists have argued that understanding situated behavior is essential for framing research on cognition (Hutchins, 1995; Suchman, 1987; Lave & Wenger, 1991), and it is somewhat paradoxical that given the highly contextualized nature of design activity, research on design expertise have typically ignored the role of situational and social factors to conduct laboratory style investigations where such factors are controlled for. This led Ball & Ormerod (2000b; 2000a) to call for an applied or cognitive ethnography in the study of design cognition. Thus, protocol analysis studies of design seem to cry out for more ecologically valid research about how the design process takes place in the real-world.

Taking this criticism of protocol analysis into account, it is necessary to study the creative process 'online' in other ways than through forced 'think-aloud' protocols conducted in the laboratory. One such 'online' methodology would be to study the creative process, as it occurs 'live' in the real-world. Dunbar (e.g., 1995; 1997; 2000; 2001b; Dunbar & Blanchette, 2001) has recently created a methodology for studying cognition in science, called the in vivo-in vitro method. The name is borrowed from the biologist's vocabulary on biology research. For example, a virus can be examined both in the Petri dish ('in vitro') and when it infects a host organism ('in vivo'). Similarly, Dunbar proposes, the same cognitive processes can be examined both in the laboratory, using controlled experiments, and as they occur 'live' in the real-world. This allows the cognitive researcher to investigate a phenomenon in a naturalistic fashion, and then go back into the psychological laboratory and conduct controlled experiments on what has been identified in the naturalistic settings (Dunbar, 2001b). This way, the methodology attempts to maintain both the ecological validity highlighted as essential by a number of researchers (e.g., Neisser, 1976; Hutchins, 1995; Christensen, 2002), as well as the experimental rigor that is possible in the psychological laboratory. In vivo research makes use of so-called messy data (Chi, 1997), which refers to such things as verbalizations, observations, videotapes and gestures studied in naturalistic settings. The in vivo – in vitro approach has been used with success in studying expertise in scientific domains such as physics, fMRi research, and astronomy (Trafton, Trickett, & Mintz, 2005; Trickett, 2004; Trickett, Fu, Schunn, & Trafton, 2000a; Trickett & Trafton, 2002; Trickett, Trafton, & Schunn, 2000b), as well as other domains of expertise, such as meteorology and submarine operations (Trickett, Trafton, Saner, & Schunn, 2005). Until now the methodology has not been applied in the study of design. But recently the present author has used this methodology to study engineering design cognition, and below I will focus on the in vivo part of this methodology, and show how it can be used to study design cognition while avoiding some of the limitations and pitfalls of the usual design cognition methodologies.

In Vivo Research on Design

The present version of in vivo research was constructed to study design cognition – notably thinking and reasoning - as it takes place in naturalistic design situations amongst expert practicing engineering designers. I identified a major international company working in medical plastics who had shown consistent design skill and creativity over a number of years. The product development department had won multiple design awards. The company agreed to take part in this study, and I was given access to the company and all aspects relating to a particular design project that was about to start up (spanning more than 2 years), including interviews with members of the project, access to the product database, email correspondence, access to meetings at all levels, including brainstorming meetings, observations of end-user product evaluation sessions, decision making meetings at both the micro and macro level, and more. I followed the design project for the first 8 months (primarily the concept design phase) of the design project. Initially a number of interviews were conducted to familiarize myself with the company and the way the project I would be following was organized. The goal was to identify points in time where creative design thinking occurs and capture this on audio or video tapes that could then be analyzed for the processes involved in the thinking and reasoning in design cognition. The time points in question would preferably be recurrent on a regular basis (e.g., occurring at regular times every week) and contain a cross-section of design activities, so as to allow for the study of multiple different design activities, allow for analysis of development over time (i.e., development across different time points as the design process progressed), and allow for the practical issue that I could schedule attendance to these time points in advance, rather than having to be present at the company at all times, as would frequently be the case in ethnographic studies. Further, the time point should be set in a group setting to ensure that natural dialogue would take place. Dunbar (1995; 1997) had discovered that in the domain of molecular biology, a suitable time point was the regularly scheduled laboratory meetings held by many scientists, especially in the natural sciences. Lab meetings consist of a senior scientist along with his or her Post Doc.s and PhD students, and Dunbar found that lab meetings contain a range of cognitive activities, such as hypothesis generation, proposal of new experiments and criticism of existing ones, and sometimes the development of entirely new concepts. He found that these meetings "... provided a far more veridical and complete record of the evolution of ideas than other sources of information" (Dunbar, 2001b). This made the lab meetings well suited as an object of study where science could be studied in a naturalistic context.

An analogous object of study in engineering design turned out to be product development meetings. The design project I would be studying incorporated 19 people who were loose or permanent members of the project. This large group was also organized into smaller units focusing on different aspects of the overall design. For example, one such sub-group focused on producing completely novel features of the product, and consisted of 5 core members (representing multidisciplinary functions, e.g., engineering, architecture, production). This subgroup (like all the subgroups involved in the project) held weekly product development meetings. Because the designers were talking out loud there was an external record of thinking and reasoning. Thus by recording product development meetings it is possible to gain access to ‘online’ thinking and reasoning without influencing the way the designers think. Using this method it is possible to directly monitor thinking and reasoning rather than uncovering reasoning through post-hoc interviews, questionnaires or think aloud protocols (Dunbar, 1997). Pilot studies in these subgroup product development meetings showed that the design activity taking place in these groups consisted of a broad cross-section of what characterizes design thinking and reasoning in general. The primary function of these subgroup product development meetings were creative development of design artifacts – that is actual creating and problem solving in collaboration – and the activity included brainstorming, concept development, design problem solving, planning of data collection and the next steps of design process, testing and evaluating mock-ups and prototypes, sketching activity, experiments, discussions and knowledge exchange about end-users, production methods and more.

A concern when conducting in vivo research is that because such research takes place in a naturalistic environment, it is likely that large amounts of irrelevant data will be captured. A risk facing the in vivo researcher is that of drowning in irrelevant data. Unlike artificial experimental settings, where the experimenter actively sets up a very particular task and context to study a particular phenomenon, in vivo research has to try to locate a suitable object of study in the real world. This approach is likely to capture irrelevant data that has to be weeded out during an often quite extensive data collection and data analysis work load. To reduce this concern of capturing too much irrelevant data, it is important to pick the object of study carefully, so that the captured irrelevant data can be kept to a minimum, and drowning in data can be avoided. To examine whether the subgroup product development meetings primarily concerned design cognition, I coded for content in these subgroup product development meetings to find out how much of the time was allocated to design thinking and reasoning. The average results indicated that, in the meetings I observed, 6% of the time concerned off-task verbalizations (such as office gossip, jokes, banter between the designers), 3% were spent summarizing the findings of past meetings (usually at the beginning of the meeting), 3% were spent planning future meetings (typically at the end of the meeting), 10% concerned planning future data collection or experiments, and a full 78% of the meetings concerned design thinking and reasoning in the here-and-now. Thus, the majority of the time spent on these meetings appeared to focus on design thinking and reasoning. Note that these exact percentages would probably be somewhat different in a different organizational context, different design project, or different phase of the design process than I studied. These percentages are merely to illustrate that this particular object of study is promising in the study of design thinking and reasoning, in that it captures relatively little irrelevant data, and looking for similar objects of study in other design projects holds promise. Further, it is meant to illustrate that it may be beneficial to conduct tests of how much irrelevant data one is likely to capture given a particular object of study.

Besides these subgroup product development meetings, other types of meetings were also held, carrying different functions. For example, leaders of each subgroup would meet with the head of the design project on a bimonthly basis to discuss strategy and status of the project. But pilot studies of these strategy meetings revealed that the content of these meetings to a much lesser extent focused on thinking and reasoning about creating design artifacts. The subgroup product development meetings were thus selected as a highly suitable object of study. It's content including a broad cross section of design activities in general, design thinking and reasoning occupied the vast majority of the time at these meetings, the regularity of the meetings, and the activity was team based and included a suitable number of people (typically 4-6) to allow for meaningful interaction. These types of product development meetings appear to be somewhat typical in engineering design teams, and there is no reason to assume that this highly suitable object of study is special for the organization or design project I was studying.

Data Collection

In vivo research requires a great deal of background knowledge of the domain in question, since the data involves experts thinking and reasoning about their usual tasks. Therefore it is necessary for the researcher to develop knowledge of the basic vocabulary and structure of the task, in order to understand what is going on. Therefore I conducted interviews with members of the subgroup as well as the project leader to familiarize myself both with the type of design product in question, the organization of the project and subgroup, the nature and steps of the design process about to begin and so on. Further, I read information about existing products in the same domain, sat in on strategy and decision making meetings, conducted pilot studies and in other ways familiarized myself with the domain and typical design process, and the vocabulary and habits of the designers that I could expect to encounter. Following this initial data gathering and familiarization, I started to collect data on the object of study (the subgroup product development meetings).

Prior to each meeting, I conducted a semi-structured interview with one of the designers to find out what the status of the project was, what was going to be the topics of today's meeting, and what they were currently working on, along with any design problems they were having. I then attended the meeting as an observer only. The meeting was videotaped, and the conversation between the designers was audio taped. When recording in vivo there appears to be a tradeoff between amount of data that can be collected, and the invasiveness of the data collection procedure (that can potentially influence the process if the designers become too self-conscious or stressed of being recorded). A non-invasive method is audio-taping only, which obviously lacks a lot of potentially important information about design objects present, motor activities and gestures, gaze of the designers etc, while capturing only verbalizations. A highly invasive method collecting some of this potentially important data could involve multiple cameras set to record total-room view, desk-tops, gestures of individual designers, and details of any note-taking or sketching behavior. Such an approach will probably influence the behavior of the designers, unless care is taking to hide all recording equipment as much as possible and allow for long trial periods to allow the participants to adapt to the artificial feel of the situation. I chose to collect an amount of data that would be relatively non-invasive, while still collecting most of the important variables. A single camera was set up high above and a short distance from the table where the designers sat during the meeting, but zoomed in so that all objects on the table could be discerned, and all sketching and note-taking activities could be captured, albeit not in detail. All people present were in the frame to allow for an examination of who was currently talking if this could not be discerned on the audiotape. Bodily gestures and general gaze could be discerned most but not all the time, depending on bodily posture of the individual designers (e.g., gaze could not be perfectly discerned when looking away from the camera). Facial expressions could not be discerned. An omnibus microphone linked to the videotape was placed center table to allow for recording of all verbalizations.

No special instructions (e.g., instructions to ‘think-aloud’) were given to participants at the meeting – they were simply informed that they should proceed with the meeting as they normally would. As an observer I took notes of information not readily available in the video frame, and collected any handouts. Following each meeting all mock-up and prototypes that had been present during the meeting was videotaped in close-up, sometimes with one of the designers explaining in voice-over the function of the object. Sketches were also recorded or copied when possible. This, together with the videotape, allowed for noting what design object (e.g., sketch, prototype or part of sketch) was currently being referenced in the verbalizations. The interviews and additional information gathered provided supplemental sources of information. The primary object of study was the videotaped meetings.

Data Analysis

Following data collection all verbalizations are transcribed. Once transcribed, the data can then be analyzed as a series of statements following standard verbal protocol analysis fashion (Ericsson & Simon, 1999). These statements can potentially reveal a lot about the cognitive mechanisms operating during the creative and reasoning processes, as Dunbar (Blanchette & Dunbar, 2000; 2001; Dunbar, 1997; 1999; 2001a) has shown. The transcription process is time consuming, and typically takes 7-10 hours labor per hour of video/audio.

The transcribed data can then be segmented (divided into units) according to a suitable grain size (i.e., size of each segment, such as proposition, sentence, episode). For much design thinking and reasoning research, such a grain size could be dividing the data into ‘complete thought’ segments (e.g., Hughes & Parkes, 2003). This entails separating verbal statements into segments containing verb phrases which are indicative of mental operations. Each segment will typically be either a single sentence or fraction of a sentence, yielding hundreds of segments per hour of recording. Each segment can then be given a time stamp, and additional non-verbal codes can be added to segments if necessary (e.g., gaze, gestures, referenced object, and so on can be coded from the video data). These segments are the primary unit of analysis.

In order to test hypotheses and theories of design thinking and reasoning, a coding scheme has to be developed. It is very difficult to convey the steps involved in choosing specific codes, since it depends entirely on the researcher’s theoretical orientation, the hypotheses or questions being asked, the task and domain (Chi, 1997). Developing and operationalizing a coding scheme is a task too complex to be described here in detail, but in essence this coding scheme development procedure follows standard verbal protocol analysis, and the reader is referred to Ericsson & Simon (1999) for more details. Rather, I will provide an extended example from ongoing research on engineering design illustrating different types of codes in the next section. Choosing a coding scheme should be done a priori so as to reduce the chance that post hoc theory will influence the data (Ericsson et al., 1999), but the first theory-laden choice of coding scheme may be too general for application on particular verbal data. Therefore, once having chosen a coding scheme, it needs to be decided what verbalizations in the data constitute evidence that they can be translated into a particular code. In other words, the codings should be operationalized in relation to the context and type of data at hand. For example, if one wants to study differences in analogical distance between different analogies in the verbalizations, it is one thing to have a general theoretically interesting distinction between 'local' analogies and 'distant' analogies, and quite another to know how to code for this distinction in a particular data set. In molecular biology, Dunbar (1995; 1997) operationalized this distinction by creating three categories: 'within organism', 'between organism' and 'non-biological or distant' analogies.

A few general comments of special relevance to in vivo data should be noted here. First, in vivo data is typically much less specific than data collected under artificial constraints in the experimental laboratory. This means that somewhat large amounts of irrelevant data will be present – even when care is taking in selecting relevant objects of study. This irrelevant data can be weeded out by applying preliminary codes that focuses in on the relevant parts of transcripts. For example, applying a code for off-task as opposed to on-task verbal behavior can remove irrelevant passages where the designers talk office and personal gossip, make jokes, banter, and other verbalizations not related to the task at hand. Another example is that transcripts can be divided into episodes. An episode is a chunk of segments that share a common theme (e.g., they all concern planning the next meeting, or they all deal with evaluating a particular prototype). By dividing transcripts into episodes, certain types of episodes can be excluded from further coding, in so far as they are irrelevant to the hypotheses being tested. But obviously care should be taken in selecting episodes for exclusion from data analysis, since this could potentially raise doubts about whether the chosen subset of data is a valid representation of the remainder of the transcripts. The nature of in vivo data further requires that the researcher pay particular attention to reliability analyses. Reliability is important in any methodology studying design cognition of course, but may be particularly important in in vivo data because of the somewhat high degree of contextual variance (as opposed to the relative contextual stability in experimental settings). Inter-rater reliability checks of individual codes using independent coders should be conducted using Cohen’s Kappa measures rather than the mere percent agreement that some researchers have reported. Percent agreement will make agreement seem much higher than warranted especially when locating phenomena that are relatively rare (‘needle-in-a-haystack’) in a large data set. Since this is often the case in in vivo data, even an exceedingly high percent agreement can be problematic. A satisfactory level of inter-rater reliability using Cohen’s Kappa should be above .70. Other types of reliabilities are also important; for example, when possible it is a good idea to recode the same hypotheses using a different coding scheme and grain size (assumed to tap into the same hypotheses), to see if the in vivo results holds up (e.g., Chi, 1997). As can be gathered by the above description, the data analysis and coding part of in vivo research is extremely labor intensive.

Example: Mental Simulation and Uncertainty in Real-world Design

The move from hypotheses to coding scheme is difficult to describe in general terms and the issue is too complex to deal with in this conference proceeding. Further, this part of in vivo research is not much different from standard verbal protocol studies, and so the reader is referred to Ericsson & Simon (Ericsson et al., 1999) for more details. So instead of describing the process in general terms, I will offer a concrete example of a coding scheme using a few different types of codes from my own data of engineering designers. These codes are from work in progress and the codes have been selected for illustrative purposes, meaning that the background of the hypotheses and the results are not explained in detail. The main focus here is on the move from hypothesis to coding scheme, and on providing examples of codes.

The hypothesis to be tested, deals with the relation between information uncertainty and mental models. A mental model is a representation of some domain or situation that supports understanding, reasoning, and prediction (Gentner, 2002). Mental models rely on qualitative relationships, such as signs and ordinal relationships, and relative positions (e.g., Forbus & Gentner, 1997). Relevant to engineering design, mental models have been used to explain human reasoning about physical systems, including devices and mechanisms (Schwartz & Black, 1996; Hegarty & Just, 1993; Hegarty, 1992). An important feature of mental models is that they frequently permit mental simulation. A mental simulation refers to the sense of being able to dynamically 'run' a simulation internally to observe functioning and outcome of a system or device. 'Runnability' implies a sense of being able to simulate system behavior and predict outcomes even for situations where the subject has no previous experience. This has been termed 'mental simulation', 'mental model runs' (Gentner, 2002) and 'conceptual simulation' (Trickett, 2004) – and here these terms are used synonymously.

Mental model runs have some disadvantages as a thinking strategy, notably inaccuracy and imprecision (Gentner, 2002). However, the potential advantage of using mental model runs in design include being able to reason about how physical systems will operate under changed circumstances/with altered features, without having to resort to actually physically constructing such a system or device. This implies quick and cheap ways of testing possible alternatives. This is particularly useful in creative domains, such as science, art and design, where uncertainty is an inescapable part of the problem space since the task involves constructing novelty. Constructing novelty implies moving into the beforehand unknown possibilities and impossibilities of the subject matter (Christensen, 2002). There are multiple ways of attempting to deal with the inherent uncertainty in design, including experimentation and other data collection, analogical thinking, and the actual construction of objects – but mental model runs may be yet another way. Mental model runs may help in the reasoning and thinking about such possibilities and impossibilities, thus reducing some of the uncertainty associated with design. Some support for this had been found in the domain of science, where use of mental models has been linked to information uncertainty and ambiguity. Trickett (2004; Trickett et al., 2002) found that the majority of mental simulations in scientific data analysis were used to evaluate hypotheses (i.e., an areas of scientific thinking fraught with uncertainty), and argued that mental model runs were used as a strategy to help resolve uncertainty. Mental simulations were used as frequently as or more frequently than any other strategy, and thus played a significant role in scientists' consideration and evaluation of hypotheses.

The present analysis was an attempt to extend this hypothesis into the domain of engineering design to see if it would hold up under different circumstances. The hypothesis being tested was thus whether information uncertainty leads to mental model runs as an attempt to reduce this uncertainty. The constructs to be measured are thus 'information uncertainty' and 'mental simulations'.

The engineering design transcripts used as data were 9 hours of video taken from the data collection described above as 'subgroup product development meetings' in the product development department of a major company in medical plastics. All 9 hours of data were from the same subgroup. These 9 hours of video were transcribed and segmented according to complete thought. The segmentation produced a total of 7414 segments covering 7 different transcripts.

Added to the transcripts were information about design objects present at the meeting to ease the coding of which design objects were currently being referenced in the protocols.

A coding scheme was developed to first limit the data set to product development in the here-and-now (i.e., reduce the transcripts to include only relevant segments), and second, to code for information uncertainty and mental model runs (we will primarily focus on the second step). The first step of the coding included coding for off-task verbalizations, segments dealing with planning future meetings or data collection, and segments dealing with referencing past meetings. The percentages of the transcript of each of these codes were reported above. This left 78% or 5806 segments of on-task here-and-now design thinking and reasoning. The second step involved coding for information uncertainty and mental simulation.

Mental Simulation

The code for mental simulations were adapted from Trickett's (2004; Trickett et al., 2002) coding scheme of scientists running mental models during data analysis. A mental model run is a mentally constructed model of a situation, phenomenon or object that can be grounded in memory or in a mental modification of the design objects currently present. This allows the designers to think and reason about new possible states of the design object and its perceptual qualities, features and functionality without actually having to physically change the object. But mental simulations do not just concern the technical aspects of the design object, but can include a host of other types of simulations of changed circumstance. One frequently occurring type concerned simulating contextual shifts, such as end-user behavior and preferences under changed circumstances (e.g., using a novel design object). The key feature in a mental simulation is that it involves a simulation 'run' that alters the representation, to produce a change of state (Trickett, 2004). This means that the simulation is not merely a question asked (e.g., changing features or functions of the design object);

it also provides a kind of answer (e.g., will it work, how should it be produced). Mental simulations thus represent a specific sequence starting with creating an initial representation, running the representation (it is modified by spatial transformation where elements or functions are for example extended, added or deleted), followed lastly by a changed representation. These three elements (initial representation, run and changed representation) are not mutually exclusive and can occur in the same utterance/segment, although frequently they will cover several segments. Each segment was coded as ‘mental simulation’ (1) or ‘no mental simulation’ (0). See table 1 for an example of a mental simulation.

Initial representation	Could you add something so that you couldn’t close this thing because there would be something in the way when you try to fold this way…
Run	But if this thing goes this way, then it is in a position to allow the ear to enter... But then I just don’t know how it should be folded… ’cause if it is folded this way then it will come out here…then it should be folded unevenly some how…You should fold it oblique.
Changed representation	It wouldn’t make any difference one way or the other. It would fold the same way, and come out on this side the same way.

Table 1: Example of a mental simulation

The mental simulation code is a qualitative code making it quite time-consuming since there is no quick and dirty way of identifying mental simulations in a transcript. The coders must code each segment in turn, noting elements of mental simulations as they go along. Further, the code requires that the coder understands much of the context for each segment, meaning that it is necessary to know about design and about what is being developed in this particular design transcript. But past research has yielded high inter-rater reliability for this code.

Information Uncertainty

To illustrate different kinds of codes, two different measures of information uncertainty will be used, one relying on syntax, the other on a combination of verbal and visual information taken from the video.

Information uncertainty using syntax. One way to code for uncertainty is to use a purely syntactical approach. This approach was adapted from Trickett et al. (2005) who used hedge words to locate segments displaying uncertainty. These hedge words included for example words like ‘probably’, ‘sort of’, ‘guess’, ‘maybe’, ‘possibly’, ‘don’t know’, ‘[don’t] think’, ‘[not] certain’, ‘believe’ and so on. Segments containing these hedge words were located and coded as ‘uncertainty present’ (1) if a scrutiny of the individual segment confirmed that the hedge word concerned uncertainty. All other on-task here-and-now segments were coded as ‘no uncertainty’ (0). Syntactical codes are quite easy to apply, but they can only be applied to a limited number of categories. See table 2 for examples of segments coded as ‘uncertainty present’ and ‘no uncertainty’ respectively.

Utterance	**Code**
’Cause I’m **not sure** whether you would fold it around the back.	Uncertainty
I **think** so too, but before we get too cocky, let’s make a model…	Uncertainty
Well, I **guess** it’s a combination of moist and heat isn’t it? I **suppose** it has to be.	Uncertainty
It has to push from the start	No uncertainty
Yes, but the problem is that you can’t hit it later …‘cause its too small	No uncertainty
It...then we have...then we loose the possibility of folding it back.	No uncertainty

Table 2: Examples of information uncertainty using syntax

Information uncertainty using verbalizations and video in combination. A different way of approaching the measure of information uncertainty is to look specifically at the objects of design thinking, or 'pre-inventive structures' (Finke, Ward, & Smith, 1992). These objects can take many different forms, including prototypes, sketches, mock-ups, or simply be ideas that are unsupported by external representations (neither in 3d physical form or on paper). It could be argued that these different kinds of design objects have different degrees of information uncertainty, in that they represent different levels of specification of the concept in question. In that line of thinking, an idea left unsupported by sketches or prototypes is more 'uncertain' than the prototype where technical features and functions are much more specified. Ideas, sketches and prototypes are all 'ambiguous' in a general sense in that they can be reinterpreted and changed somewhat rapidly, and in the sense that they represent an object-in-the-making, rather than a finished form. But the ambiguity and uncertainty may be somewhat less for prototypes than for ideas, with sketches somewhere in between. Sketches primarily support visual representation but is less specified in other modalities (haptic, gustatory, olfactory, and auditory). Therefore in design, we would expect that experts working with external support systems of sketching and prototypes would be facing less artifact uncertainty, than when no such external support exists ('idea only'). Further, sketching would provide more uncertainty than prototypes. Thus, another way of measuring information uncertainty is to code for the kind of design object being referenced. In the present transcript three different kinds of design objects occurred frequently: Prototypes, sketches and 'ideas' (i.e., objects of design thinking that were unsupported by external representation). This distinction is referred to as 'type of preinventive structure' below. Included in the transcripts were information about the design objects present at each meeting (sketches, prototypes etc.). For each segment it was first coded whether the focus of attention of the person speaking was one of these design objects present in the room. This was coded using the video recording of the design session (not the verbal data). Focus of attention was operationalized as either actual handling or holding a particular object; pointing to a particular object; or gazing toward a particular object (if this was possible to discern from the video). In effect the ‘focus of attention’ code acted as a helping variable in coding type of preinventive structure. Then coders coded whether each segment of the verbal data referred to an ‘idea’ (1), sketch (2), prototype (3) or other (4-removed from analysis), aided by the ‘focus of attention’ variable. Note it is of course perfectly possible to look at or handle one type of object and think or talk about another. In all cases the verbalized objects had precedent, meaning that if there was a difference between referenced object between focus of attention and verbalization, the object from the verbalization was chosen. Coding the preinventive structure variable was quite time consuming given that both video analysis and then verbal protocol analysis were required, but ‘focus of attention’ from the video data greatly aided the coding of the verbal protocols since most of the segments had a synchronicity between focus of attention and verbal reference.

Reliability

Following coding various forms of reliability were conducted. Inter-rater reliability was done on 17% of the data (two full transcripts), with all disagreements resolved by discussion. All inter-rater reliability tests reached satisfactory Kappa values. The syntactical uncertainty measure and the mental simulation codes both had very high Kappa values (>.90). Further, two split-half reliability analyses were conducted to test for ordering effects. Each transcript was split in half, and all analyses were re-done using the first halves and second halves separately. The transcripts were then rank ordered in terms of data collection date, and the first half of the transcripts were separated from the last half of the transcripts, and analyses were re-done on each of these halves. All split-half reliability tests yielded comparable results.

Results

The results revealed that mental simulations were extremely common in engineering design. Chi-square analyses revealed that segments containing syntactical information uncertainty had significantly more mental simulations than segments without uncertainty, supporting the hypothesis that information uncertainty and mental model runs are linked. Another chi-square showed significant differences between ideas, sketches and prototypes. Subsequent 2x2 chi-squares revealed that idea and sketches did not differ, but both had significantly more mental model runs than prototypes. This lends some support to the link between uncertainty measures as type of preinventive structure and mental model runs, although ideas and sketches did not differ. These results converge to lend support to the hypotheses that a link exists between information uncertainty and mental model runs in real-world engineering design. The link was strong enough to show up using two different codes for uncertainty under naturalistic circumstances in real-world design, thus demonstrating a strong and psychologically meaningful effect. However, since these results are correlational in nature we cannot draw any firm conclusions as to causality. We thus need more research before we can conclude that mental model runs are used *as a strategy* to reduce information uncertainty in design. The present results suffer from possible sampling biases in that only a small number of sessions and subjects were involved. More research both in vivo and in vitro should be conducted to replicate these findings.

An important advantage in using in vivo research is the fact that the data is not collected in order to test one particular hypothesis, but rather can be used to test a range of hypotheses. The nature of the data collection allows for an infinite number of re-codings of the transcribed data. These re-codings can concern finer grained analyses of the same or similar hypotheses using different codes. But the same set can also be used again in testing *other* hypotheses about thinking and reasoning in concept design, for example concerning analogical thinking, aesthethics, design planning and so on. In the domain of science this can be illustrated in the works of Susan Trickett and colleagues. They collected data on scientific data analysis in the domains of physics, astronomy and cognitive psychology, and used *the same* data sets to analyze hypotheses about conceptual simulation when evaluating hypotheses (Trickett et al., 2002), anomalies in data analysis (Trickett et al., 2000b), and change of representation in visual data analysis (Trickett et al., 2000a).

Due to the extensive data analysis and coding involved, in vivo research will typically involve only relatively few hours of recordings to be analyzed. Further, for the same reasons, usually a rather small number of different contexts are studied. This limited data variance and data amount can potentially threaten the generalizability of the results, due to increased risk of sampling error along with statistical power issues. Therefore, as mentioned, Dunbar recommends supplementing the in vivo research with in vitro controlled experiments that can better deal with these sampling and low N issues. These issues aside, in vivo research remains particularly suited to tackle the lack of ecological validity in some design cognition research.

Conclusion

The in vivo methodology holds promise to improve on some of the limitations of typical design cognition methodologies. In vivo research attempts to study design thinking and reasoning ‘live’ or ‘online’ as it takes place in the real world. In engineering design, it is argued that subgroup product development meetings may be suitable objects of study, in that pilot studies in multidisciplinary design teams reveal that the content includes a broad cross section of design activities in general, and because design thinking and reasoning occupies the majority of the time at these meetings. By recording verbalizations at such meetings (or other suitable objects of study), transcribing, segmenting and coding the data, it is possible to test hypotheses about design cognition in the real-world. In contrast to more traditional design methodologies, this approach has some advantages. In vivo methodology captures design thinking and reasoning ‘live’ as it occurs, as contrasted with some design methodologies focusing on problematic retrospective data. Further, although in vivo research shares much of the data analysis features of protocol analysis it avoids the problematic forced verbalizations typically used in verbal protocol studies. Rather, in vivo research relies on natural dialogue taking place between designers. While the typical protocol study takes place in an experimental lab setting, in vivo research focuses on real world design with expert designers working on their normal tasks, in their usual context and using personalized tools, working with their regular network and teams, over extensive periods of time. This ensures that in vivo design research will prove to have a much better ecological validity than standard experimental and protocol design research. However, in vivo design is not without problems. It can be somewhat problematic due to the labor intensive data analysis and coding issues, which may put in vivo research at risk of sampling errors and low N problems, if too few cases are subjected to analysis. To reduce this potential threat to the generalizability of the results, it is recommended that in vivo research is supplemented with standard experimental lab studies that may add experimental rigor and significantly increase the number of analyzed cases.

References

Ball, L. J. & Ormerod, T. C. (2000a). Applying ethnography in the analysis and support of expertise in engineering design. *Design Studies, 21,* 403-421.

Ball, L. J. & Ormerod, T. C. (2000b). Putting ethnography to work: The case for a cognitive ethnography of design. *International Journal of Human-Computer Studies, 53,* 147-168.

Blanchette, I. & Dunbar, K. (2000). How analogies are generated: The roles of structural and superficial similarity. *Memory and Cognition, 28,* 108-124.

Blanchette, I. & Dunbar, K. (2001). Analogy use in naturalistic settings: The influence of audience, emotion, and goals. *Memory and Cognition, 29,* 730-735.

Chi, M. T. H. (1997). Quantifying qualitative analysis as verbal data: A practical guide. *The Journal of the Learning Sciences, 6,* 271-315.

Christensen, B. T. (2002). *The creative process and reality. An analysis of search and cognition in the creative process and a call for an ecological cognitive framework for creativity.* (vols. 5, no. 3 - Psykologisk Studieskriftserie) Aarhus, Denmark: Department of Psychology, University of Aarhus.

Craig, D. L. (2001). Stalking Homo Faber: A comparison of research strategies for studying design behavior. In C.M.Eastman, W. M. McCracken, & W. C. Newstetter (Eds.), *Design knowing and learning: Cognition in design education* (pp. 13-36). Amsterdam: Elsevier.

Cross, N. (2001). Design cognition: Results from protocol and other empirical studies of design activity. In C.M.Eastman, W. M. McCracken, & W. C. Newstetter (Eds.), *Design knowing and learning: Cognition in design education.* (pp. 79-104). Amsterdam: Elsevier.

Cross, N., Christiaans, H., & Dorst, K. (1996). Introduction: The Delft Protocols Workshop. In N.Cross, H. Christiaans, & K. Dorst (Eds.), *Analysing design activity* (pp. 1-16). Chichester: John Wiley & Sons.

Davies, S. P. (1995). Effects of concurrent verbalization on design problem solving. *Design Studies, 16,* 102-116.

Dunbar, K. (1995). How scientists really reason: Scientific reasoning in real-world laboratories. In R.J.Sternberg & J. E. Davidson (Eds.), *The nature of insight.* (pp. 365-395). Cambridge, MA, US: The MIT Press.

Dunbar, K. (1997). How scientists think: On-line creativity and conceptual change in science. In T.B.Ward, S. M. Smith, & J. Vaid (Eds.), *Creative thought: An investigation of conceptual structures and processes.* (pp. 461-493). Washington, DC, US: American Psychological Association.

Dunbar, K. (1999). How scientists build models: InVivo science as a window on the scientific mind. In L.Magnani, N. Nersessian, & P. Thagard (Eds.), *Model-based reasoning in scientific discovery.* (pp. 89-98). NY, US: Plenum Press.

Dunbar, K. (2000). How scientists think in the real world: Implications for science education. *Journal of Applied Developmental Psychology, 21,* 49-58.

Dunbar, K. (2001a). The analogical paradox: Why analogy is so easy in naturalistic settings yet so difficult in the psychological laboratory. In D.Gentner, K. J. Holyoak, & B. N. Kokinov (Eds.), *The analogical mind: Perspectives from cognitive science.* (pp. 313-334). Cambridge, MA: The MIT Press.

Dunbar, K. (2001b). What scientific thinking reveals about the nature of cognition. In K.Crowley, C. D. Schunn, & T. Okada (Eds.), *Designing for science: Implications from everyday, classroom, and professional settings.* (pp. 115-140). Mahwah, N.J., US: Lawrence Erlbaum Associates,Inc.,Publishers.

Dunbar, K. & Blanchette, I. (2001). The invivo/invitro approach to cognition: the case of analogy. *Trends in Cognitive Sciences, 5,* 334-339.

Eastman, C. M. (1970). On the analysis of intuitive design processes. In G.M.Moore (Ed.), *Emerging methods in environmental design and planning*. Cambridge, MA: MIT Press.

Ericsson, K. A. & Simon, H. A. (1980). Verbal reports as data. *Psychological Review, 87,* 215-251.

Ericsson, K. A. & Simon, H. A. (1999). *Protocol analysis: Verbal reports as data*. (Third ed.) Cambridge, MA, US: The MIT Press.

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

Forbus, K. D. & Gentner, D. (1997). Qualitative mental models: Simulations or memories? In.

Gentner, D. (2002). Psychology of mental models. In N.J.Smelser & P. B. Bates (Eds.), *International encyclopedia of the social and behavioral sciences* (pp. 9683-9687). Amsterdam: Elsevier.

Hegarty, M. (1992). Mental animation: Inferring motion from static displays of mechanical systems. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 18,* 1084-1102.

Hegarty, M. & Just, M. A. (1993). Constructing mental models of machines from text and diagrams. *Journal of Memory & Language, 32,* 717-742.

Hughes, J. & Parkes, S. (2003). Trends in the use of verbal protocol analysis in software engineering research. *Behaviour & Information Technology, 22,* 127-140.

Hutchins, E. (1995). *Cognition in the wild*. Cambridge, MA: The MIT Press.

Lave, J. & Wenger, E. (1991). *Situated learning: Legitimate peripheral participation*. Cambridge, UK: Cambridge University Press.

Lloyd, P., Lawson, B., & Scott, P. (1995). Can concurrent verbalization reveal design cognition? *Design Studies, 16,* 237-259.

Neisser, U. (1976). *Cognition and reality*. San Francisco, CA, US: W. H. Freeman and Co.

Norman, D. A. (1988). *The psychology of everyday things*. New York, NY, US: Basic Books, Inc.

Perkins, D. N. (1981). *The mind's best work*. Cambridge, MA: Harvard University Press.

Schooler, J. W. & Melcher, J. (1995). The ineffability of insight. In S.M.Smith, T. B. Ward, & R. A. Finke (Eds.), *The creative cognition approach.* (pp. 97-133). Cambridge, MA, US: The MIT Press.

Schooler, J. W., Ohlsson, S., & Brooks, K. (1993). Thoughts beyond words: When language overshadows insight. *Journal of Experimental Psychology: General, 122,* 166-183.

Schwartz, D. L. & Black, J. B. (1996). Analog imagery in mental model reasoning: Depictive models. *Cognitive Psychology, 30,* 154-219.

Suchman, L. A. (1987). *Plans and situated actions: The problem of human-machine communication.* New York, NY, US: Cambridge University Press.

Trafton, J. G., Trickett, S. B., & Mintz, F. E. (2005). Overlaying images: Spatial transformations of complex visualizations. *Foundations of Science*.

Trickett, S. B. (2004). *Movies-in-the-mind: The instantiation and use of conceptual simulations in scientific reasoning.* Unpublished Doctoral dissertation from George Mason University.

Trickett, S. B., Fu, W.-T., Schunn, C. D., & Trafton, J. G. (2000a). From dipsy-doodles to streaming motions: Changes in representation in the analysis of visual scientific data. In *Proceedings of the 22nd Annual Conference of the Cognitive Science Society.* (pp. 959-964). Mahwah, NJ: Erlbaum.

Trickett, S. B. & Trafton, J. G. (2002). The instantiation and use of conceptual simulations in evaluating hypotheses: Movies-in-the-mind in scientific reasoning. In *Proceedings of the 24th Annual Conference of the Cognitive Science Society* (pp. 878-883). Mahwah, NJ: Erlbaum.

Trickett, S. B., Trafton, J. G., Saner, L., & Schunn, C. D. (2005). 'I don't know what's going on there': The use of spatial transformations to deal with and resolve uncertainty in complex visualizations. In M.C.Lovett & P. Shah (Eds.), *Thinking with data* (Mahwah, NJ: Erlbaum.

Trickett, S. B., Trafton, J. G., & Schunn, C. D. (2000b). Blobs, Dipsy-Doodles and other funky things: Framework anomalies in exploratory data analysis. In *Proceedings of the 22nd Annual Conference of the Cognitive Science Society* (pp. 965-970). Mahwah, NJ: Erlbaum.

Zimring, C. & Craig, D. L. (2001). Defining design between domains: An argument for design research á la Carte. In C.M.Eastman, W. M. McCracken, & W. C. Newstetter (Eds.), *Design knowing and learning: Cognition in design education* (pp. 125-146). Amsterdam: Elsevier.

Article 4

The Relationship of Analogical Distance to Analogical Function and Pre-inventive Structure: The Case of Engineering Design

Christensen, B. T. & Schunn, C. D. (submitted)

Abstract

Analogy was studied in real-world engineering design, using the in vivo method. Spontaneous analogizing was found to occur frequently, entailing a roughly equal amount of within and between domain analogies. In partial support for theories of unconscious plagiarism (Brown & Murphy, 1989; Marsh, Landau, & Hicks, 1996) and Ward's (1994) Path-of-least-resistance model, it was found that the reference to exemplars (in the form of prototypes) significantly reduced the number of between-domain analogies between source and target, compared to using sketches or no external representational systems. Analogy served three functions in relation to novel design concepts: identify problems, solve problems or explain concepts. Problem identifying analogies were mainly within-domain; explanatory analogies were mainly between domain; while problem solving analogies were a mixture of within and between domain analogies.

Introduction

Analogy involves accessing and transferring elements from familiar categories (often referred to as the 'source' analogue) to use it in constructing a novel idea ('target'), e.g., in an attempt to solve a problem or explain a concept (e.g., Gentner, 1998). In its most general sense, analogy is the ability to think about relational patterns (Holyoak, Gentner, & Kokinov, 2001). Most of the research on analogical reasoning has used experiments with naive novice subjects working on simple tasks in the psychological laboratory. Although this research has proven valuable and produced many interesting results, it is important to try to establish whether these findings will generalize to real-world analogizing in studies with more ecological validity. This is especially important when experimental laboratory findings are somewhat counter-intuitive from an applied perspective, such as the finding of poor rates of spontaneous access when the similarity between source and target is mainly structural (Gick & Holyoak, 1980; Gick & Holyoak, 1983; Anolli, Antonietti, Crisafulli, & Cantoia, 2001). Also, experimental effects may be so small in the real world as to be of no importance.

Analogical reasoning is assumed to be a general human capacity (Holyoak & Thagard, 1995) involved in most domains, although perhaps notably creative problem solving domains such as science, design and art. Engineering design, perhaps especially the early conceptual stage of the design process, is one such creative domain. Design theorists (e.g., Roozenburg & Eekels, 1996; Casakin & Goldschmidt, 1999; Goldschmidt, 2001) have argued for the importance of analogy in design, and several design tools or techniques make extensive use of analogy, including Synectics (e.g., Gordon, 1961), and TRIZ (e.g., Terninko, Zusman, & Zlotin, 1998). Further, anecdotes of inventors and engineers making breakthrough discoveries or inventions following distant analogical transfer abound in the creativity literature. A few of the most famous include George de Mestral developing Velcro after examining the seeds of the burdock root that had attached themselves to his dog. Gutenberg is said to have developed the printing press modeled after a wine press; Dunlop to have developed the pneumatic tire from garden hose; Westinghouse to have developed train air brakes from air-powered rock drills; and James Watt to have developed the steam engine from a tea kettle. Regardless of whether these stories are true, the sheer number of famous stories illustrates the importance engineering designers have placed on analogy.

During the early stage of design (the so-called concept design phase) the designer works on constructing a central (or core) technical concept around which the entire design will be built (Jansson & Smith, 1991; Dahl & Moreau, 2002). Much of the design activities during this phase can be characterized as problem structuring activities (Goel & Pirolli, 1992; Goel, 1995), or problem finding, frequently entailing both constructing of multiple possible concepts (e.g., by utilizing creative process management techniques such as brainstorming (Osborn, 1963)), and exploring some of these possible concepts in more detail. As such, the early stages of design frequently entail a mixture of breath first and depth first search approaches (Ball & Ormerod, 1995).

The creative cognitive processes working in concept design are frequently supported by external representational systems, such as sketching (McGown, Green, & Rodgers, 1998) and the construction of mock-ups and prototypes (Römer, Pache, Weisshahn, Lindemann, & Hacker, 2001). In a survey among engineering designers, Römer et al. (2001) found that 96% reported using sketches and models during the development of solution concepts. Notably this was also the case during the early stages of design (task clarification and concept design) where 95% reported using sketches and 58% reported using prototypes. However, it is a relatively unexplored avenue in cognitive science how creative cognitive phenomena (such as analogy, imagery, and mental models) interact with the presence or absence of external support systems. In the literature, such interaction studies have typically been limited to lab experiments contrasting sketching with no sketching in order to examine characteristics of imagery, such as the possibility of restructuring (e.g., Anderson & Helstrup, 1993; Finke, 1990; Verstijnen, van Leeuwen, Goldschmidt, Hamel, & Hennessey, 1998; Verstijnen, van Leeuwen, Hamel, & Hennessey, 2000). This literature on imagery has revealed that, at least under some conditions, cognition with the support of sketching is superior to cognition without external support (Verstijnen et al., 1998; Verstijnen et al., 2000; Roskos-Ewoldson, Intons-Peterson, & Anderson, 1993).

Returning back to analogical transfer, different forms of external support may also have different affordances for analogical transfer during design. More concrete external representations may produce more superficial kinds of transfer, whereas absent or more abstract external representations may produce more distant or more functionally relevant analogical transfer. Therefore, it is also interesting to extend the contrast between sketches and no sketches to even more concrete support systems than sketches, such as prototypes. Early prototyping is used extensively in design communities today (Römer et al., 2001), and the construction and evaluation of prototypes provide a rather different kind of support to the cognitive system than does sketching. For example, while sketches support visual perception well, it does not support other sense modalities (haptic, gustatory, olfactory, and auditory) as well as prototypes do.

There are multiple reasons why it is interesting to study whether the presence or absence of different kinds of external representations may be useful for analogizing in the conceptual stages of design. From a theoretical perspective, past research on restructuring has revealed that there are differences between sketching and imagery in their ability to support perceptual restructuring and information extraction (as noted above) and similar differences between analogical thinking with and without external representation may be revealed. Considerable past research on analogical transfer has focused on problems with a single 'correct' answer. It is less clear how analogical transfer will function in much more ill-structured domain like design. Here concrete representations could constrain analogical inferences more heavily to superficial matches or they could better support selecting from prior experiences to find some useful representation to ground the open-ended problem.

On a more practical level, knowledge of which types of external support are useful for analogical transfer in design can help designers structure the early stages of the design process in ways that has the potential of improving the creative outcome. 3D-CAD systems are being increasingly employed in design communities, sometimes at the expense of traditional sketching and prototyping. Therefore it is important to examine what the cognitive benefits of the traditional intuitive and effort-saving representational systems are to the creative engineering process, so that (if possible) these benefits can be transferred to the CAD systems, or alternatively, that the traditional approaches are retained in use. Thus, knowledge of why and how analogy is best supported by sketches and prototypes may help in making new and improved design tools (such as computer systems) for the conceptual design phase.

Spontaneous Analogical Access

In analogical transfer, spontaneous access refers to accessing a source without receiving hints or instructions to use this source. A consistent finding in the research literature is that transfer increases with similarity (e.g., Holyoak & Koh, 1987; Novick, 1988; Ross, 1987; 1989; Simon & Hayes, 1976). But whereas analogical transfer has been found to be closely related to structural similarity, analogical access often strongly depends on superficial similarity between source and target (Gentner, Rattermann, & Forbus, 1993; Holyoak et al., 1987; Ross, 1987; Novick, 1988). Despite the widespread assumption that analogy is important in design, a large number of experimental studies have found that spontaneous access is relatively rare when the source is provided by the experimenter, unless salient superficial similarity exists between source and target. Given the importance analogy has been afforded in cognitive science, it is somewhat surprising that salient superficial similarity or explicit instructions to use previous knowledge is required for analogical access to occur when relevant structural information is stored in long-term memory (Gick et al., 1980; Gick et al., 1983; Anolli et al., 2001). In design, Casakin & Goldsmith (1999) have shown how providing analogous visual displays to design problems produce better quality solutions when subjects are required to use the analogous visual displays than when no such instructions are provided. This was the case for both expert and novice designers and again demonstrates the importance of instructions in analogical access. However, some suggestive evidence that designers do produce some measure of spontaneous access comes from Ball, Ormerod and Morley (2004). In an experimental setting, they asked expert and novice designers to complete a simple engineering design task while thinking-aloud. Protocol analysis revealed that both experts and novices produced analogies, although experts produced more analogies than novices. But generally the research on spontaneous access has shown disappointing results. Anolli et al. (2001) argued these findings suggested that analogical access is not a spontaneous process, but, as Ross, Ryan & Tenpenny (1989) noted, studies have not shown that people never spontaneously access relevant information - only that in cases where they are expected to do so, they often do not.

The lack of experimental evidence for spontaneous access is surprising for another reason. Dunbar (Blanchette & Dunbar, 2000; 2001; Dunbar, 1995; 1997; 2001a) has conducted several studies of analogical transfer in real-world science and politics using the in vivo methodology (see Dunbar, 1995; 1997) and he has found that scientists and politicians frequently access analogues spontaneously (e.g., Dunbar, 2001a) – even for analogues sharing mainly deeper structural features (see also Bearman, Ball, & Ormerod, 2002 for similar results on a task in management decision making). This research finding obviously stands in sharp contrast to the experimental findings; a contrast Dunbar called the analogical paradox. For some reason experimental research and real-world research seem to reach opposite conclusions concerning frequency of spontaneous access. Experiments on spontaneous access are supposed to be studying a simplified version of the real-world, but several differences exist between the two research strains that could potentially explain the paradox. For one thing Blanchette & Dunbar (2000; 2001; Dunbar, 2001a) has highlighted that in the real-world experts generate their own analogies, while in the typical experimental laboratory subjects are provided with specific analogue sources. Similarly in design, designers may be able to spontaneously produce structural analogies because they have generated the source problems in prior design work.

Analogical Distance

In analogical transfer, the ’distance’ between source and target may be large or small. For example, a designer trying to develop door handles for the auto industry may make an analogy to other door handles in the auto industry (within-domain, or local analogies), or make an analogy to telephones or oysters in developing the design (between-domain, or distant analogies) (see Dunbar, 1995; Dunbar & Blanchette, 2001). Local analogies involve greater superficial similarity between source and target, compared to lesser amounts of superficial similarity in distant analogies. This increase in superficial similarity may make local analogies easier to access (e.g., Gentner et al., 1993; Holyoak et al., 1987). Both local and distant analogies involve structural similarity. But, because distant analogies involve two vastly different bodies of knowledge, it may be more difficult to ensure successful transfer of solution elements in design problem solving from source to target because the domains may differ in multiple subtle ways. It may thus be more difficult to successfully transfer between source and target from vastly different domains (Johnson-Laird, 1989).

The use of distant analogies may be positively related to originality in design. Although experimental evidence on the subject is scarce, Dahl (2002) found that the percentage of distant analogies used during design had a significant positive effect on the estimated originality of the resulting product in design. The anecdotes quoted above point in similar directions, since all the anecdotes of groundbreaking design involve exclusively distant analogies. Further, in an experimental study of visual analogy in design, Casakin (2004) found that both novices and experts produced more between-domain than within-domain analogies. The experimental setup involved providing subjects with visual analogous displays and instructing them to use analogies, and this choice of experimental setup may have significantly affected the results. These research findings and anecdotes would lead one to assume that distant analogies are very common in design, and play an important part in the generation of creative products.

In science, similar anecdotal evidence of distant analogizing leading to breakthrough discoveries exist (see e.g., Shepard, 1978; Ghiselin, 1954). However, in his studies of real-world analogy in microbiology, Dunbar (1995; 2001a) found that distant analogies did not play a significant part in discovery. Dunbar divided analogies into local, regional, and distant, and found that while distant analogies were only used to explain concepts to other scientists, local analogies were used to fix problems with experiments and design new experiments as well as explain. Regional analogies (within biology, but between species) were, however, used when forming hypotheses, as well as when designing new experiments and explaining concepts to other scientists. Distant analogies were very rare in comparison to local and regional analogies.

So, it is unknown whether real-world designers use mainly within-domain or between-domain analogies. In-so-far as the real-world science findings will generalize, we would expect distant analogies to be rare in real-world design. But in-so-far as Casakin's experimental findings will generalize, we would expect designers to use mainly distant analogies.

The Influence of External Support Systems on Analogical Transfer

Most theories of design stress the dual nature of the processes involved, sometimes referred to as generative vs. exploratory processes. These phases or processes are central in the Geneplore model (Finke, Ward, & Smith, 1992) which basically describe the creative process as an iterative process of generation and exploration under product constraints. Similar models illustrating the dual nature of the inventive process have been proposed by Anderson & Helstrup (1993), Roskos-Ewoldsen (1993), and is captured by Goldschmidt (1994; 1991) in the distinction between 'seeing as' and 'seeing that', and Schön's (e.g., Schon & Wiggins, 1992) notions of 'seeing moving seeing'.

The dual nature of the design process reflects the dual nature of the design objects involved in design and invention. On the one hand, these objects are in-the-making or being constructed by cognitive processes. But on the other hand, the same objects are being explored as objects in-the-world (or at least as objects to-be-in-the-world). In the generative phase, mental representations (so-called ‘pre-inventive structures’) are constructed. These pre-inventive structures have properties that promote creative discovery. In the exploratory phase, these properties are exploited in that the pre-inventive structures are interpreted in meaningful ways (Finke, Ward & Smith, 1992). The pre-inventive structures can, for example, be visual patterns, object forms or mental models, and the properties that make them promote creative discovery include novelty, ambiguity, emergence, implicit meaningfulness and divergence (the capacity for finding multiple uses in the same structure).

These pre-inventive structures are precursors for the final externalized creative products, and they are generated, explored, modified and re-interpreted during the creative process. Although the term ‘pre-inventive structure’ typically refer to unsupported cognitive structures (Finke, Ward & Smith, 1992), it may also refer to cognitive structures with external support of sketches or prototypes.

Pre-inventive structures have mainly been studied as mental representation without support from external representations. But, as noted above, a growing literature has focussed on comparing sketching with not-sketching conditions. This makes sense since the properties of pre-inventive structures highlighted by Finke, Ward & Smith (1992) are also properties of pre-inventive structures with the support of sketching (indeed, the property of ‘ambiguity’ is a prime reason why many cognitive researchers find sketching so interesting). Far less studied, however, are pre-inventive structures with the external support of 3D physical objects, such as prototypes. Unlike sketches, prototypes support not only visual modalities, but potentially all other modalities as well. All the properties highlighted as central for pre-inventive structures can indeed be properties of pre-inventive structures w/support of prototypes. Although it may seem counter-intuitive, prototypes (at least in the early stages of design) can be ambiguous and contain emergent properties (supporting reinterpretive processes such as functional inference and contextual shifting), novelty, divergence, implicit meaningfulness and so on. In concept design, prototypes may change appearance rapidly, be reinterpreted in terms of functional inference, contextual shifting, and may otherwise be explored and changed following their conception qua prototype. As such, a prototype is both a 'fluid' structure able to change meaning and appearance, and an object frozen in time and form. That prototypes are not mere end-products of the process, but rather play an intrinsic part in supporting creative processes during design can be seen in the fact that prototypes can be constructed for several different reasons, including to see if a particular concept will work at all. Thus, constructing a prototype is not a straightforward matter, and may afford surprises and even fail, e.g., when unexpected design constraints make the construction impossible. So, sketching and prototyping can be considered an integral part of the design process – as pre-inventive structures – rather than mere terms for the end-products of the design process. Prototyping and sketching thus share properties with unsupported cognitive pre-inventive structures, in that they are "…formed without full anticipation of their resulting meaning and interpretation. In addition they are to be distinguished from the final externalized creative products, which, in contrast, are often fully interpreted and extensively refined" (Finke et al., 1992, p. 22-23).

Exemplar Influence on Analogical Transfer

Since objects from similar domains share more superficial similarity than objects from dissimilar domains, and superficial similarity is one of the key driving forces of analogical access, we would expect that the presence or availability of within-domain exemplars would increase the likelihood of within-domain analogizing (Ward, 1998). In other words, the presence of within domain examples may make it hard for subjects to break away from local analogies, since superficial similarity dominates access, and distant analogies will be less superficially similar than local analogies. Providing prior within-domain examples should thus bias people's creations toward features contained in those examples (Marsh et al., 1996). This within-domain biasing could for example be the case when designers use external support of prototypes during the concept phases in engineering design, as compared to conditions without such support.

The idea that fixating elements (either in memory or the environment) may constrain performance or creativity is well-known, including phenomena such as 'functional fixedness'(restricting the uses of objects to well-known functions (Maier, 1931)), and 'mental set' (situational induced obstacles to problem solving (Luchins, 1942)). Further, it has been found that providing (Jansson et al., 1991; Dahl et al., 2002; Marsh, Ward, & Landau, 1999; Ward, 1994) or retrieving (Ward, 1994) existing examples may inhibit generative creative processes, and lead to a higher proportion of property transfers from examples into the subject's own work (e.g., Marsh et al., 1996), even when subjects are explicitly instructed to avoid such transfer (e.g., Smith, Ward, & Schumacher, 1993). Source monitoring of this property transfer is poor especially in generative tasks (e.g., Marsh, Landau, & Hicks, 1997), which originally led to the label 'unconscious plagiarism' or cryptomnesia (i.e., unconscious influence of memory causes current thoughts to be (wrongly) experienced as novel or original inventions) (Brown et al., 1989; Marsh & Bower, 1993; Marsh & Landau, 1995; Marsh et al., 1999). Ward (1994; 1995; 1998) proposed a path-of-least-resistance model to account for some of these findings, which states that the default approach in tasks of imagination (especially when few constraints must be satisfied) is to access a specific known entity or category exemplar (gravitating towards basic level), and then pattern the new entity after it.

In support of this model, Ward (1994; Ward, Patterson, Sifonis, Dodds, & Saunders, 2002) found that people who report basing their novel constructions on specific exemplars are less original than people who use other strategies. Property transfer in generative tasks has proven robust across a variety of settings, including engineering design tasks conducted in the lab (Jansson et al., 1991; Dahl et al., 2002). Jansson & Smith (1991) had subjects (either mechanical engineering students or professional designers) work on simple design problems (such as how to construct a car-mounted bicycle rack), with (the fixation group) or without (the no fixation group) a specific example provided by the experimenter. They found that the fixation group included more properties from the examples. However, it should be noted that a failure to replicate this finding has been reported (Purcell & Gero, 1992). Dahl (2002) had undergraduate engineering students design new products that would solve problems of the commuting diner (e.g., difficulties with spillage, consumption and storage of food during automotive driving). Subjects who saw an example sketch of a drive-in window tray transferred significantly more properties, and generated fewer distant analogies, than subjects who saw no sketch.

However, the effect sizes were somewhat small, and therefore it remains unclear whether these results will generalize to engineering design experts working on their own tasks in their natural work environment. It has never been tested whether exemplars that are a natural part of the design process, such as external support systems in the form of sketches or prototypes generated by the designers themselves, also constrain analogical distance. These findings from the analogical transfer and fixation literature led to the following hypothesis: Use of within-domain external support (such as prototypes and sketching) during the concept design phase will lead to a lower proportion of between-domain analogies, compared to using no such external support.

This prediction leaves open the question of whether there may be a difference in average analogical distance between external support relying on sketching vs. prototypes. Comparing sketches to prototypes, it could be argued that sketching is characterized by density and ambiguity (Goel, 1995), and concentrates (like diagrams, but unlike drafting) primarily on illustrating structural aspects of objects-to-be. As such, it could be argued that sketching involves less superficial similarity, when compared to prototypes. This argument led to the hypothesis that sketches will lead to a higher proportion of between-domain analogies relative to within-domain analogies when compared to external support with high resolution and superficial detail such as prototypes.

In comparing sketches and prototypes in a naturalistic setting, it is important to consider whether there is an inherent temporal or order confound—i.e., can any difference in performance with ideas vs. sketches vs. prototypes be attributed to <u>when</u> they are used in the design process. We will examine this potential confound empirically in the current study. But it is also useful to say a few words about the temporal structure of design processes.

Pre-inventive Structures and the Temporal Structure of the Design Process

It is tempting to equate the move from initial to goal states in a design problem space with a move from the abstract idea to the concrete physical object, especially since the initial state in design problems (by definition) lacks a concrete solution, and that the goal state entails such a concrete solution. Such a mapping would seem to suggest that the design process progresses from an abstract problem state to a concrete goal state going gradually and sequentially through intermediate states of concreteness, from ideas to diagrams and sketches, on to high resolution illustrations, on to 3-dimensional mock-ups and finally ending in the fully functional prototype.

However, such a characterization of the concept design phase would be misleading for several reasons: Even in so far as the mapping between a design problem space and degree of object concreteness holds up, this mapping should not imply a ‘incremental and sequential move’ to ever higher degrees of concreteness. Such incremental and sequential moves may be characteristics of well-defined problem spaces, but in ill-structured domains (such as design) (Reitman, 1965; Simon, 1973), the process is instead characterized by changes in representation (‘jumps in the problem space’) (Kaplan & Simon, 1990), discarded concepts and non-linear movement (Perkins, 1981), and a frequent alternation between internal mental procedures, and external materializing actions (Römer et al., 2001). Therefore, linear incrementality does not characterize the move from abstract to concrete in design problem spaces.

Further, the degree of abstractness or concreteness in intermediate states of the design process does not correspond to closeness of solution. Abstract design concepts may be closer to being a useful concept than (flawed or rejected) prototypes. Even the initial design state can hardly be considered completely abstract insofar as ideas do not come into being out of nothing (‘ex nihilo’; even though dictionary definitions may have us believe this (Boden, 1990; Perkins, 1988)). For any given designer, a typical design process starts out with a large knowledge base of functionally and physically similar concrete design objects. It therefore seems ill advised to characterize most initial design states as completely ‘abstract’.

Functions of Analogy in Engineering Design

Analogies are constructed for different purposes. In his studies of microbiology labs, Dunbar (1997; Dunbar, 2001a) distinguished 4 types of functions for analogies: forming hypotheses, designing experiments, fixing experiments, and explaining concepts to other scientists. These functions are however, at least in part, specific to science, and do not apply to design. Although engineering design certainly can involve experimentation, other kinds of activity are more prevalent and important, such as the construction, modification and evaluation of novel and useful objects. Another classification of the function of analogies comes from Ward (1998), who classified analogies in invention or design as either ‘explanatory’ or ‘inventive’. Along similar lines, Bearman, Ball and Ormerod (2002) examined analogy using a simulated real-world management decision making task, involving undergraduate students. In this context, they distinguished between two different functions of analogies: problem-solving and illustration. Both of these functions are well-suited for the ill-structured domain of engineering design.

Adapting these prior distinctions to the design setting, we distinguish among three functions of analogies: explanation, problem solving, and problem identification. Engineering design is frequently conducted in teams, rather than individually, whereby communicating novel ideas to other members of a team becomes an important part of the process. Explanations through analogy can be a way of enhancing and ensuring comprehension, while avoiding misunderstanding when dealing with novelty. Communicative alignment quickly becomes a problem when novel ideas have to be shared among group members, because team members may not be certain that they are referring to the same design object, since this object may take the form of nothing more than an idea. It could be hypothesized that when a design object takes the form of an idea (i.e., without external support of sketches or prototypes) the need for communicative alignment will be greater than when external support is present, thereby producing more explanatory analogies. Thus explanation or illustration using analogy is certainly a function to be expected in engineering design. Another function analogy is expected to serve, is that of problem solving. Indeed, this function is perhaps the primary reason researchers have focussed on analogy in design and science. In addition to these two functions, problem identification may be an important function especially in the early conceptual stages of engineering design. When developing novel concepts, it is necessary to try to foresee whether a novel idea or concept would work under particular circumstances. In this case, analogy may play some part in evaluating novel concepts, in that it is possible to transfer, not only solutions, but also potential problems from sources with which the subject has past experience. Here the elements to be transferred from source to target involve potential design problems that the new concept may display. For example, a designer may comment on a new design, that it is not going to work, since he has previously constructed a similar design, and that previous design had inherent design flaws which this new design also posses. The function of this type of analogy involves identifying potential problems in the evaluation of new designs, based on past experience.

The frequency of these three kinds of analogical functions in real-world design (explanation, problem-solving, and problem identification) remains to be seen. Based on Dunbar (1997; 2001a) and Bearman et al. (2002) results, we should at least predict that a substantial proportion of the analogies would involve explanations. In science, Dunbar found that almost half of the analogies were explanatory, and in management decision making, Bearman et al. found that 27% of the analogies served the function of illustration.

What can be predicted is that analogical distance will interact in particular ways with analogical function. A significant part of explanatory analogies would be expected to be between-domain analogies, since explanations of novel design objects may entail relating the novel object to well-known exemplars outside the present domain in order to successfully communicate ideas to others. In science, Dunbar has argued that between-domain analogies are primarily explanatory in function. Between-domain analogizing may be necessary in explaining novel design concepts exactly because the concept is new to the domain. Further, problem identification in evaluation may involve primarily within-domain analogies, since within-domain analogies may be more accessible (due to superficial similarity), available (due to within domain expertise), as well as be the analogy-of-choice in identifying problems since within domain analogies may increase the chances of successful transfer. Finally, because engineering design involves the production of novel and useful solutions, solving problems by relating to past within-domain knowledge may frequently not be enough to construct an original product. Therefore, a mixture of within and between domain analogies are expected when the function of the analogy is to solve a design problem.

Methods

Data was collected using Dunbar's (1995; 1997; Dunbar et al., 2001) in vivo methodology. In vivo methodology allows the researcher to study expert thinking and reasoning 'live' or 'online' in the real-world. To this end, a major international company focusing on engineering design in the domain of medical plastics was selected for its persistent creativity displayed over many years. The company's R&D department had won multiple design awards for a number of different designs. Upon contacting the R&D department, a particular design project about to start up was chosen as the focus. The design project involved a total of 19 expert engineering designers organized in 3 subgroups focusing on different aspects of the design object, and involved designing a new and improved product in a domain where the company already had multiple products and extensive experience. The 3 subgroups were organized as multidisciplinary teams, involving different functions. The design project would span over 2 years.

When conducting in vivo research, it is necessary to locate a suitable object of study (or time point), where, in this case, design thinking and reasoning can be studied. A suitable object of study should include a broad cross-section of design activities, include mainly design activities in the here-and-now (e.g., rather than retrospective accounts of designing), and involve natural dialogue between designers so as to avoid potentially problematic 'think-aloud' instructions (see also Christensen, 2005). For example, Kevin Dunbar studied scientific thinking and reasoning by recording lab group meetings because he found that they "... provided a far more veridical and complete record of the evolution of ideas than other sources of information" (Dunbar, 2001b). An analogous object of study in design turned out to be subgroup product development meetings. Each subgroup in the project held product development meetings on a regular basis (e.g., a weekly basis). Because the designers were talking out loud there was an external record of thinking and reasoning. Thus by recording product development meetings it is possible to gain access to 'online' thinking and reasoning without influencing the way the designers think. Pilot studies in these subgroup product development meetings showed that the design activity taking place in these groups consisted of a broad cross-section of what characterizes design thinking and reasoning in general (Christensen, 2005). The primary function of these subgroup product development meetings were creative development of design artifacts – that is actual creating and problem solving in collaboration – and the activity included brainstorming, concept development, design problem solving, planning of data collection and the next steps of design process, testing and evaluating mock-ups and prototypes, sketching activity, conducting experiments, discussions and knowledge exchange about end-users, production methods and more. Further, it was found that the vast majority of the design activity at these meetings concerned design thinking and reasoning in the here-and-now. Therefore, the subgroup product development meetings were chosen as the most suitable object of study.

The current paper examines the meetings of one particular subgroup. This subgroup had the task of developing completely novel features for the new product. It consisted of 5 core members (1 female and 4 male) representing different functions (industrial designer, lab technician, project manager) and backgrounds (machine engineering, architecture, machinist). They all had extensive experience in medical plastics and design (10, 10, 20, 27, 35 years). Besides these core members, the team would invite experts with specialized knowledge from other parts of the company to participate in the group meetings when required. This subgroup was referred to as the 'New Feature' group.

Background information was collected including interviews with the designers and the head of the design project, observations of project strategy meetings, observing sessions where end-users described problems they were having with existing products, and reading written materials on the project and product domain etc. After collecting this background information, the subgroup product development meetings were video-taped using a single camera capturing design objects present on the table between the designers and object handling (e.g., holding prototypes or sketching activity), albeit not in detail. Further, gestures and general direction of gaze of the designers could be discerned from the video. The conversation between the designers was audio-taped. During the meetings, the experimenter was present as an observer only. No special instruction to think-aloud was given. The designers were merely asked to continue with the meeting as they normally would.

Following each meeting, design objects (sketches and prototypes) that had been present were videotaped in close-up, sometimes with one of the designers explaining the functioning of the object in voice-over. Each meeting lasted between 30 minutes and 2 hours. Subgroup product development meetings were recorded during the first 5 months of the design project (during the concept design phase). The recordings were transcribed and segmented according to complete thought. A total of 7 transcripts covering approximately 9 hours of video were used in the present data analysis yielding a total of 7414 segments.

Each segment was supplemented with information of what (if any) design object presently in the room that was the focus of attention of the person speaking (i.e., typically a design object such as a sketch or prototype located on the table between the designers). This information was coded for each segment using the video recording of the design session (not the verbal data). Focus of attention was operationalized as either actual handling or holding a particular object; pointing to a particular object; or gazing toward a particular object (if this was possible to discern from the video). The segmented data was the main unit of analysis.

The designers developed multiple design concepts, and most were discarded again. The designers would work on several different design concepts at each meeting, although usually two or three would be the main focus of each session.

Protocol Coding

The transcripts were initially reduced by coding for off-task behavior (e.g., jokes, banter between the designers, office gossip or events unrelated to design), and episodes dealing with summarizing past meetings or planning future meetings or data collection. This coding removed 1602 segments from further analysis. Then the transcripts were coded for type of pre-inventive structure and analogy. All analogies were then coded further for analogical distance and analogical purpose. All transcripts were coded by the first author. Reliability checks were conducted by an independent coder who had received training both in protocol analysis in general and in this coding scheme using spare data from a different subgroup.

Coding of Type of Pre-inventive Structure

For each segment, we coded whether the verbally referenced design object was unsupported by external representations (‘idea’), supported by sketches (‘sketch’) or supported by 3D physical objects in the form of prototypes (‘prototype’). These types of design objects in-the-making will be referred to as ‘pre-inventive structures’ (Finke, Ward & Smith, 1992). These three types exhausted the types of design objects used in the present transcripts. In addition a ‘finalized existing product’ (‘other’) category was used to capture references to existing products on the market, either from the same company or from competitors. These ‘other’ segments, referring exclusively to products already on the market, were excluded from further analysis because they were not considered creative design objects in-the-making.

Coding of Analogy

All segments were coded for analogies, following the method developed by Dunbar (1997; 1995). Any time a designer referred to another base of knowledge to explain, create, modify or evaluate a design, it was coded as an analogy. An analogy was defined as consisting of both an explicit mapping and an explicit transfer. Mapping and transfer can take place in separate segments. For example, an analogy is a statement such ‘This reminds me of one of John’s old ideas, where he put holes in the side of the box. *Could you do something similar here*?’ Here the first sentence maps the old idea to the present context, and the second sentence ensures explicit transfer from the old idea to the new one. Without the last sentence (in italics), this example would not have been counted as an analogy due to a lack of explicit transfer. Without explicit structural transfer the mapping may lead to nothing more than a statement of similarity (A is like B), and this was not counted as an analogy in the present coding scheme.

Coding of Analogical Distance

All analogies were coded for analogical distance. Two levels of analogical distance were used; within-domain and between-domain. Within-domain analogies were defined as analogical mappings within the domain of medical plastics (for example to existing products from the same company or from competitors). Between-domain analogies were analogies made to domains outside of medical plastics (for example to the auto industry, biology, or sports).

Coding of Analogical Function

All analogies were coded for function. Three categories were used. The function could concern identifying possible problems in a new design, taken from an analogous source ('identify problem'). The transfer from source to target thus involved transferring information about potential problems with a design. Another function involved solving design problems ('solve problem'), where the transfer from source to target included elements that would potentially solve a particular design problem that the target was having. Finally, the function of the analogy could be to explain a design to the other designers ('explain'). All analogies could be classified using these three categories.

Results

Inter-rater Reliability

Reliability coding was done on two full transcripts (approximately 18% of the data). All disagreements between coders were resolved through discussion. For each code, a reliability kappa coefficient was calculated.

	kappa
Type of pre-inventive structure	.731
Analogy	.708
Analogical distance	1.000
Analogical function	.802

All codes reached a satisfactory level (>.70), with perfect agreement for analogical distance.

Spontaneous use of Analogies

A total of 102 analogy segments were found in the 7 transcripts (M= 11.3 analogy segments per hour of verbal data). The range of analogy segments was 4 to 21 per transcript, showing that analogies were commonly used by the designers during product development meetings.

Analogical Distance

The analogies were 55% within and 45% between-domain. All transcripts contained both within-domain analogies and between-domain analogies. It thus appears that within-domain and between-domain analogies were used in roughly similar quantities, and that both occurred frequently. Examples of between domain sources included potato print, zippers, credit cards, children’s slides, milk containers, shoes, toilet paper, cars, Christmas decorations, water wheels, picture puzzles, Venetian blinds, and lingerie. In short, a large number of distant domains, seemingly with little or no relation or superficial similarity to medical plastics, were accessed and used during design problem solving.

Pre-inventive Structures and Analogy Use

Figure 1 presents proportion of analogies in segments associated with ideas, sketches, and prototypes. Chi-square results indicated that there were significant differences between conditions ($\chi^2(2)=7.961$, $p<.019$). Subsequent 2x2 chi-squares indicated that ideas had significantly more analogies than prototypes ($\chi^2(1)=7.875$, $p<.006$), while sketches did not differ significantly from either ideas or prototypes (see figure 1). The designers thus produced more analogies when referring to unsupported ideas, than when they referred to prototypes.

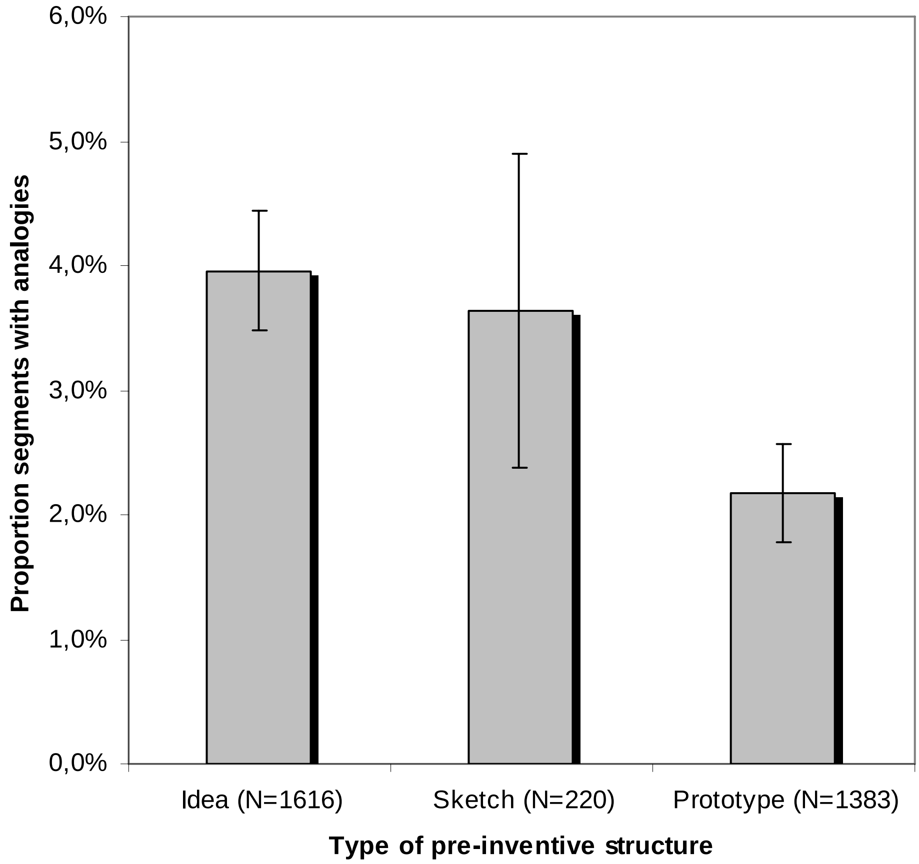


Figure 1. Proportion segments with analogies by type of pre-inventive structure (with SE bars).

To examine the relation between analogical distance and pre-inventive structure, a chi-square between analogical distance (between-domain, within-domain, no analogy) and preinventive structure (ideas and prototype) was conducted (sketches were removed from this analysis due to having an expected count less than 5). Results showed significant differences in analogical distance ($\chi^2(2)=19.255$, $p<.001$). Within domain analogies accounted for 1.7% (S.E.=0.3%) and 1.8% (S.E.=0.4%) for idea and prototype segments respectively. However, between domain analogies accounted for 2.2% (S.E.= 0.4%) and 0.4% (S.E.=0.2%) for idea and prototype segments respectively. This pattern illustrates that ideas and prototypes produced an equal amount of within-domain analogies, but ideas produced significantly more between-domain analogies than prototypes (see figure 2). The overall result that ideas produced more analogies in total compared to prototypes thus appears to be caused by prototypes producing less between-domain analogies.

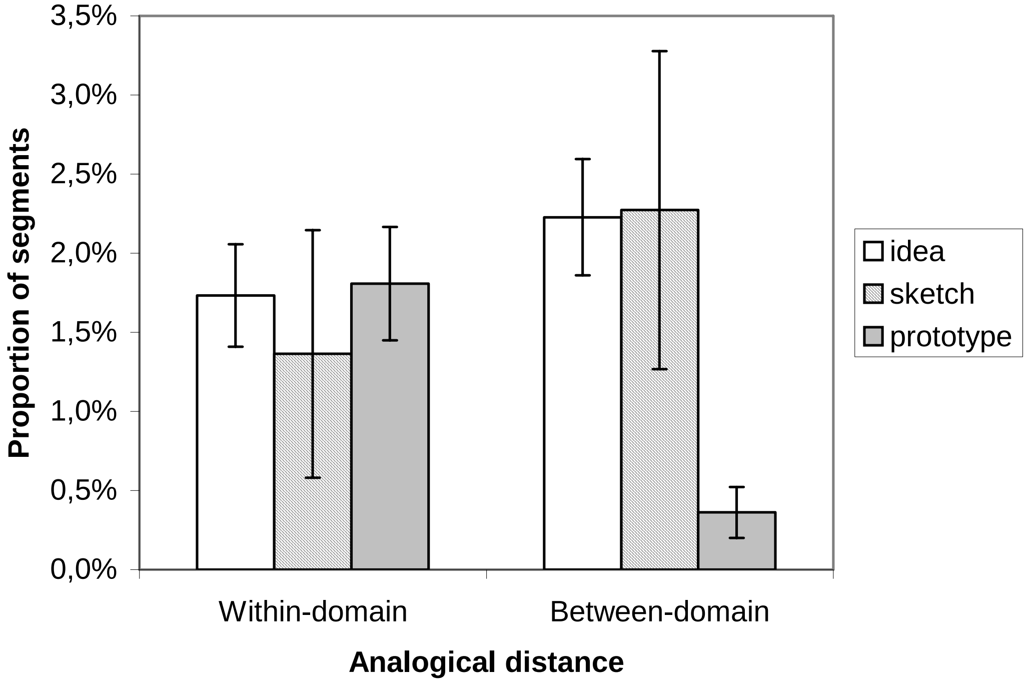


Figure 2. Proportion within-domain and between-domain analogies in the data set by pre-inventive structure (with SE bars).

Chi-square tests involving sketches were made impossible due to expected counts of less than five. So to examine whether the ratio of within to between domain analogies differed between sketches and ideas/prototypes, Fisher's Exact tests were conducted on the number of within and between domain analogies for each pair of pre-inventive structures (see figure 3). Results indicated that, as before, ideas differed from prototypes (two-tailed Fisher's exact test, $p<.001$), with more between domain analogies for ideas. Furthermore, sketches differed from prototypes (two-tailed Fisher's exact test, $p<.019$), with sketches producing more between-domain analogies relative to within-domain analogies. Sketches did not differ significantly from ideas (two-tailed Fisher's exact test, $p=1.000$). Thus, both ideas and sketches produced more between-domain analogies relative to within-domain analogies compared to prototypes.

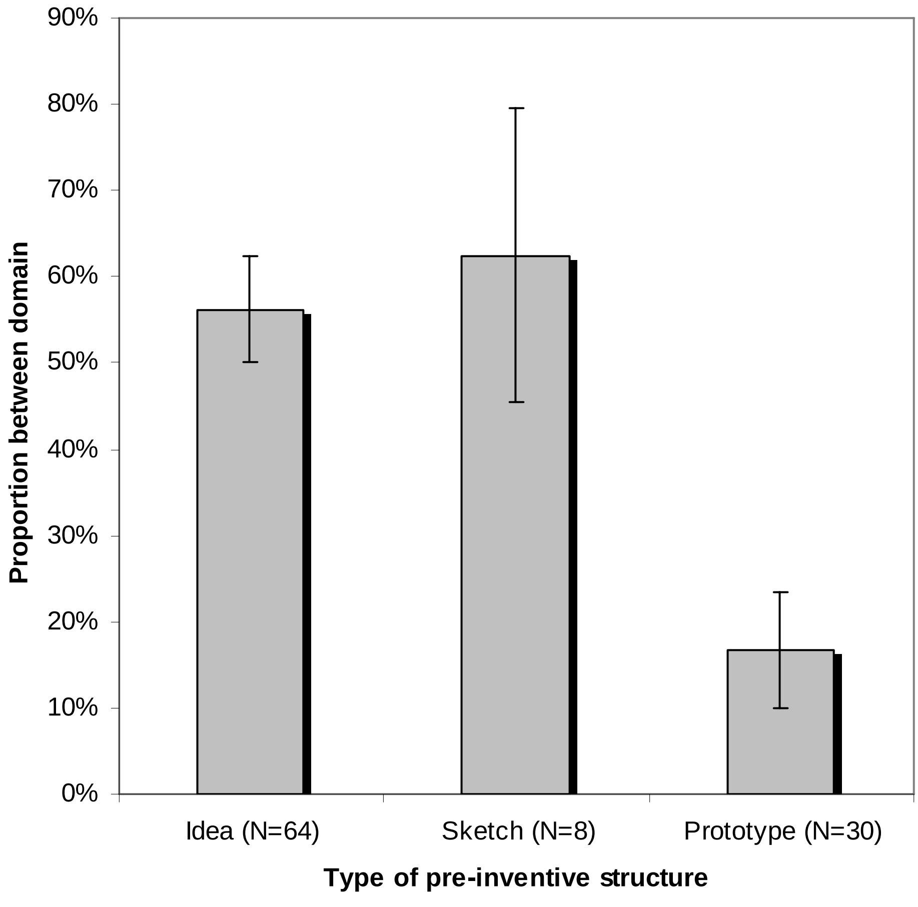


Figure 3: Proportion of between domain analogies by pre-inventive structure (with SE bars).

Analogical Function

The functions of analogies were distributed roughly evenly across the three categories, with 28% of analogy segments focussing on identifying problems, 40% on solving problems, and the remaining 32% with explaining. Analogy thus served all three functions in the present data set.

To examine whether the ratio of within and between-domain analogies varied with analogy function, a chi-square was conducted (see figure 4). There were significant differences between the three types of functions ($\chi^2(2)=20.934$, $p<.001$). Subsequent 2x2 chi-squares revealed that there were significant differences between all pairs of analogical functions; identify problem-solve problem ($\chi^2(1)=6.720$, $p<.010$), solve problem-explain ($\chi^2(1)=6.190$, $p<.013$), and identify problem-explain ($\chi^2(1)=20.834$, $p<.001$). As predicted, the results indicate that analogies used to identify problems primarily utilized within-domain sources, whereas explanatory analogies mainly use between-domain sources. Problem solving analogies concern a mixture of within and between domain analogies.

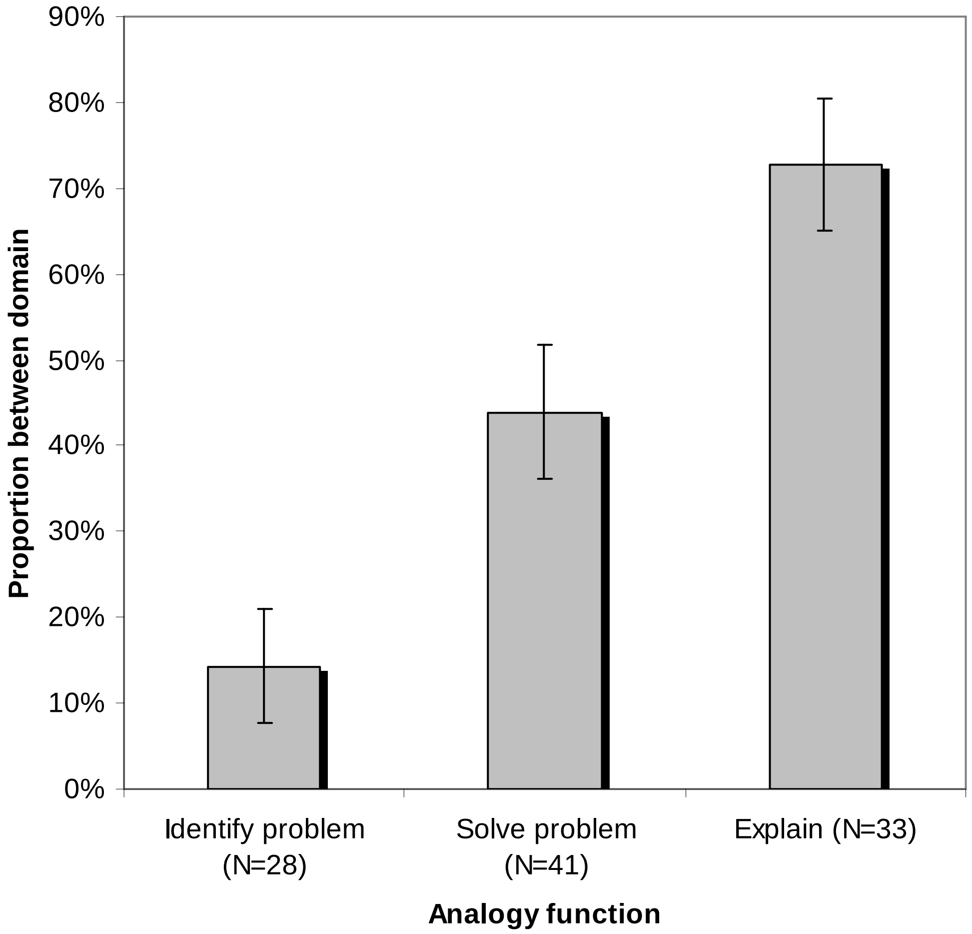


Figure 4: Proportion between domain analogies by analogical function (with SE bars).

A category of special interest to design problem solving is the category related to solving problems utilizing between-domain analogies. This category has traditionally been linked to radical novelty in design problem solving, as evidenced in anecdotes. Two examples[1] are provided to illustrate this kind of analogy in the present data set.

The designers were trying to create a kind of portable slide or tube that could be used to transport liquid from one container to another at a somewhat blunt angle without support. At the

[1] The examples do not provide exact details about the product in question to ensure anonymity for the company, but care has been taken to correctly illustrate the structural requirements. Except to change two concepts that could have endangered anonymity, the protocol exerpts have not been altered in any way.

same time, the device had to be soluble in water over a few hours or days. However, these requirements were problematic in that at least two of the constraints appeared in conflict. Portability required the device to be either somewhat small or at least foldable. But the function of supporting liquid at a blunt angle meant that the obvious choice of material was a hard and sturdy one. Similarly there were a conflict between immediate durability and solubility. How to combine hard and sturdy with foldable and soluble? This led to the following exchange:

Designer

A	what if you made…now Im just saying something…the stuff you make Venetian blinds of for example…they can be bent.
B	yes they can…
A	…folded…
B	Yes
A	…and when you…when they want to…but you can't dump them in water
B	No, 'cause they're made of steel
A	But they are…they have this curve, they can be folded into next to nothing…it's
B	Yes, that's right, but it is …it is really really thin steel, and that's why it goes click-click…

Following this exchange, the designers left the Venetian blinds analogy, not to return. The problem solving attempt apparently failed even though several constraints were met by the analogy (foldable; solid and sturdy) because a further constraint (water soluble) was not met by the suggested analogy.

Another example of between-domain problem solving analogies involved a different design concept. Again a design requirement was it had to be water soluble, but this time the design involved a container for small amounts of liquid, capable of holding the liquid for a few minutes time before falling apart. The problem was finding a suitable material for the design. Most plastics the team discussed were too durable. The following exchange ensued:

Designer

A	But…what's it called…one of these things…[points to an envelope]…if you made it in a paperbag, that will take some time….maybe …perhaps you should…if you could make it look flushable…
B	...but it's going to be hard to flush
A	Well yes but you could…it has to be made flushable…
C	We're going to make it out of paper!

The designers worked on this idea for a novel kind of material for their product for quite some time to follow, conducting experiments using the envelope mentioned in the transcript. Later in the meeting it was decided to examine different types of paper, and to develop prototypes of the product. Apparently the analogy to the envelope was a successful one. In conclusion: the designers did use between-domain analogies in their active problem solving, supporting the view that distant analogies may play an important part in engineering design. However, it should be noted that analogies also served other functions than to solve problems (such as explain concepts and identify problems). Further, within-domain analogies were used just as frequently in solving problems, showing that between-domain analogies were not the only type of analogies used in solving problems in the present data set.

To examine whether communicative alignment was a bigger problem for segments referring to unsupported ideas than for sketches or prototypes, explanatory analogies for ideas vs. sketches and prototypes (grouped) were compared using a chi-square (problem identification and problem slving analogies were removed from the data set). Results indicated a small but marginally non-significant difference ($\chi^2(1)=5.846$, $p=.054$) in the expected direction, even though twice as many analogy segments concerned ideas than sketches and prototypes combined (see figure 5). The hypotheses that referencing ideas should lead to more explanatory analogies than referring to sketches or prototypes was not supported although a small effect was present, but more research will have to be conducted to establish whether more data will prove the small effect reliable and significant.

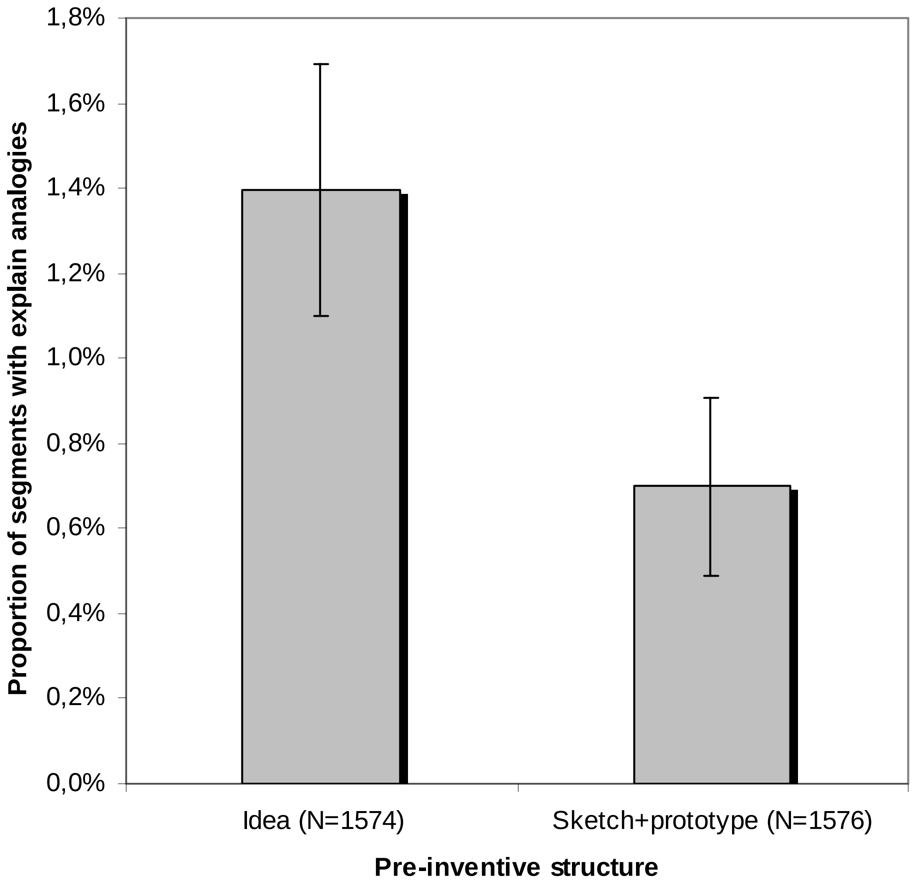


Figure 5: Proportion of segments with explain analogies by pre-inventive structure (with SE bars).

Split-half Reliability Tests

All significant results were subjected to split-half reliability tests. The data was split in two ways. First, to test for temporal effects between transcripts, the transcripts were ranked according to data collection date, and the first half of the collected transcripts was separated from the second half. Second, to test for temporal effects within each transcript, each transcript was split in two halves, and reliability tests were conducted for the first halves and second halves separately. Reliability results showed that in the majority of cases each half of the data yielded significant results in the same direction as the total data set. In some cases, due to low N problems, one of the halves did not yield significant results, or Chi-squares could not be conducted due to expected counts less than 5. But it is important to note that in all cases the directionality of the data was the same as in the total data set. Thus, no time-course confounds were found in the present data set.

Discussion

The current study examined the design processes of a creative, real-world design team in depth to test the interaction between pre-inventive structures and analogical transfer of different forms. While the overall base-rates of different activities may be particular to this particular team or the particular objects they were designing, we take the patterns of the functional roles of pre-inventive structures with respect to analogical transfer to be a general, albeit preliminary, picture of design cognition.

The present study of analogy suggests that spontaneous analogies occur in real-world engineering design, as it has been shown previously in other real-world studies of analogies (e.g., science and politics). This data further support the notion that analogy is used in real-world creative domains, and contrasts the expectation from experimental studies of spontaneous access, were spontaneous analogies are rarely found unless salient superficial similarity between source and target exists.

The present study also showed that, unlike real-world science, between-domain analogies are quite frequent in engineering design, and almost as frequent as with-in domain analogies, suggesting they serve important functions in design cognition (especially for explanations but also for problem solving) as would be predicted from the large number of anecdotes and design tools claiming that between-domain analogies are crucial to design. But as has been found in the domain of science, within-domain analogies also play an important role in everyday designing.

More analogies were made when the designers were not using external representations than when they were referring to prototypes. This finding could be explained by a differential number of between-domain analogies between ideas and prototypes. Previous studies and theories have suggested that making within-domain exemplars available during the creative constructive process tend to lead subjects to unconsciously plagiarize these exemplars when they try to construct novel objects. The present research extended this finding by showing that the within-domain exemplars may even be the designers' own prototypes constraining the creative process, as evidenced in a reduction in the number of between-domain analogies the designers made. When the designers were referring to prototypes they made very few between-domain analogies, and notably fewer than when referring to either sketches or ideas unsupported by external representations. This finding supports the hypothesis that within-domain exemplars constrain creativity by providing paths-of-least-resistance for design analogizing. In other words, if exemplars are present the designers are less likely to think about other domains than the present one. Apparently constraining exemplars in creativity include not only accidental exemplars in the immediate environment, but also include external representations made and used by the designers themselves to (paradoxically) help the creative process along! Notably, however, sketching appeared not to constrain analogical distance in the present data set, which could perhaps in part explain why designers are so prone to sketching.

This finding has potentially important implications for how the early stages of design should be structured. In so far as the designer's current design goal involves generating novel and original products, a tentative recommendation could be to use sketching and idea-generation unsupported by external representation in the early stages of design, and perhaps postpone prototyping until several promising concepts have been developed.

The functions of analogy in engineering design, as revealed by the present study, include both explanation, problem solving, and problem identification. As found in previous research in science, a significant part of the analogies made had an explanatory function in relation to explaining concepts to other designers. It is possible that one of the primary reasons for the importance of explanatory analogies is that in design involving the development of novel concepts, communicative alignment in a group becomes a main concern when the object being referenced exist only in the mind of other team members. This explanation could however not be confirmed in the present data set, even though an effect in the expected direction was present. A large proportion of analogies was made to solve design problems, as would be expected by previous research, and design theorists. But support for a third and new function was also found, in that several analogies concerned identifying problems in novel designs. Here, rather than transferring solutions to the novel design, what was transferred was the expected existence of potential design problems, taken from past sources with which the designers had experience. This type of analogical transfer served a function in the quick evaluation of newly developed design concepts.

These three types of analogy functions in design had differential ratios of within to between domain analogies. As hypothesized, problem identification analogies were mainly within domain; explanatory analogies were more frequently (and mainly) between-domain; while problem solving analogies concerned a mixture of within and between domain analogies. Unlike Dunbar’s (1995; 2001a) findings in science, between-domain analogies were not only made to explain concepts, but concerned problem solving as well. These results indicate the importance of between-domain analogies in real-world engineering design as also claimed in anecdotes and design techniques, but at the same time stress that analogy serves several other functions in design.

References

Anderson, R. E. & Helstrup, T. (1993). Multiple perspectives on discovery and creativity in maind and on paper. In B.Roskos-Ewoldson, M. J. Intons-Peterson, & R. E. Anderson (Eds.), *Imagery, creativity, and Discovery: A cognitive perspective* (pp. 223-252). Amsterdam: Elsevier Science Publishers B.V.

Anolli, L., Antonietti, A., Crisafulli, L., & Cantoia, M. (2001). Accessing source information in analogical problem-solving. *Quarterly Journal of Experimental Psychology: Human Experimental Psychology, 54A,* 237-261.

Ball, L. J. & Ormerod, T. C. (1995). Structured and opportunistic processing in design: A critical discussion. *International Journal of Human-Computer Studies, 43,* 131-151.

Ball, L. J., Ormerod, T. C., & Morley, N. J. (2004). Spontaneous analogising in engineering design: a comparative analysis of experts and novices. *Design Studies, 25,* 495-508.

Bearman, C. R., Ball, L. J., & Ormerod, T. C. (2002). An exploration of real-world analogical problem solving in novices. In W.Gray & C. D. Schunn (Eds.), *Proceedings of the 24th annual conference of the Cognitive Science Society* (pp. 101-106). Mahwah, NJ: Erlbaum.

Blanchette, I. & Dunbar, K. (2000). How analogies are generated: The roles of structural and superficial similarity. *Memory and Cognition, 28,* 108-124.

Blanchette, I. & Dunbar, K. (2001). Analogy use in naturalistic settings: The influence of audience, emotion, and goals. *Memory and Cognition, 29,* 730-735.

Boden, M. A. (1990). *The creative mind: Myths and mechanisms*. New York: Basic Books.

Brown, A. S. & Murphy, D. R. (1989). Cryptomnesia: Delineating inadvertent plagiarism. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 15,* 432-442.

Casakin, H. (2004). Visual analogy as a cognitive strategy in the design process: Expert versus novice performance. *Journal of Design Research, 4*.

Casakin, H. & Goldschmidt, G. (1999). Expertise and the use of visual analogy: implications for design education. *Design Studies, 20,* 153-175.

Christensen, B. T. (2005). A methodology for studying design cognition in the real-wold. In Proceedings from the 1st Nordic Design Research Conference.

Dahl, D. W. & Moreau, P. (2002). The influence and value of analogical thinking during new product ideation. *Journal of Marketing Research, 39,* 47-60.

Dunbar, K. (1995). How scientists really reason: Scientific reasoning in real-world laboratories. In R.J.Sternberg & J. E. Davidson (Eds.), *The nature of insight.* (pp. 365-395). Cambridge, MA, US: The MIT Press.

Dunbar, K. (1997). How scientists think: On-line creativity and conceptual change in science. In T.B.Ward, S. M. Smith, & J. Vaid (Eds.), *Creative thought: An investigation of conceptual structures and processes.* (pp. 461-493). Washington, DC, US: American Psychological Association.

Dunbar, K. (2001a). The analogical paradox: Why analogy is so easy in naturalistic settings yet so difficult in the psychological laboratory. In D.Gentner, K. J. Holyoak, & B. N. Kokinov (Eds.), *The analogical mind: Perspectives from cognitive science.* (pp. 313-334). Cambridge, MA: The MIT Press.

Dunbar, K. (2001b). What scientific thinking reveals about the nature of cognition. In K.Crowley, C. D. Schunn, & T. Okada (Eds.), *Designing for science: Implications from everyday, classroom, and professional settings.* (pp. 115-140). Mahwah, N.J., US: Lawrence Erlbaum Associates,Inc.,Publishers.

Dunbar, K. & Blanchette, I. (2001). The invivo/invitro approach to cognition: the case of analogy. *Trends in Cognitive Sciences, 5,* 334-339.

Finke, R. A. (1990). *Creative imagery: Discoveries and inventions in visualization*. Hillsdale, NJ, US: Lawrence Erlbaum Associates, Publ.

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

Gentner, D. (1998). Analogy. In W.Bechtel & G. Graham (Eds.), *A companion to cognitive science* (pp. 107-113). Malden, MA, USA: Blackwell Publ.

Gentner, D., Rattermann, M. J., & Forbus, K. D. (1993). The roles of similarity in transfer: Separating retrievability from inferential soundness. *Cognitive Psychology, 25:* 524-575.

Ghiselin, B. (1954). *The creative process: A symposium*. Berkeley: University of California Press.

Gick, M. L. & Holyoak, K. J. (1980). Analogical problem solving. *Cognitive Psychology, 12,* 306-355.

Gick, M. L. & Holyoak, K. J. (1983). Schema induction and analogical transfer. *Cognitive Psychology, 15,* 1-38.

Goel, V. (1995). *Sketches of thought*. London: Bradford book.

Goel, V. & Pirolli, P. (1992). The structure of design problem spaces. *Cognitive Science, 16,* 395-429.

Goldschmidt, G. (1991). The dialectics of sketching. *Creativity Research Journal, 4,* 123-143.

Goldschmidt, G. (1994). On visual design thinking: the vis kids of architecture. *Design Studies, 15,* 158-174.

Goldschmidt, G. (2001). Visual analogy: A strategy for design reasoning and learning. In C.M.Eastman, W. M. McCracken, & W. C. Newstetter (Eds.), *Design knowing and learning: Cognition in design education* (pp. 199-220). Amsterdan: Elsevier.

Gordon, W. J. J. (1961). *Synectics: The development of creative capacity*. New York: Harper & Row.

Holyoak, K. J., Gentner, D., & Kokinov, B. N. (2001). The place of analogy in cognition. In D.Gentner, K. J. Holyoak, & B. N. Kokinov (Eds.), *The analogical mind: Perspectives from cognitive science* (pp. 1-20). Cambridge, MA, USA: MIT press.

Holyoak, K. J. & Koh, K. (1987). Surface and structural similarity in analogical transfer. *Memory & Cognition, 15,* 332-340.

Holyoak, K. J. & Thagard, P. (1995). *Mental leaps: Analogy in creative thought*. Cambridge, MA: MIT Press.

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

Johnson-Laird, P. N. (1989). Analogy and the exercise of creativity. In S.Vosniadou & A. Ortony (Eds.), *Similarity and analogical reasoning.* (pp. 313-331). New York, NY, US: Cambridge University Press.

Kaplan, C. A. & Simon, H. A. (1990). In search of insight. *Cognitive Psychology, 22:* 374-419.

Luchins, A. S. (1942). Mechanisation in problem solving. The effect of Einstellung. *Psychological Monographs, 54,* no. 248.

Maier, N. R. F. (1931). Reasoning in humans II, The solution of a problem and its appearance in consciousness. *The Journal of Comparative Psychology, 8,* 181-194.

Marsh, R. L. & Bower, G. H. (1993). Eliciting cryptomnesia: Unconscious plagiarism in a puzzle task. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 19,* 673-688.

Marsh, R. L. & Landau, J. D. (1995). Item availability in cryptomnesia: Assessing its role in two paradigms of unconscious plagiarism. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 21*, 1568-1582.

Marsh, R. L., Landau, J. D., & Hicks, J. L. (1996). How examples may (and may not) constrain creativity. *Memory and Cognition, 24,* 669-680.

Marsh, R. L., Landau, J. D., & Hicks, J. L. (1997). Contributions of inadequate source monitoring to unconscious plagiarism during idea generation. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 23,* 886-897.

Marsh, R. L., Ward, T. B., & Landau, J. D. (1999). The inadvertent use of prior knowledge in a generative cognitive task. *Memory & Cognition, 27,* 94-105.

McGown, A., Green, G., & Rodgers, P. A. (1998). Visible ideas: information patterns of conceptual sketch activity. *Design Studies, 19,* 431-453.

Novick, L. R. (1988). Analogical transfer, problem similarity, and expertise. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 14:* 510-520.

Osborn, A. F. (1963). *Applied imagination*. (3rd revised ed.) New York: Charles Scribner's Sons.

Perkins, D. N. (1981). *The mind's best work*. Cambridge, MA: Harvard University Press.

Perkins, D. N. (1988). The possibility of invention. In R.J.Sternberg (Ed.), *The nature of creativity: Contemporary psychological perspectives.* (pp. 362-385). New York, NY, US: Cambridge University Press.

Purcell, A. T. & Gero, J. S. (1992). Effects of examples on the results of a design activity. *Knowledge-Based Systems, 5,* 82-91.

Reitman, W. R. (1965). *Cognition and thought: An Information-Processing approach*. New York: John Wiley & Sons.

Römer, A., Pache, M., Weisshahn, G., Lindemann, U., & Hacker, W. (2001). Effort-saving product representations in design--results of a questionnaire survey. *Design Studies, 22,* 473-491.

Roozenburg, N. F. M. & Eekels, J. (1996). *Product design: Fundamentals and methods*. Chichester: John Wiley & Sons.

Roskos-Ewoldson, B. (1993). Discovering emergent properties of images. In B.Roskos-Ewoldson, M. J. Intons-Peterson, & R. E. Anderson (Eds.), *Imagery, creativity, and discovery: A cognitive perspective* (pp. 187-221). Amsterdam: Elsevier Science Publishers B. V.

Roskos-Ewoldson, B., Intons-Peterson, M. J., & Anderson, R. E. (1993). Imagery, Creativity, and discovery: Conclusions and implications. In B.Roskos-Ewoldson, M. J. Intons-Peterson, & R. E. Anderson (Eds.), *Imagery, creativity, and discovery: A cognitive perspective* (pp. 313-328). Amsterdam: Elsevier science Publications B.V.

Ross, B. H. (1987). This is like that: The use of earlier problems and the separation of similarity effects. *Journal of Experimental Psychology: Learning, Memory, & Cognition, 13,* 629-639.

Ross, B. H. (1989). Distinguishing types of superficial similarities: Different effects on the access and use of earlier problems. *Journal of Experimental Psychology: Learning, Memory, & Cognition, 15,* 456-468.

Ross, B. H., Ryan, W. J., & Tenpenny, P. L. (1989). The access of relevant information for solving problems. *Memory & Cognition, 17,* 639-651.

Schon, D. A. & Wiggins, G. (1992). Kinds of seeing and their function in designing. *Design Studies, 13,* 135-156.

Shepard, R. N. (1978). Externalization of mental images and the act of creation. In B.S.Randawa & W. E. Cofman (Eds.), *Visual learning, thinking, and communication* (pp. 133-189). New York: Academic Press.

Simon, H. A. (1973). The structure of ill-structured problems. *Artificial Intelligence, 4,* 181-201.

Simon, H. A. & Hayes, J. R. (1976). The understanding process: Problem isomorphs. *Cognitive Psychology, 8,* 165-190.

Smith, S. M., Ward, T. B., & Schumacher, J. S. (1993). Constraining effects of examples in a creative generations task. *Memory and Cognition, 21,* 837-845.

Terninko, J., Zusman, A., & Zlotin, B. (1998). *Systematic innovation An introduction to TRIZ.* Boca Raton: St Lucie Press.

Verstijnen, I. M., van Leeuwen, C., Goldschmidt, G., Hamel, R., & Hennessey, J. M. (1998). Creative discovery in imagery and perception: Combining is relatively easy, restructuring takes a sketch. *Acta Psychologica, 99,* 177-200.

Verstijnen, I. M., van Leeuwen, C., Hamel, R., & Hennessey, J. M. (2000). What imagery can't do and why sketching might help. *Empirical studies of the arts, 18,* 167-182.

Ward, T. B. (1994). Structured imagination: The role of category structure in exemplar generation. *Cognitive Psychology, 27,* 1-40.

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

Ward, T. B. (1998). Analogical distance and purpose in creative thought: Mental leaps versus mental hops. In K.J.Holyoak, D. Gentner, & B. N. Kokinov (Eds.), *Advances in analogy research: Integration of theory and data from the cognitive, computational, and neural sciences* (Sofia: New Bulgarian University.

Ward, T. B., Patterson, M. J., Sifonis, C. M., Dodds, R. A., & Saunders, K. N. (2002). The role of graded category structure in imaginative thought. *Memory and Cognition, 30,* 199-216.

Article 5

How can Participation in Decision Making enhance Creativity in Work Groups?

Christensen, B. T. & Jønsson, T. (submitted for review for anthology in preparation),
H. J. Jeppesen, T. Jønsson, J. Mammen and T. Rasmussen (eds.) Dynamics of participation in work organizations (working title).

Abstract

Participation in decision making has frequently been hypothesized to lead to creativity in work groups. It is argued that cognitive theories and concepts of problem solving and creativity may help ground theories of the relationship between participation and creativity. From theoretical extensions of creative search theories, participation in decision making is hypothesized to lead to creativity through three mediating variables: better integration of diverse groups (leading to more creative product quality), freedom (leading to more creative search activity and quantities of creative products), and commitment to group decisions (leading to more innovation).

Introduction

It is often claimed in the organizational psychological literature that participation in decision making (PDM) leads to creativity and innovation in work groups and organizations. One of the major arguments for introducing participation in work groups has been the argument that participation can lead to enhanced creativity and innovation (Shapiro, 2000; Plunkett, 1990; Kanter, 1983; West, 1990). King (quoted in Anderson, 1992) reviewed many of the studies into the effects of leadership style on innovation, and concluded that the general consensus supports a democraticparticipative style as being a facilitator of innovativeness in work groups. In a large-scale study testing 19 different factors, Campion, Medsker & Higgs (1993) found that selfmanagement and participation were the most predicative of effective group work (although the effects were still modest). But although this link between PDM and creativity has been tested and in part documented by research, theories with clearly stated causal explanations of *why* participation in decision making (PDM) would lead to creativity and innovation are somewhat less common than claims of the existence of the link between PDM and creativity. The mechanisms causing this facilitating effect can seem hazy and unspecified at times in the literature. Why exactly should increased employee involvement in decision making lead to enhanced creativity and innovation?

Some of the suggested mediating factors include such different explanations as enhanced intrinsic motivation (Amabile, 2001; Conti & Amabile, 1999), reduction in resistance to change (De Dreu & West, 2001), pooling of unshared knowledge (e.g., Latham et al., 1994; Jackson, 1996) and better utilization of individual differences in cognitive style (Kirton, 1989), and improved work environment for creativity (e.g., Isaksen, Lauer, Ekvall, & Britz, 2001). It has also been proposed that increased degrees of freedom in work (e.g., Amabile, 2001), minority dissent (e.g., De Dreu et al., 2001), and increased social support and organizational encouragement (e.g., Mumford & Gustafson, 1988; West & Anderson, 1996; Kanter, 1983) are mediating factors between participation and creativity.

Although these theories of how PDM may lead to creativity in work groups seem to point in many different directions, some of the explanatory models could be argued to be converging on somewhat similar explanations. The present article is an attempt to develop new theory and ground some of the hypotheses about the relation between PDM and creativity in cognitive theories of creative problem solving. This grounded theory should be in line with present research findings on participation and creativity. First we will define what we mean by creativity, innovation and participation. Then cognitive theories of creative search are reviewed, and their relation to participation in decision making is suggested.

Definition of Creativity and Innovation

Creativity refers to an area of study ranging over creative persons, products, processes and the study of creative context (or ’press’) (Rhodes, 1961; Mooney, 1963). With such a diverse field of investigation, one could imagine that consensus on a single definition would be scarce. However, in the last couple of decades, some consensus has appeared in the creativity research community concerning what to study: Creativity occurs when someone creates an original and useful product (Mayer, 1999). This definitional emphasis on creative products has helped expand the previous understanding of creativity as merely a trait or a single stage in the creative process (e.g., ’insight’), to the understanding that creativity is first and foremost a matter of changing the world (Feldman, Csikszentmihalyi, & Gardner, 1994) by creating novel and useful products. The products being brought into existence needs to be novel and original, as well as appropriate, useful and adaptive. However, the definition also requires ‘someone’ who carries out the creative act in a series of steps, and of course a context within which the acts occur. The emphasis on creative products has enabled researchers to study creativity in not only individuals, but also groups and organizations. Although the relation between PDM and creativity in principle could be (and have been, see King, 1990 for a review) studied at all these levels, the level chosen for review and examination here is that of the work group.

When focussing on the work group rather than the creative individual the explanatory models for how creative products come into existence is usually directed towards factors in the environment (called antecedent factors) and in the creative process (e.g., West, 1990). Most research in the area has sought such explanations either in the context (e.g., work environment or creative climate), or in the characteristics of the creative process.

Explanations that look at how the *organizational context* or work environment influences creativity typically try to pinpoint a series of mediating variables in the organization that have a strong impact on processes and people, and thereby the production of creative products and innovations. When trying to determine whether PDM enhances the organizational context for creativity, the typical way of doing so would be to try to determine whether PDM facilitates each mediating variable or not.

Explanations linking PDM to the *creative process* typically link to the requirements of either known mechanisms or stages in the creative process, in an attempt to determine how PDM would facilitate the process. The creative process is traditionally described as consisting of at least four stages: preparation, incubation, illumination, and verification (e.g., Wallas, 1926; Ghiselin, 1954; Koestler, 1964; Csikszentmihalyi & Sawyer, 1995) although slightly different names have been used. These stages highlight the problem *solving* nature of creativity, which theorists have argued does not do justice to creative processes that pinpoint creative problems in the first place (Getzels & Csikszentmihalyi, 1976). This has led to an increasing emphasis on problem finding behavior in addition to creative problem solving (Jay & Perkins, 1997; Hoover & Feldhusen, 1994; Runco & Chand, 1994). The traditional description of the creative problem finding and solving process ends with the verification stage, where the product is tested and evaluated, and finalized. However, in real-world organizations creative products needs to be implemented – put into operation – to finalize the process. However, the implementation aspect of creativity includes some processes and requirements we may not intuitively associate with creativity itself, such as a need to communicate the product to the world or domain and pursuade others of it's novelty and usefulness (Simonton, 1988), along with social and practical aspects of it's implementation, such as overcoming resistence to change and creating commitment to the new product. The different requirements and processes operating in the creative process and the implementation stages has led to a conceptual separation between *creativity* which involves the above mentioned creative problem finding and –solving stages and *innovation* which involves the succesful implementation of creative ideas within an organization (Amabile et al., 1996; De Dreu et al., 2001). This form of innovation, where the implemented product is created by the work group itself, has been called 'emergent innovation' (e.g., King & Anderson, 1990; West, 1990) to separate it from innovation that is merely imported from the outside or imposed by management. Some authors have argued that creative processes (such as discovery) and the implementing aspects of innovation may require quite different management strategies (Argote, Gruenfeld, & Naquin, 2001) and organizational structures (Holbek, 1988; for similar arguments on organizational learning, see Weick & Westley, 1996). So even though the progression from the creative process into innovation may seem like a natural one, the distinction is conceptually and practically important. In the present article we will look at how cognitive explanations of problem solving and creativity may help explain the link between PDM and creativity and innovation, while at the same time incorporating previous research and theories related to both organizational context and the creative process.

We still need to define what kind (or rather, what level) of creativity we are talking about. It is possible to discern different levels of creative products by classifying to whom it is novel and useful (e.g., Johannessen, Olsen, & Lumpkin, 2001; Cohen, 1989), whereby creativity reception by other groups becomes an integral part of creativity estimates (Vissers & Dankbaar, 2002). For example, a scientific research group working on creating a novel scientific theory is working on a product that is novel and useful not only to the group itself, but also to the entire research community, the domain and field (Csikszentmihalyi, 1990; Csikszentmihalyi, 1988). Along this line of thinking, a problem solving group (e.g., Quality Circles) working on creating or improving a new product to be implemented and used in daily operations across an entire organization is involved at a lower level of creative endeavour. And finally, an autonomous work group that creates a new way of carrying out their own work, has created a product that is perhaps novel and useful only to the work group itself, which would place the product at yet a lower level of creativity. There are some important differences in the processes between different levels. In particular the implementation stage may differ quite markedly between a scientific research group trying to convince the scientific research community that their discovery is a legitimate one, and an autonomous work group that implements a novel approach to their own work into their daily procedures. This illustrates that the difference between levels mainly concerns whom the product is likely to ‘spread’ to; the group itself; other groups; domain and society. With this caution in mind we will narrow the level of creativity used in the present integrative review down to work group innovation (i.e., lower levels of creativity), where the novel and useful product are implemented in the organization. This level of investigation is in line with other researchers in the area, such as West (e.g., De Dreu et al., 2001; West et al., 1996) and Amabile (e.g., 1988; 2001; Amabile et al.,1996). The extent to which this level of creativity generalizes to higher levels (such as scientific research groups making discoveries or product development groups creating breakthrough inventions for the market) is an empirical question.

Definition of Participation

Having defined creativity and innovation, we also need to define participation, and specify the types and aspects thereof we are talking about in relation to creativity. Heller et al. (1998, p. 15) defined participation as "…a process which allows employees to exert influence over their work and the conditions under which they work". An important distinction in this connection is between direct and representative participation. Direct participation involves the direct involvement of individual employees (e.g., problem solving groups), whereas representative participation involves employee representatives, often selected by unions. For the present purposes we will only deal with direct participation, as that is the kind of participation that seems to be involved in creativity in work groups. In terms of creativity the important thing is for subjects to be involved directly in decision making processes and discussions. Bearing this in mind, a group of elected representatives which are involved directly in decision making processes will also potentially be more creative *in that group setting* – but the work force who have elected or appointed them is not likely to be. Examples of the kind of participation in work groups we will deal with are autonomous and semi-autonomous work teams (also called decision-making work teams), product development teams, Quality Circles and problem solving groups, and project teams. Obviously the degree of participation in decision making varies quite a lot between these types of groups, and the particular organizational context within which they are placed. Comparing Quality Circles with autonomous groups makes the variability clear: Quality Circles are temporary, often voluntary, parallel participation structures with no formal authority invested in them. QC's must seek to have its ideas and solutions accepted by management by virtue of the quality of the solution forwarded and supporting argument (Cordery, 1996). Autonomous groups, on the other hand are permanent, formally constituted work groups with formal authority that interact directly. These groups have the capability of implementing innovations into their own work process without obtaining approval from outside.

But despite these differences we shall try to determine how the participative qualities they share influence creativity and innovation in the work groups.

Cognitive Theories of Problem Solving and Creativity

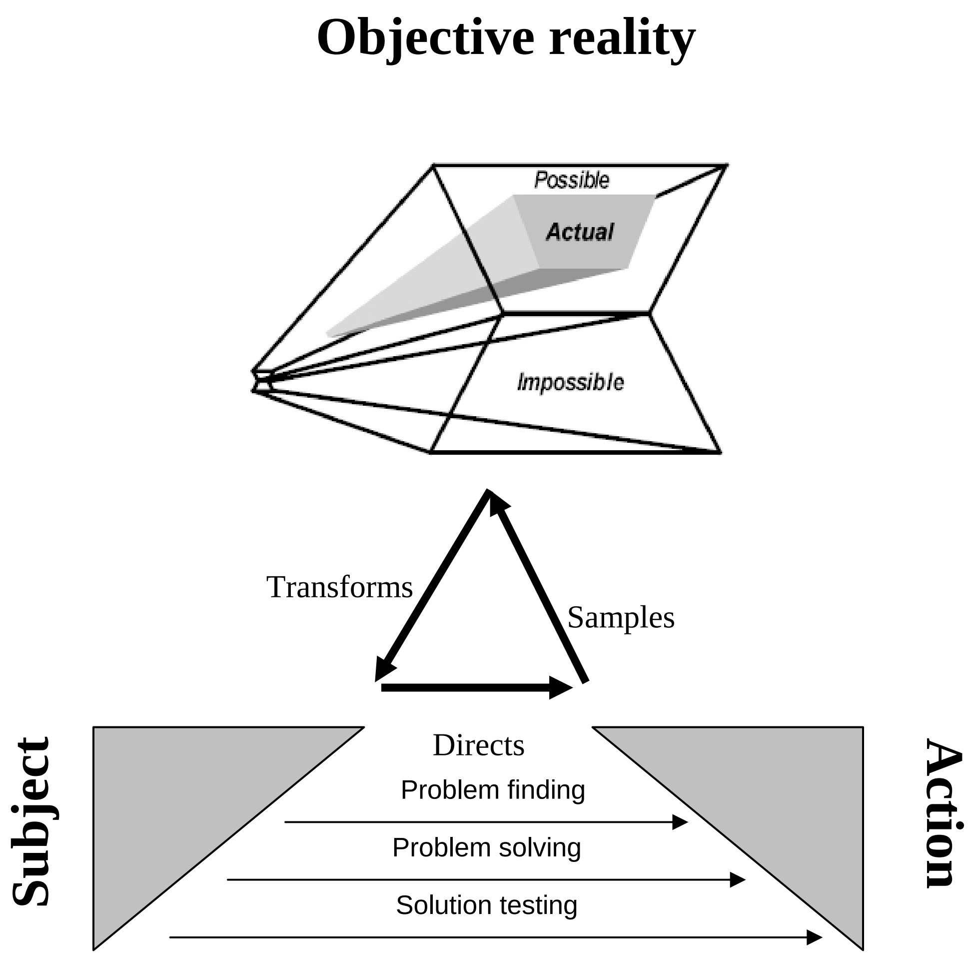
When looking at cognitive theories of creativity, information processing theories of problem solving is a natural place to start. Although the information processing approach to the study of problem solving has been around for a number of years, it is still a dominant paradigm for the cognitive study of problem solving and creativity. It dates back to 1972 when Newell and Simon published their monumental 'Human Problem Solving', in which they attempt to develop a theory of humans as Information Processing Systems (IPS), with problem solving being viewed as a search for a solution to a problem in a problem space. Basically Newell and Simon's model characterizes problem solving as a constrained and guided search through a space of alternative possibilities. Decision-making is viewed as a rational and logical act, viewing the similarities between human and computer processing as being more than a mere metaphor. Both are viewed as systems that process information over time, proceeding in a more or less logical fashion. Human problem solving is viewed as goal-directed search consisting of symbolic manipulation by rule-following systems. An IPS consist of an active processor, input (sensory) and output (motor) systems, and internal Long Term Memory, Short Term Memory, and External Memory. The problem space, and the processes occurring within it, is limited by the capacity of the cognitive system. According to Newell & Simon (1972), a person is confronted with a problem when he wants something and does not know immediately what series of actions he can perform to get it. In many respects, the brilliance of Newell and Simon's problem solving paradigm can be attributed to their concept of space. They define the space in which problems are solved, as not merely the present layout of the problem and world, but extend it into the possible states. Problem solving is a search in a space of possibilities. As Perkins (Perkins, 1994, p. 140) has expressed it: "We think in terms not just of actualities but possibilities of varying payoff and promise." The space in which problems are solved is a space of all possible (allowable) states, when starting from the initial state. The problem space is the internal representation of the space used by the subject in his problem solving, to be distinguished from the environment itself (referred to as task environment). The problem space need not correspond completely with realizable external states (Newell & Simon, 1972, p. 77). Included in the thought processes are also the set of behaviors the IPS considers that prove infeasible, illegal, or in some other way impossible. As such, the subject’s wishes and dreams as well as his more realistic thoughts are included in the problem space (ibid., p. 60).

By engaging in a search the IPS is looking for a path through the problem space that will take him or her to the goal state. As the number of possible paths are usually very high (just consider the possibility space – the number of possible states – in a game of chess) the IPS uses heuristics. Heuristics are rules of thumb (e.g., means-end analysis) that limit the number of options, and are thus strategies for reducing a large number of possible states to a lower number, hopefully without excluding the path to the goal. The search stops when a desired goal is achieved. The criteria needs only to be satisfied (rather than optimized) for the IPS to halt the search process.

Although Newell & Simon’s theory dealt with problem solving in general, several researchers have tried to extend it to a theory of creative problem solving as well (e.g., Perkins, 1981; Boden, 1990; Kaplan & Simon, 1990). Most of the basic insights in problem space theory is maintained in creative problem solving theories: The problem solver is a limited capacity system; problem solving takes place in a knowledge based problem space of possibilities; searching entails skills and strategies.

Newell & Simon’s approach to problem solving stress the importance of knowledge of the task environment, stored in memory. The basic argument that quality and large amounts of knowledge is necessary for and promotes creativity is increasingly acknowledged in modern creativity theories (e.g., Perkins, 1988; Ward, Smith, & Vaid, 1997; Weisberg, 1986). Creativity does not come into being out of nothing ('ex nihilo'), but is grounded in knowledge of the world and it's possibilities. This appreciation of knowledge is also seen in an increased recognition of analogical transfer as a part of creativity. Finding analogous solutions to problems in other domains is an inherently creative process, that can create novel and useful products (Gick & Holyoak, 1980; Holyoak & Thagard, 1995; Holyoak & Thagard, 1997). It need not be less creative just because the source analogue is derived from somewhere (i.e, the same or another domain). However, the opposite relationship between knowledge and creativity has also been put forth: that knowledge may be detrimental to creativity. The problem is that creativity is a very special kind of performance that creates *novel* products, and the argument that the generation of a product is knowledge-based opens up for the discussion of whether a knowledge-based product can be 'novel'. If creativity is viewed as radical novelty without any links to the past, knowledge-driven processes can easily be seen as concerning something other than creativity. The evidence for this claim seem to come from research showing that activating certain information or knowledge components can temporarily inhibit the production of novelty, as cognitive theories of functional fixedness (Duncker, 1945) and mental ruts (Smith, 1995a; Smith, 1995b; Smith & Blankenship, 1991; Smith & Vela, 1991; Smith, Ward, & Schumacher, 1993) have argued. However, this research mainly show temporary detrimental effects of particular kinds of knowledge – and this should not detract from the general conclusion that having quality knowledge of the domain in question in and of itself is a vital and necessary component of creativity.

Information processing theories of problem solving and creativity have been criticized on a number of counts, such as the extensive rationality in the process, and the fact that problem solving and creativity mainly involves changing the mind of the creator, rather than changing the world (thereby largely ignoring the fact that imagined solutions and insights can be wrong). We will not reiterate the criticism here, but rather point out that most of the basic insights of the information processing approach about problem solving can be maintained even in ecological cognitive approaches to creativity (usually seen in opposition to information processing theories), involving perhaps less rational kinds of search. For example, Christensen (2002) attempted to show how an ecological cognitive theory of creativity following in the tradition of Neisser (1976) and Barsalou (1999) could essentially use the basic insights from Newell and Simon to explain creativity as a search in a space of possibilities. With inspiration from Neisser's (1976) perceptual cycle, the developed model was called 'the creative cycle' (Christensen, 2002), and it explains creativity as a cyclical process between cognition and reality, mediated by creative action. However, ecological cognitive frameworks like Neisser's perceptual cycle have most frequently been used to explain cognitive phenomena such as perception and learning (i.e., where the individual does not to any significant extent alter the world) rather than creativity (where that is necessarily the case). When it comes to processes where the subjects alters the real-world (i.e., brings possibilities of the world into being) some remarks are merited as to what constitutes cognition, activity and ontology. The creative cycle appears as follows:



The model is briefly reviewed here – for a more thorough explanation, please refer to Christensen (2002). Basically creativity is viewed as a directed and active search in the real-world. In this connection the real-world is not limited to what is presently and objectively (i.e., positively) existing in the world (such as objects and events), but is extended to the objective although not-existing possibilities and impossibilities of this world. In that sense, creative search action can be said to be *sampling* these possibilities and impossibilities, even though the very same processes are generative, in the sense that they involve recombining cognitive elements, such as entities, events, and categories. The present explanation places possibilities and impossibilities as real-world qualities, thereby creating a synthesis between realist and constructivist theories of creativity, mediated by creative search activity. There are various kinds of creative search activity, including problem finding, problem solving, and solution testing activities. Activity in not limited to ’external’ behavioral search activities (i.e., behavior that actualizes – brings possibilities into being), but include cognitive simulations of variations of the real-world as well, much as the information processing theories hypothesize.

In conclusion, both information processing theories of problem solving and creativity and ecological cognitive models of creativity can incorporate similar explanatory models: creativity is basically a search among possibilities; the search is a constrained search; the cognitive system is limited; the search entails search strategies. Such cognitive models of creativity as search focus primarily on individual subjects (or the interactions between individual subjects and the environment) at the expense of social interaction. Therefore an attempt to explain the impact of participation in decision making on creativity in work groups using creative search models have to involve some theory development that encompasses work group interaction. Here we will make a few conjectures in that regard which will be elaborated in the following sections.

One conjecture that needs to be made concerns the relation between the individual representation or ‘problem space’, and the group representation or ‘problem space’. Here we suggest the following basic relation:

- Individual problem spaces contain both some knowledge that is shared and some that is unshared among team members.
- The total amount of unshared knowledge in a group will vary between groups.
- Communication among team members will help share previously unshared knowledge, thereby increasing the quality of individual problem spaces.

- Without communication unshared knowledge will not be distributed among team members.

Another conjecture concerns constraints imposed by group management:

- Team management imposes constraints on the legitimate search space (i.e., where to search is part decided by team management).
- Team management imposes constraints on search strategy (i.e., how to search is in part decided by team management).

Below we present theoretical extensions based on creative search theories (including the presented conjectures) that may help explain how participation in decision making may lead to creativity in work groups. Support for the theoretical extensions from existing research is also reviewed.

A Theory of Cognitive Exchange: Participation in Decision Making helps utilize Diversity through Integration

According to dominant theories of creative problem solving, having knowledge (and skills) of the domain in question, and it's objective possibilities and impossibilities will help the creative process, as more possible variations of the world can be simulated and attempted, in the search for novel and useful products. Having diversity in work teams is one way to improve the problem space and quality of the knowledge and skills which are used to search for creative solutions. The argument revolves around taking advantage of individual differences in order to create new knowledge; people are to a great extent different: they have different knowledge, experiences, expertise, skills, traits, roles they take on in groups, and ways of doing things. Rather than viewing this diversity as a possible source of conflict and inefficient production (which it can be, see

Jackson, 1996; De Dreu & De Vries, 1997), it is viewed as a source of unshared knowledge and characteristics that can be utilized in creativity and innovation.

Groups can be diverse in a number of ways (e.g., age, sex, skills, experience, values). A useful way of categorizing these kinds of diversity is as either *task-related* (e.g., department and unit membership, formal credentials, education level, physical skills and abilities, job experience) or *relations oriented* (gender, race, ethnicity, age, national origin, social status, attitudes, values, personality) diversity (Jackson, 1996). It is the task-related kind of diversity that has the potential for increasing the problem space, and enhancing the creative potential.

There are two different kinds of explanation of why task-related diversity would lead to creativity and innovation. The first could be termed the *knowledge complementarity hypothesis.* It simply states that pooling knowledge and skills from diverse fields related to the area of interest will increase group performance (e.g., Latham, Winters, & Locke, 1994). The second kind of explanation of why diversity would lead to creativity concerns *social interaction* in the group. Studies of minority dissent suggests that a dissenting opinion in a group will give rise to sense-making attempts of why someone would think that way. This in turn leads to the dissemination, exploration and taking into account of diverse perspectives and new ways of looking at things (Vissers et al., 2002). Of course these two kinds of explanation could be said to be closely related; one focusing on diverse knowledge pooling – the other the social interaction effects of pooled diverse knowledge. The social interaction kind of explanation points to an important aspect of group problem solving: the communication of the unshared task-related diversity.

Simply possessing task-related diversity is not enough for a group. Unless the individual (unshared) knowledge is shared and discussed with team members, no increase of the individual problem space and problem solving activity can be expected. Adding people with diverse knowledge to a group does not increase the common problem space unless they actively engage in group problem solving, idea development, idea sharing and active exchange of information. Task-related diversity in teams needs to be integrated; solutions and unshared knowledge needs to be discussed and evaluated; knowledge needs to be distributed through social interaction and communication. This need for integration in order for diversity to fulfill it's potential for enhancing the quality of creative solutions can come from participation in decision making. PDM can at least under some circumstances lead to more discussion and information sharing of unshared information. Therefore, when task-related diversity needs to be shared between individuals in a group, PDM can perform an integrative role, where the social interrelatedness of the group increases, along with the exchange of information. Without integration, task-related diversity is not shared and utilized. This explanation places a heavy emphasis upon the social integrative function of PDM[1] (communication and interpersonal relationships), and focuses less on the decision making abilities of the group (see e.g., West, 2002). Social integration must not be misunderstood to mean anything along the lines of 'friendship' or other positive connotations of close social interrelatedness, but is purely a measure of the degree of sharing information and skills in the group. This may entail constructive conflict as well, as research on minority dissent has shown (e.g., De Dreu et al., 2001).

The role of participation in this connection is, thus, to increase exchange of unshared task-related information through integrative processes. One could call this a 'theory of cognitive

[1] This is similar to what was examined by West & Anderson (1996). They used a scale to measure 'participation in the team', which included questions relating to the social interrelatedness and communication in the team – rather than questions of their capability to make decisions.

exchange’. It is clear that open discussions, task-related conflicts, and increased social interrelatedness all help this integration. The cognitive-exchange theory predicts that given high work group integration (which *can*, at least under certain circumstances, be improved by PDM), a significant amount of task-related diversity in the group will lead to a high quality of creative products. In so far as the high work group integration leads to an increase in creative problem solving activity, a higher quantity of creative products may also result.

Although such a theoretical model of participation, diversity and creativity has not been tested directly in empirical studies, the various aspects of the theory do find support in existing theories and research. Support for the different hypothesized relations is reviewed below.

An extensive amount of research has shown that minority dissent and group diversity can in fact be creativity enhancing, and a way to counter some of the possible negative consequences of group decision-making (e.g., the ’groupthink’ phenomenon, (Janis, 1972), and tendency to conformity and seeking compliance (e.g., Baron, Kerr, & Miller, 1992)).

At a group level the argument that avoiding too much conformity and utilizing individual differences will lead to creativity and innovation is typically examined by looking at interaction patterns in the group[2]. E.g., does the group attempt to reach a common consensus by ignoring minorities – or is the opposite the case? Minority dissent in groups is an important area of research in this connection. Minority dissent occurs when a minority in a group publicly opposes the beliefs, attitudes, ideas, procedures, or policies assumed by the majority of the group (e.g., De Dreu et al., 2001; Nemeth & Kwan, 1987; Nemeth, 1986). De Dreu & West (2001) found that minority dissent

[2] The argument can be carried out at an organizational level as well. Here the argument becomes one of avoiding too much centralization and formalization in organizations, in order to bring out the creative potential hidden in individual differences (e.g., Ekvall, 1996).

predicted innovation in teams only when the teams had high levels of participation in decision making as well. It was concluded that minority dissent stimulates creativity and divergent thought, which, through participation, manifest as innovation.

A long list of studies have found that diversity in groups increases performance on creative decision-making tasks (see Jackson, 1996 for a brief review) and innovation (West et al., 1996; Campion, Medsker, & Higgs, 1993). Although there have been inconsistent findings, these can probably be attributed to the fact the studies used relations-oriented diversity, rather than task-related diversity. As Jehn et al. (1999, p. 742) has argued: "These inconsistent findings should not be all that surprising. No theory suggests that a workgroup's diversity on outward personal characteristics such as race and gender should have benefits except to the extent that diversity creates other diversity in the workgroup, such as diversity of information or perspective. […]. Even when workgroups do possess that 'other' diversity (e.g., information or perspective), performance benefits should be expected only to the extent that workgroup members successfully manage the difficulties of interacting effectively with dissimilar others […]".

How is task-related diversity related to participation? West (1990; West et al., 1996) argued that whereas diversity primarily increases the quality of creativity and innovations, participation in decision making primarily increases the quantity of innovations, through easing implementation. In two studies De Dreu & West (2001) found that an interaction effect of PDM and minority dissent was predicative of work group innovation. Furthermore, participation alone was only a significant predictor in one of the studies. This *could* be interpreted as support for the above hypothesis that participation is necessary for the utilization of the diverse opinions brought out through minority dissent (although other interpretations are certainly possible). In other words, PDM is not thought to be related to the early stages of the creative process except under very special circumstances (e.g., the continued sharing and attending to one another's ideas) (West, 2002). However, these results seem to be in agreement with a theory of cognitive exchange, stating that when PDM performs integrative functions, task-related diversity in groups will lead to more creative products. But arguing against a theory of cognitive exchange, it should be noted that some studies have found that participation alone in some cases predict creativity or work group innovation (e.g., West et al., 1996; De Dreu et al., 2001, study 1).

Several lines of research support the integrative role PDM can play in groups. Participative decision-making processes have been viewed from a cognitive or communicational perspective. Employees often have expertise on the execution of the work and the management have expertise on strategy, organizing, and policies of the organization. Participation may then integrate the work processes and the leadership processes in organizational decisions, by incorporating relevant knowledge from the different domains in an organization (Seibold & Shea, 2001). In a reanalysis of the review of different involvement programs by Cotton (1993), Seibold & Shea suggests that the desirable effects of participation are contingent on the scope of the communication network: The more spheres and domains of the participation that is connected to the decision-making process, the better and more influential is the decision taken (Seibold et al., 2001).

The integrative feature of participation is in focus in an experimental study by Scully et al. (Scully, Kirkpatrick, & Locke, 1995). They found participation to improve performance, in conditions where one of the parts had correct relevant information. In the extent that this result from the laboratory generates to the complex real world, participation can be viewed as integrating distributed relevant experience.

The cognitive/informational perspective supposes that participation increases quantity and quality of communication and thereby enhance the quality of decisions. Harrison (1985) explored these assumptions in a study of 2 divisions of a large social service agency in an Australian city. 234 employees and 30 immediate leaders from 38 work units participated. The highly participative groups reported significantly better communication than low participation groups. The communication in high participation groups were characterized by more desire to interact, more receiving of interaction, more accurate interaction, less gatekeeping of information (Harrison, 1985, p. 111). This cross-sectional study establishes a close connection between quality and quantity of communication and PDM.

Open discussions will in many (if not most) cases result from implementing participation in decision making. IDE (Industrial Democracy in Europe research group) has developed a taxonomy over influence (table 1). This taxonomy is used to define the extent of participation, from no participation at all (level 1) to full autonomy (level 6) (Heller, 1998). Open debates become an almost integral aspect of participation at the higher levels of participation (level 4 and 5). At the other end of the scale, open debates are unlikely to occur, or will at least be ’inauthentic’, as it will avoid genuine influence-sharing. The IDE study found that higher levels of participation were related to higher frequencies of disagreement. Wilpert (1998) explains this with the increased articulation of the interests of different actors and parties in the participatory decision making process. This articulation of interests also increases the chances that *different* interests are brought up in discussion, resulting in more conflict among stakeholders. This conflict need, however, not be in opposition to creativity (as research on minority dissent also

1. I am not involved at all
2. I am informed about the matter beforehand
3. I can give my opinion
4. My opinion is taken into account
5. I take part with equal weight
6. I decide on my own

Table 1: The IDE-taxomony, from Heller (1998)

shows). The view that conflicts is a natural part of as well as conducive to cooperation is presented and pursued in the research on constructive controversy by Dean Tjosvold (e.g., 1998). The contingency for the conflict to be constructive is that the ultimate goal of the involved parties is cooperative rather than competitive. I.e. the parties are interdependently connected in realizing the goal, rather than competing about the goal in a situation, that resembles a zero-sum game (cf. Deutsch, 1949; 1973). Tjosvold (1987) suggests in a theoretical analysis, that participation legitimates employees to bring problems and issues up in a decision making agenda. The potential of participation to improve both productivity and morale lies in the organization's and the implied parties' ability to manage for and utilize constructive controversy.

Taken together participative organizations experience more disagreement as well as better problem solving. This suggests that the participative organization experiences more 'functional' (what is here referred to as task-related) conflicts. But one must be aware that diversity must be managed well to avoid the potential negative effects of personal (non-task-related) conflicts.

A number of researchers have pointed out that group processes resting on diversity must be managed carefully. Shapiro (2000) noted that if different employees are not involved equally, the program can struggle. He called for a different management approach to dealing with diversity in organizations. Stasser & Titus (1987) argued that much of discussion in groups is devoted to already-shared information. Only under certain circumstances do the discussions revolve around the unshared knowledge of the participants. Argote, Gruenfeld & Naquin (2001) reviewed the literature on what may influence whether unshared information will be shared in group processes, and found that "…information is most likely to be shared in groups when: (a) group members are not burdened with exceptionally high information load, (b) diversity of opinion exists in the group, (c) group members are perceived as having special expertise, (d) groups are relatively small in size, and (e) tasks are perceived as having a demonstrably correct answer such as making a diagnosis." (Argote et al., 2001, p. 381).

Apart from poor intra-group management (e.g., due to lack of training or guidance), lack of resources may also hinder the desired integrative processes of PDM in taking place. Two points need to be made in relation to resources. First, participation in decision making will allow a team greater control over how they utilize the resources they have been allocated (time, money etc.). Insofar as they have the sufficient knowledge and expertise, the resources may be used more effectively. However, and this is point number two, if the resources are scarce – for example if the organization implements participation in decision making with the purpose of cutting costs (e.g., by eliminating middle management staff), the reduction of resources may have a detrimental effect on creativity. Therefore, sufficient resources need to be allocated to a team before participation in decision making can be expected to improve creativity.

In conclusion, there seem to be some support for a theory of cognitive exchange which states that in so far as participation in decision making leads to integrative functions in a group with sufficient task-related diversity, increased quality of creative products may result.

A Theory of Freedom for Creative Search: Participation in Decision Making Increases Freedom and Creative Search Activity

When the creative process is viewed as an active search for novel and useful products among possibilities and impossibilities, an important question becomes what constrains this search. Creativity becomes a matter of exploring, experimenting, and searching the possibilities of the world, but since we do not have direct access to these possibilities, we need to generate and simulate variations of the real-world, in an attempt to throw light upon them. This cognitive simulation process is constrained by a number of things; the purpose of the search, the cognitive elements you have available, the search strategy you are employing, and constraints in the environment.

An additional constraint applies in work team settings. It is hypothesized that subjects in a work setting will only attack creative problems when they have the freedom to do so. ’Freedom’ in this context means that they have the authority and control to attack the problem, and to implement any solution they may come up with. Freedom also implies that subjects are able to move around a search space (i.e., it is considered a legitimate part of their work), in order to find creative problems and solutions.

Several researchers have noted that freedom seem to be essential for creativity (Johnson-Laird, 1988; Ekvall, 1996; Amabile, 2001), even though little empirical research has been carried out to explore the relationship between freedom, creativity and participation. Theorists focussing on the creative climate of organizations have especially highlighted the importance of freedom. Creative climate theorists (such as Göran Ekvall and Teresa Amabile) stress the importance of freedom in creative organizations, which is mirrored in the fact that the associated psychometric instruments (‘Creative Climate Questionnaire’ and ‘KEYS’) used to measure the creative climate include the dimensions of freedom or autonomy. To Amabile, autonomy and freedom concerns “the day-to-day-conduct of the work and a sense of ownership and control over their own work and their own ideas” (Amabile et al., 1996, p. 1161). Ekvalls agrees that freedom is important. His concept of freedom is defined as “the independence in behavior exerted by the people in the organization” (Ekvall, 1996, p 149). To Ekvall, freedom is the opposite of passive rule-bound behavior – i.e.

active transcendence of or challenging rules (ibid.). Rule-bound behavior and lack of autonomy causes the individuals and groups in the organization to ‘walk the same paths’ in problem solving. This mean trying the same problem solving strategies rather than to create new and better ways of understanding and solving the problem. The construct has gained some support from a study by Amabile and Gryskiewicz (1987). In interviews with 120 R&D scientists, 74% mentioned freedom as an important factor to creativity. Freedom can be linked to the construct of risk taking in that the ability to take risk is dependent upon having the freedom to do so. Keywords are a high tolerance of uncertainty and the ability to act quickly on opportunities (Ekvall, 1996). Amabile also supports the important role of encouragement of risk taking at all levels in the creative organization (Amabile et al., 1996). Amabile stresses that one of the reasons freedom and risk-taking is important is that they enable creative exploration to take place. Exploration in a work setting involves searching for problems and solutions and having the freedom to decide over, primarily, the means and ways one wants to carry out a creative exploration (but not necessarily freedom to decide ends or outcomes). This can be contrasted with a non-explorative approach to problem solving (an ‘algorithmic’ approach), which means that the steps and ways to solve the problem follows established rules and procedures – much like calculating by algorithms (Amabile, 1983).

How is freedom related to PDM? A major effect of implementing participation in decision making is the increased freedom of a group or team to decide how to carry out it’s work (and in some cases also concerning outcomes). In turn, such freedom enables the team to explore, experiment and try out new and potentially useful products and ways of working. By allowing employees to participate in decision making, problem solving and other areas traditionally viewed as ’managerial tasks’ new possibilities open up. One could say that PDM comes with the ability to change things. Guidelines and procedures are no longer merely something laid down by management, but rather subject to discussion and change. And it is quite possible, and even expected, for employees to experiment, discuss and explore both old and new approaches to their work. The resulting learning, problem solving and developing potential is one of the most sought after aspects of implementing participation in organizations. By allowing the individual to participate in, and exert influence on, his own daily work life to a greater extent, new actions become possible, and freedom increases.

A consequence of this capability for decision-making is the awareness that implementation of novel approaches into their own work process no longer requires management approval (as long as they can be kept within existing resources, and structural requirements, such as organizational regulations for work procedures).

Being given the freedom and responsibility of a greater part of their daily work life, comes with the awareness that any problems that arise are theirs to solve. This includes any problems they can pinpoint themselves. Increased capability for decision making can thus lead to increased innovation, and the tendency to notice and/or identify, create and put into use novel and useful products and procedures that would have been left alone, had the capability not been there. In this way the freedom to implement solutions sparks off problem findings activities, where the subject notices problems in his daily work life, and then attempts to solve them creatively. This implies the tacit dimension of problem finding in PDM. As individuals move around in their work spaces, the freedom to change routines and structures will lead to noticing problems, and ways of doing things better.

It should be noted here that the particular structure of the PDM group has important implications for the ability to implement solutions. For example, Quality Circles established to find solutions to problems, may end their creative process with a solution that have to be implemented across groups in the organization – a situation demanding managerial and co-worker support, and requiring persuasion. But at least as far as autonomous groups that create with the purpose of implementation into their own daily work life is concerned, the above consequences seem valid.

Behind the theory that freedom will increase creative search activity is an assumption that could be summarized as follows: When individuals are given freedom and control under reasonable conditions they will attempt to utilize expertise and knowledge to improve their own situation. This goes for creativity and learning alike. ’Reasonable conditions’ include resources (such as time and money), but also organizational and team support in general. The theory could be referred to as the ’freedom for creative search’ hypothesis. Freedom may thus increase the quantity (but not necessarily quality) of creative products generated. The amount of these creative products that are implemented (and thus lead to more work group innovations) will depend on factors such as the quality of the products, work team support, commitment to group decisions, and the availability of resources.

Having reviewed two cognitive extensions of creative search theories hypothesizing why PDM may lead to increased creativity in work groups, we will now turn to a different kind of explanation of the relation between PDM and creativity that does not follow closely from cognitive search theories.

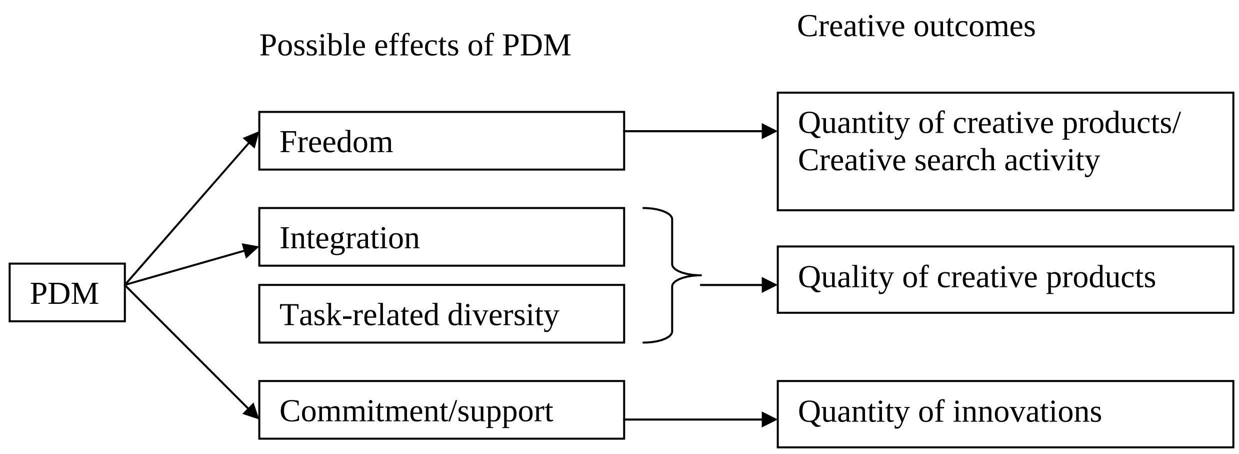
Participation in Decision Making Increases Commitment and Responsibility to Group Decisions and Implementations.

Creative search theories focus primarily on individual creativity, and do not concern the implementing stages of innovation to any great extent. However, in relation to work group creativity, an important part of the process involves the implementation of innovation in the work group. Some researchers have argued that a primary benefit of PDM in relation to creativity is that PDM enhances commitment and responsibility to team decisions (e.g., Steel & Lloyd, 1988), and reduces resistance to change (King, Anderson, & West, 1991; West et al., 1996). These notions build on classical research in social and industrial psychology (see West & Anderson, 1996, for a brief review). This finding has led researchers to argue that a primary function of PDM in relation to creativity is to ensure that creative products are successfully implemented (e.g., West, 2002). In this respect, PDM does not influence creative processes, but rather ensures that creative products are successfully implemented, and possibly that *more* creative products are implemented. The sharing of information, increased social interrelatedness, and influence on the process characteristic of PDM leads to greater commitment to and investment in group decisions (King et al., 1991; Kanter, 1983).

Hypotheses Derived from the three Explanations of the Relation between Participation and Creativity

Above we have sketched out three different ways participation may influence creativity: through cognitive-exchange, through freedom for creative activity, and through commitment to group decisions. What should be clear from the above discussions is that the three explanatory models each influence creativity in different ways. In that respect they each lead to different kinds of ’outcome variables’. The cognitive-exchange theory holds that integration and task-related diversity together leads to an increased quality of creative products (but not necessarily quantity of creative products, or an increase of the quality of implemented innovations). ’Freedom for creative search’ states that increased freedom leads to increased search activity, as well as an enlarged problem space to search. This, in turn, leads to an increase in the quantity of creative products (but not necessarily increased quality – or increased quantity of innovations). And finally, commitment to group decisions leads to an increase in the quantity of implemented innovations (but not necessarily an increase in creative search activity, of the number or quality of creative products generated).

The primary finding is thus that there are three different ways participation in decision making may influence the creative process, and each of these explanations will have a different impact on creative outcome variables (creative search activity, creative products and implemented innovations). See below for a sketch.



An open question concerns possible interaction effects between the three different ways participation in decision making may influence creativity. This type of explanation opens up for making further hypotheses of what would happen if various constellations of mediating factors were to occur. A simple hypothesis would for example be that only when commitment to group processes exist will an increase in freedom lead to an increase in the quantity of implemented innovations. However, this is not the place to completely outline all possible interaction effects between these variables. Future research will have to make predictions of specific interaction effects, and test their validity. For now we will settle with the overall hypothesis, that when all three types of mediating factors are present (freedom, integration & task-related diversity, commitment to group processes) only then will all of the outcome variables be positively affected (creative search activity, quality and quantity of creative products, quality and quantity of implemented innovations). If one or more of the mediating factors does not result from participation in decision making, one or more of the outcome variables will not be positively affected.

Conclusions and Future Research

Organizational psychological literature has hypothesized and in part established that participation in decision making is positively related to creativity in work groups. Several causal explanations for this relationship have been offered. To ground some of these different causal explanations in cognitive theories it is suggested that creative problem solving is viewed as a search in a knowledge-based problem space of possibilities and impossibilities, involving constraints and search strategies. In work group creativity this also involves constraints on search strategies and legitimate search spaces imposed by team management and communication among team members. From these theoretical conjectures, a fairly coherent theory of the relationship between participation and creativity could be put forth.

This framework lists three different ways participation in decision making can influence aspects of the creative process and outcomes. First, *cognitive exchange* theory predicts that given high work group integration (which *can*, at least under certain circumstances, be improved by PDM), a significant amount of task-related diversity in the group will lead to a high quality (but not necessarily a high quantity) of innovations. Second, *freedom for creative search* predicts that insofar as participation in decision making leads to more freedom and control for a group, this will positively benefit the creative processes in the group. This occurs as the objective and legitimate search space for the group is enlarged, and the creative search activity (exploration and experimentation; problem finding, problem solving, solution testing activity) of the individuals are positively affected. This, in turn, will lead to a greater quantity (but not necessarily quality) of creative products. Third, *commitment to decision making* predicts that insofar as participation leads to greater commitment to group decisions, a greater quantity of work group innovations can be expected, as more of the creative products are implemented.

A prediction from the present review is that each of the three theories lead to different outcome variables, all associated with creativity. Only when all three (integration and diversity, freedom, and commitment) are present in a given organizational implementation of participation in decision making, can it be expected that creative search activity, and the quality and quantity of creative products and innovations are all positively affected. Although supported by some previous research, the present theoretical model remains somewhat speculative, and the details should be tested directly in experiments. Further, future research will have to determine what kind of interaction effects are to be expected given different combinations of the presence of diversity and integration, freedom, and commitment.

References

Amabile, T. M. (1988). From individual creativity to organizational innovation. In K.Gronhaug & G. Kaufmann (Eds.), *Innovation: A cross-disciplinary perspective* (pp. 139-166). Oslo 6,Norway: Norwegian University Press.

Amabile, T. M. (1983). The social psychology of creativity: A componential conceptualization. *Journal of Personality and Social Psychology, 45,* 357-376.

Amabile, T. M. (2001). How to kill creativity. In J.Henry (Ed.), *Creative Management* (2nd ed., pp. 4-10). London: Sage.

Amabile, T. M., Conti, R., Coon, H., Lazenby, J., & Herron, M. (1996). Assessing the work environment for creativity. *Academy of Management Journal, 39*, 1154-1184.

Amabile, T. M. & Gryskiewicz, S. S. (1987). *Creativity in the R&D Laboratory* (Rep. No. Technical Report 30). Center for Creative Leadership.

Anderson, N. (1992). Work group innovation: State-of-the-art review. In D.M.Hosking & N. Anderson (Eds.), *Organizational change and innovation: Psychological perspectives and practices in Europe* (pp. 149-160). Florence,KY,US: Taylor & Francis/Routledge.

Argote, L., Gruenfeld, D., & Naquin, C. (2001). Group learning in organizations. In M.E.Turner (Ed.), *Groups at work: Theory and research. Applied social research* (pp. 369-411). Mahwah,NJ,US: Lawrence Erlbaum Associates,Inc.,Publishers.

Baron, R. S., Kerr, N. L., & Miller, N. (1992). *Group process, group decision, group action*. Pacific Grove, CA, US: Brooks/Cole Publ. Co.

Barsalou, L. W. (1999). Perceptual symbol systems. *Behavioral and Brain Sciences, 22:* 577-660.

Boden, M. A. (1990). *The creative mind: Myths and mechanisms*. New York: Basic Books.

Campion, M. A., Medsker, G. J., & Higgs, A. C. (1993). Relations between work group characteristics and effectiveness: Implications for designing effective work groups. *Personnel-Psychology., 46:* 823-850.

Christensen, B. T. (2002). *The creative process and reality. An analysis of search and cognition in the creative process and a call for an ecological cognitive framework for creativity.* (vols. 5, no. 3 - Psykologisk Studieskriftserie) Aarhus, Denmark: Department of Psychology, University of Aarhus.

Cohen, L. M. (1989). A continum of adaptive creative behaviors. *Creativity Research Journal, 2,* 169-183.

Conti, R. & Amabile, T. M. (1999). The impact of downsizing on organizational creativity and innovation. In R.E.Purser & A. Montuori (Eds.), *Social creativity. Volume 2* (pp. 209-234). Cresskill, NJ, US: Hampton press, Inc.

Cordery, J. L. (1996). Autonomous work groups and Quality Circles. In M.A.West (Ed.), *Handbook of Work Group Psychology* (pp. 225-246). New York: John Wiley & Sons, Inc.

Cotton, J. L. (1993). *Employee Involvement: Methods for improving performance and work attitudes*. Newbury Park, CA : Sage Publishers.

Csikszentmihalyi, M. (1988). Society, culture, and person: A systems view of creativity. In R.J.Sternberg (Ed.), *The nature of creativity: Contemporary psychological perspectives.* (pp. 325-339). New York, NY, US: Cambridge University Press.

Csikszentmihalyi, M. (1990). The domain of creativity. In M.A.Runco & R. S. Albert (Eds.), *Theories of Creativity* (pp. 190-212). Newbury Park, CA: SAGE Publications Incorporated.

Csikszentmihalyi, M. & Sawyer, K. (1995). Creative insight: The social dimension of a solitary moment. In R.J.Sternberg & J. E. Davidson (Eds.), *The nature of insight.* (pp. 329-363). Cambridge, MA, US: The MIT Press.

De Dreu, C. K. W. & De Vries, N. K. (1997). Minority dissent in organisations. In C.K.W.De Dreu & N. K. De Vries (Eds.), *Using conflict in organizations* (pp. 72-86). London: Sage.

De Dreu, C. K. W. & West, M. A. (2001). Minority dissent and team innovation: The importance of participation in decision making. *Journal of Applied Psychology, 86:* 1191-1201.

Deutsch, M. (1949). A Theory of Cooperaation and Competition. *Human Relations, 2* , 129-152.

Deutsch, M. (1973). *The Resolution of Conflict. Constructive and destructive processes.* New Haven and London : Yale University Press.

Duncker, K. (1945). *On Problem-solving.* Westport, CT, US: Greenwood Press, Publ.

Ekvall, G. (1996). Organizational climate for creativity and innovation. *European Journal of Work & Organizational Psychology, 5,* 105-123.

Feldman, D. H., Csikszentmihalyi, M., & Gardner, H. (1994). *Changing the world: A framework for the study of creativity*. Westport, CT, US: Praeger Publishers/Greenwood Publishing Group, Inc.

Getzels, J. W. & Csikszentmihalyi, M. (1976). *The creaive vision. A longitudinal study of Problem Finding in art*. New York, US: John Wiley & Sons, Inc.

Ghiselin, B. (1954). *The creative process: A symposium*. Berkeley: University of California Press.

Gick, M. L. & Holyoak, K. J. (1980). Analogical problem solving. *Cognitive Psychology, 12,* 306-355.

Harrison, T. M. (1985). Communication and Participative Decision Making: An Exploratory Study. *Personnel Psychology, 38,* 93-115.

Heller, F. (1998). Playing the devil's advocate : Limits to influence sharing in theory and practice. In F.Heller, E. Pusic, G. Strauss, & B. Wilpert (Eds.), *Organizational participation: Myth and reality* (pp. 144-189). Oxford: Oxford University Press.

Heller, F., Pusic, E., Strauss, G., & Wilpert, B. (1998). *Organizational participation: Myth and reality*. Oxford: Oxford University Press.

Holbek, J. (1988). The innovation design dilemma: Some notes on its relevance and solutions. In K.Gronhaug & G. Kaufmann (Eds.), *Innovation: A cross-disciplinary perspective* (pp. 253-277). Oslo 6,Norway: Norwegian University Press.

Holyoak, K. J. & Thagard, P. (1995). *Mental leaps: Analogy in creative thought*. Cambridge, MA: MIT Press.

Holyoak, K. J. & Thagard, P. (1997). The analogical mind. *American Psychologist, 52:* 35-44.

Hoover, S. M. & Feldhusen, J. F. (1994). Scientific problem solving and problem finding: A theoretical model. In M.A.Runco (Ed.), *Problem finding, problem solving, and creativity.* (pp. 201-219). Stamford, CT, US: Ablex Publishing Corp.

Isaksen, S. G., Lauer, K. J., Ekvall, G., & Britz, A. (2001). Perceptions of the best and worst climates for creativity: Preliminary validation evidence for the Situational Outlook Questionnaire. *Creativity Research Journal, 13:* 171-184.

Jackson, S. E. (1996). The consequences of diversity in multidisciplinary work teams. In M.A.West (Ed.), *Handbook of Work Group Psychology* (pp. 53-75). New York: John Wiley & Sons, Inc.

Janis, I. L. (1972). *Victims of groupthink*. Boston: Houghton Mifflin.

Jay, E. S. & Perkins, D. N. (1997). Problem finding: The search for mechanism. In M.A.Runco (Ed.), *The creativity research handbook, volume one* (pp. 257-293). Creskill, NJ: Hampton Press.

Jehn, K. A., Northcraft, G. B., & Neale, M. A. (1999). Why differences make a difference: A field study of diversity, conflict, and perfomance in work groups. *Academy of Management Journal, 44,* 741-763.

Johannessen, J.-A., Olsen, B., & Lumpkin, G. T. (2001). Innovation as newness: What is new, how new, and new to whom? *European Journal of Innovation Management, 4,* 20-31.

Johnson-Laird, P. N. (1988). Freedom and constraint in creativity. In R.J.Sternberg (Ed.), *The nature of creativity: Contemporary psychological perspectives.* (pp. 202-219). New York, NY, US: Cambridge University Press.

Kanter, R. M. (1983). *The change masters: Innovation and entrepreneurship in the American corporation*. New York, NY, US: Simon & Schuster, Inc.

Kaplan, C. A. & Simon, H. A. (1990). In search of insight. *Cognitive Psychology, 22:* 374-419.

King, N. (1990). Innovation at work: The research literature. In M.A.West & J. L. Farr (Eds.), *Innovation and creativity at work: Psychological and organizational strategies* (pp. 15-59). New York,NY: John Wiley & Sons.

King, N. & Anderson, N. (1990). Innovation in working groups. In M.A.West & J. L. Farr (Eds.), *Innovation and creativity at work: Psychological and organizational strategies* (pp. 81-100). New York,NY: John Wiley & Sons.

King, N., Anderson, N., & West, M. A. (1991). Organizational innovation in the UK: A case study of perceptions and processes. *Work and Stress, 5:* 331-339.

Kirton, M. J. (1989). Adaptors and innovators at work. In M.J.Kirton (Ed.), *Adaptors and innovators. Styles of creativity and problem solving.* (pp. 51-71). London, England: Routledge.

Koestler, A. (1964). *The act of creation*. London: Hutchinson.

Latham, G. P., Winters, D. C., & Locke, E. A. (1994). Cognitive and motivational effects of participation: A mediator study. *Journal of Organizational Behavior, 15:* 49-63.

Mayer, R. E. (1999). Fifty years of creativity research. In R.J.Sternberg (Ed.), *Handbook of creativity* (pp. 449-460). Cambridge, UK: Cambridge University Press.

Mooney, R. L. (1963). A conceptual model for integrating four approaches to the identification of creative talent. In C.T.Taylor & F. Barron (Eds.), *Scientific creativity: Its recognition and development* (pp. 331-340). New York: Wiley.

Mumford, M. D. & Gustafson, S. B. (1988). Creativity syndrome: Integration, application, and innovation. *Psychological Bulletin, 103:* 27-43.

Neisser, U. (1976). *Cognition and reality*. San Francisco, CA, US: W. H. Freeman and Co.

Nemeth, C. J. (1986). Differential contributions of majority and minority influence. *Psychological-Review., 93:* 23-32.

Nemeth, C. J. & Kwan, J. L. (1987). Minority influence, divergent thinking and detection of correct solutions. *Journal of Applied Social Psychology, 17:* 788-799.

Newell, A. & Simon, H. A. (1972). *Human problem solving*. Englewood Cliffs, NJ: Prentice-Hall.

Perkins, D. N. (1981). *The mind's best work*. Cambridge, MA: Harvard University Press.

Perkins, D. N. (1988). The possibility of invention. In R.J.Sternberg (Ed.), *The nature of creativity: Contemporary psychological perspectives.* (pp. 362-385). New York, NY, US: Cambridge University Press.

Perkins, D. N. (1994). Creativity: Beyond the Darwinian paradigm. In M.A.Boden (Ed.), *Dimensions of creativity.* (pp. 119-142). Cambridge, MA, US: The MIT Press.

Plunkett, D. (1990). The creative organization: An empirical investigation of the importance of participation in decision-making. *Journal of Creative Behavior, 24:* 140-148.

Rhodes, M. (1961). An analysis of creativity. *Phi Delta Kappan, 42,* 305-310.

Runco, M. A. & Chand, I. (1994). Problem finding, evaluative thinking, and creativity. In M.A.Runco (Ed.), *Problem finding, problem solving, and creativity.* (pp. 40-76). Stamford, CT, US: Ablex Publishing Corp.

Scully, J. A., Kirkpatrick, S. A., & Locke, E. A. (1995). Locus of Knowledge as a Determinant of the Effects of Participation on Performance, Affect, and Perceptions. *Organizational Behavior and Human Decision Processes, 61* , 276-288.

Seibold, D. R. & Shea, B. C. (2001). Participation and Decision Making. In F.M.Jablin & L. L. Putnam (Eds.), *The New Handbook of Organizational Communication. Advances in Theory, Research, and Methods* (pp. 665-703). Thousand Oaks: Sage Publications.

Shapiro, G. (2000). Employee involvement: Opening the diversity Pandora's box? *Personnel Review, 29:* 304-323.

Simonton, D. K. (1988). Creativity, leadership, and chance. In R.J.Sternberg (Ed.), *The nature of creativity: Contemporary psychological perspectives.* (pp. 386-426). New York, NY, US: Cambridge University Press.

Smith, S. M. (1995a). 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. 135-156). Cambridge, MA: The MIT Press.

Smith, S. M. (1995b). Getting into and out of mental ruts: A theory of fixation, incubation, and insight. In R.J.Sternberg & J. E. Davidson (Eds.), *The nature of insight.* (pp. 229-251). Cambridge, MA: The MIT Press.

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

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

Smith, S. M., Ward, T. B., & Schumacher, J. S. (1993). Constraining effects of examples in a creative generations task. *Memory and Cognition, 21,* 837-845.

Stasser, G. & Titus, W. (1987). Effects of information load and percentage of shared information on the dissemination of unshared information during group discussion. *Journal of Personality and Social Psychology, 53:* 81-93.

Steel, R. P. & Lloyd, R. F. (1988). Cognitive, affective, and behavioral outcomes of participation in Quality Circles: Conceptual and empirical findings. *The Journal of Applied Behavioral Science, 24,* 1-17.

Tjosvold, D. (1987). Participation: A Close Look at Its Dynamics. *Journal of Management, 13* , 739-750.

Tjosvold, D. (1998). Cooperative and competitive goal approach to conflict: Accomplishments and challenges. *Applied Psychology: An International Review, 47,* 285-342.

Vissers, G. & Dankbaar, B. (2002). Creativity in multidisciplinary new product development teams. *Creativity and Innovation Management, 11,* 31-42.

Wallas, G. (1926). *The art of thought*. London, England: Jonathan Cape.

Ward, T. B., Smith, S. M., & Vaid, J. (1997). Conceptual structures and processes in creative thought. In T.B.Ward, S. M. Smith, & J. Vaid (Eds.), *Creative thought: An investigation of conceptual structures and processes.* (pp. 1-27). Washington, DC, US: American Psychological Association.

Weick, K. E. & Westley, F. (1996). Organizational learning: Affirming an oxymoron. In S.R.Clegg & C. Hardy (Eds.), *Handbook of organization studies* (pp. 440-458). Thousand Oaks,CA,US: Sage Publications,Inc.

Weisberg, R. W. (1986). *Creativity. Genius and other myths.* New York, NY, US.

West, M. A. (1990). The social psychology of innovation in groups. In M.A.West & J. L. Farr (Eds.), *Innovation and creativity at work: Psychological and organizational strategies* (pp. 309-333). New York,NY: John Wiley & Sons.

West, M. A. (2002). Sparkling fountains or stagnant ponds: An integrative model of creativity and innovation implementation in work groups. *Applied Psychology: An International Review, 51:* 355-387.

West, M. A. & Anderson, N. R. (1996). Innovation in top management teams. *Journal of Applied Psychology, 81:* 680-693.

Wilpert, B. (1998). A view from psychology. In F.Heller, E. Pusic, G. Strauss, & B. Wilpert (Eds.), *Organizational participation: Myth and reality* (pp. 40-64). Oxford: Oxford University Press.

Abstract in English

The present thesis concerns creative cognition, and notably spontaneous analogizing and incubation effects.

Spontaneous analogizing: Previous literature on analogical problem solving has frequently found disappointingly low levels of spontaneous access, unless salient superficial similarity exists between source and target. In an experimental condition using analogous insight problems on uninformed subjects, the present thesis demonstrated a significant spontaneous access rate when access was compared to a baseline calculated from the same subjects typical problem solving behavior – even when controlling for subjects who may have become informed ('caught on') during the experiment. In an in vivo study of real-world design problem solving, frequent spontaneous analogizing was found, with roughly equal amounts of within-domain and between-domain analogizing. In partial support for theories of unconscious plagiarism and Ward's Path-of-least-resistance model, it was found that the reference to exemplars (in the form of prototypes) significantly reduced the number of between-domain analogies between source and target, compared to using sketches or no external representational systems. Analogy served three functions in relation to novel design concepts: identify problems, solve problems or explain concepts. Problem identifying analogies were mainly within-domain; explanatory analogies were mainly between-domain; while problem solving analogies were a mixture of within and between domain analogies.

Incubation effects: Previous experimental incubation effects studies have had problems finding reliable interactive incubation effects when using RAT items (remote associate test) or visual analogies. In an experimental setup using analogous insight problems, a strong analogical incubation effect was found, demonstrating that analogical problem solving shows reliable incubation effects. It is argued that the experimental incubation research paradigm suffers from a

number of problematic assumptions. The operational definition of incubation effects is off the mark when attempting to generalize to the incubation phenomenon as seen in anecdotal evidence since it includes studies not typically associated with incubation and excludes issues not dealing with performance. Incubation effects may not have a single common cause, but rather multiple, leaving non-competing theories. Incubation effects may have no special relationship to creativity, but concern problem solving in general. Despite these problems, standard incubation effect studies may still be interesting in that they violate expectations from working memory and forgetting research.

Finally, it is argued that based on cognitive explanations of creativity, participation in decision making may enhance creativity in work groups through the mediating variables of freedom, integration of diverse groups, and commitment.

Limitations of the present studies and possibilities for future research are discussed.

Resume på dansk

Denne afhandling omhandler kreativ kognition, og fokuserer specielt på spontan analogisering og incubations effekter.

Spontan analogisering: Hidtidig literatur har ofte fundet manglende evidens for spontan adgang i analogisk problem løsning, med mindre tydelig overfladisk lighed eksisterer mellem den analog der overføres til og fra. I denne afhandling blev der fundet en signifikant spontan adgang i et eksperiment med analoge indsigtsproblemer og uinformerede forsøgspersoner når adgang blev sammenlignet med en baseline udregnet efter de samme forsøgspersoners typiske problemløsningsadfærd – også selv om der blev kontrolleret for om forsøgspersonerne under eksperimentet opdagede at nogle af opgaverne kunne hjælpe dem i deres problemløsning.

Hyppig spontan analogisering blev også fundet i et in vivo studie af design problemløsning, med cirka lige store mængder tætte og fjerne analogier. I delvis støtte til teorier om ubevidst plagiering og Ward’s sti-med-mindst-mulig-modstand model blev det fundet at referencer til objekter (prototyper) signifikant reducerede antallet af fjerne analogier sammenlignet med tegninger eller ideer der ikke blev understøttet eksternt.

Analogier havde tre funktioner i relation til nyt design: identificere problemer, løse problemer og forklare begreber. Problem identifikation var primært tætte analogier; forklaringer var primært fjerne; og problem løsning var en blanding af tætte og fjerne analogier.

Inkubations effekter: Hidtidig eksperimentel forskning om inkubations effekter har haft problemer med at finde evidens for pålidelige interaktive inkubations effekter ved brug af RAT (remote associate test) opgaver eller visuelle analogier. I et eksperiment med analoge indsigtsopgaver blev der fundet en stærk analog inkubations effekt. Der argumenteres for at den eksperimentelle inkubations forskning lider under en række problematiske antagelser. Den operationelle definition af inkubations effekter rammer forkert i det omfang den søger at generalisere til inkubationsfænomenet som beskrevet i anekdotiske beretninger derom eftersom den inkluderer forskning der ikke typisk associeres med inkubation og ekskluderer områder der ikke omhandler performance. Inkubations effekter har muligvis ikke een enkelt årsag, men derimod flere, hvilket efterlader ikke-konkurrerende inkubationsteorier. Inkubations effekter har muligvis ingen speciel relation til kreativitet, men omhandler derimod problemløsning generelt. På trods af disse problemer er standard inkubations effekt studier stadig interessante idet de bryder forventninger fra teorier om arbejdshukommelsen og glemsel.

Til sidst argumenteres der på baggrund af kognitive forklaringer af kreativitet for at participation muligvis forbedrer kreativitet i arbejdsgrupper igennem de medierende variable frihed, integration af grupper med interne forskelligheder, og commitment.

Begrænsninger af disse studier og muligheder for fremtidig forskning diskuteres.