Proceedings of DETC 2003:
ASME 2003 International Conference on Design Theory and Methodology
September 2-6, 2003, Chicago, IL

DETC2003/DTM-48679

EMPIRICAL STUDIES OF DESIGN IDEATION: ALIGNMENT OF DESIGN EXPERIMENTS WITH LAB EXPERIMENTS

Jami J. Shah[1]
Department of Mechanical & Aerospace Engineering
Arizona State University
Tempe, Arizona 85287-6106

Steven M. Smith
Department of Psychology
Texas A & M University
College Station, Texas 77843-4235

Noe Vargas-Hernandez
Department of Mechanical & Aerospace Engineering
Arizona State University
Tempe, Arizona 85287-6106

David R. Gerkens
Department of Psychology
Texas A & M University
College Station, Texas 77843-4235

Muqi Wulan
Department of Industrial and Manufacturing Systems Engineering
Beijing University of Aeronautics and Astronautics
Beijing, 100083, China

ABSTRACT

Although various Idea Generation (IG) methods exist for conceptual design, the ideation process is still hardly understood. There is a need for a Design Ideation Model that explains the variables and processes occurring during IG. Cognitive Science provides models and theories, but these are usually derived from simple tasks or problems. On the other hand, Design Research simulates real world design better, but experimentation at the design level is time consuming and is difficult to isolate due to interactions of the variables involved. This paper introduces an approach for the alignment of experiments at the design level with lab experiments in cognitive psychology. Two key concepts that make this alignment possible are: ideation components (mechanisms believed to promote IG) recognized in Design Research and Cognitive Science, and uniform measures. The long-term objective of this research is the creation of a Design Ideation Model; this will require the testing and modeling of several of these ideation components. This paper presents results from Design and Lab Experiments for a selected component: incubation. Results are discussed and their significance explained in the context of the Design Ideation Model. This study found that, based on the correlation at both the Lab and Design Experiments, incubation had a positive impact on Design Ideation. Further, the alignment approach followed proved to be appropriate for the individual modeling of ideation components.

1. INTRODUCTION

The long-term objective of this study is the development of a model of Design Ideation for conceptual Engineering Design. In this empirical study, data on Design Ideation is being collected from experiments conducted at multiple levels, from highly controlled psychological experiments to design exercises in a realistic setting. A critical element is the "alignment" of experiments at multiple levels; each level varies in terms of internal (i.e. cause and effect) and ecological validity. This paper presents the details on how to achieve this alignment. It is estimated that 70% of a product's cost is defined during conceptual design (Pahl and Beitz, 1996). However, few methods exist that aid engineers at the conceptual stage; the bulk of the product development time and effort is dedicated to later stages of design. Improved methods to aid engineers at the conceptual stage would be of considerable benefit to industries, in particular techniques that help explore the design space in search of good alternatives. Without good alternatives, subsequent design stages (e.g. analysis, testing, decision making, etc.) have no significant impact. Compared to design analysis methods, design synthesis methods are scarce and less understood. Further, the few available are neither theoretically based nor empirically substantiated. The development of a model of Ideation in design will have multiple benefits. First, an increased understanding of the interaction of human variables, design problem variables and method variables, and relating cognitive processes to creative outcome. This can lead

[1] Corresponding author at jami.shah@asu.edu

educators to find better ways of teaching design synthesis. Second, the development of new theoretically based Idea Generation (IG) methods will replace ad-hoc methods. For example, companies can determine which method to use under given conditions and how to constitute design teams. Finally, the collaboration of Engineering Design research and cognitive science fields offers interesting possibilities for future developments in Design Ideation.

2. BACKGROUND

2.1. Design Idea Generation Methods

Various design IG methods have been developed over the past four decades; a comprehensive classification can be found in Shah (1998), Shah et al. 2000, Kulkarni (1998), and VanGundy (1988). These methods have no theoretical basis, and the little empirical evidence of their effectiveness is generally frail. Nowadays various companies, such as NTELLEK (Sickafus, 1998), Ideation (2003), and TRIZ Consulting (2003), market methods, tools, and training for IG. Because there is a need for these kinds of services, these companies are employed by corporations, big and small. The effectiveness of their methods needs to be objectively studied. In order to develop theoretically based design IG methods or conduct experimental studies of their effectiveness; one must look at theories of creative cognition that contain cognitive models of perception, retrieval, mental blends, etc. Such models developed by cognitive scientists are based on controlled experiments in which they isolate mental processes using simplified tasks. However, design is multifaceted involving complex tasks that result in large number of interacting processes. Also, one must relate cognitive processes to desirable outcome. Scaling up of existing cognitive models requires radical departure from the norms in cognitive science.

2.2. Cognitive Research on Idea Generation

Various cognitive models of creativity have been developed; a review of these models can be found in Shah et al. (2000), Koestler (1964), Wallas (1926), Simonton (1988), Finke et al. (1992), Smith et al. (1995), Ward et al. (1997), Smith (1995), Martindale (1995), and Langley and Jones (1988). These cognitive models/theories have been derived from controlled experiments that often use very simple tasks or problems. The suitability of using these models for design problems that are much more complex has never been investigated. Research in cognitive experimental psychology has examined the effects of incubation in various types of insight problems, such as puzzles or brain-teasers. Although these types of insight problems tend to lead thinking in the wrong direction initially, and require flexibility in one's representation of a problem, they are nonetheless convergent tasks with unambiguous solutions.

2.3. Engineering Design Research

Experimental methods, such as Case studies (Altshuller, 1984; Marples, 1960; Ward and Sobek, 1996), Protocol studies (Ericsson and Simon, 1984; Waldron and Brook, 1994; Ullman et al., 1989; Ullman et al., 1988; Christiaans and Dorst, 1991; Cross et al., 1991), and Controlled tests (Schön, 1991), have been used for studying the design process and/or its associated cognitive activities. A survey of these methods can be found in Shah et al. (2000), Christiaans and Dorst (1991), Christiaans and Venselaar (1991), and Schön (1991). Compared to cognitive research, Engineering Design research results better simulate real world design (i.e. less controlled environment, more complex tasks, closer to engineering design). One disadvantage, in general, is that experimentation and analysis is too time consuming (e.g. protocol studies). For each extra variable and interaction considered, the work required increases considerably, this because the same designer or team cannot be tested with the same problem but different methods. Another disadvantage is that the results, being empirical, have natural limitations, for example, results cannot be extrapolated to different conditions since there is little understanding of the behavior of the variables involved. In the past, no standard framework, or measures, has been available to present and compare results. In the present experiment, we examined the effects of incubation on open-ended tasks, including divergent thinking (in the laboratory task), and a design task that was intended to require creative thinking.

2.4. Cognitive Components

To overcome the limitations of cognitive research and Engineering Design research, a quasi-experimental approach was developed at Arizona State University (ASU) (Shah et al., 2000; Shah et al., 2003). This required the study of cognitive processes related to key components (Kulkarni, 2000) of design IG and interactions between them. These key components (e.g. incubation) are mechanisms that are believed to intrinsically promote IG or to help designers overcome specific mental blocks. Many IG methods contain various of these components - deliberate or unintentional. Explanations for *why* and *how* components are effective has been researched (Kulkarni and Shah, 1999) from accepted models of atomic processes and structures already available in cognitive science. Evaluating specific IG methods in their entirety is complicated, the reason for this being that many components are at play simultaneously. The alternative was to identify these components and test them individually. The effectiveness of specific IG methods could then be predicted by the components present in the method. This approach helped better understand IG in general, but since it was purely empirical (i.e. had no theoretical foundation), it had major limitations: (1) inability to discriminate between necessary and superfluous components of Design IG methods, (2) prohibitive number of experiments required, (3) inability to extrapolate experimental results to different environment, design problem, and human variables. To overcome these problems, the present study is theoretically based. Cognitive processes related to key components of Design IG and interactions between them are studied. The processes are related to outcome (i.e. the metrics used are common). These three elements (i.e. the components, their interactions, and the outcome) help in developing a model of Design Ideation. The availability of such a theory will allow us to not only evaluate existing Design IG with fewer surgically targeted experiments but to also refine them by eliminating their superfluous components. One may even be able to develop new methods derived from this theory.

3. FUNDAMENTAL ISSUES

A large body of knowledge on creative processes exists in cognitive science. However, the models in cognitive theories are derived from highly controlled Lab Experiments involving simple and isolated tasks. As shown in Figure 1, there is little similarity between the conditions for these experiments and design concept generation in the real world. Relying purely on Lab Experiments and models of "atomic" cognitive processes derived from such experiments is, therefore, not appropriate for our study. On the other hand, direct experiments of Design IG methods, such as those conducted in the past, simulate real world design better but are unable to discriminate between necessary and superfluous components, require prohibitive number of experiments, and are unable to explain the performance of methods under different conditions. The major difference between real world design and simulated Design Experiments are organizational factors, such as incentives, organizational structure, resource constraints, etc. These are not within the scope of this study. The main issue is how to combine the strengths of Cognitive Psychology and Engineering Design research to obtain a model of Design Ideation. This paper explains the overall research approach and shows partial results that exemplify the first stage (i.e. how to align Lab Experiments and Design Experiments and how to compare results). Future work will discuss other stages in the development of a model of Design Ideation.

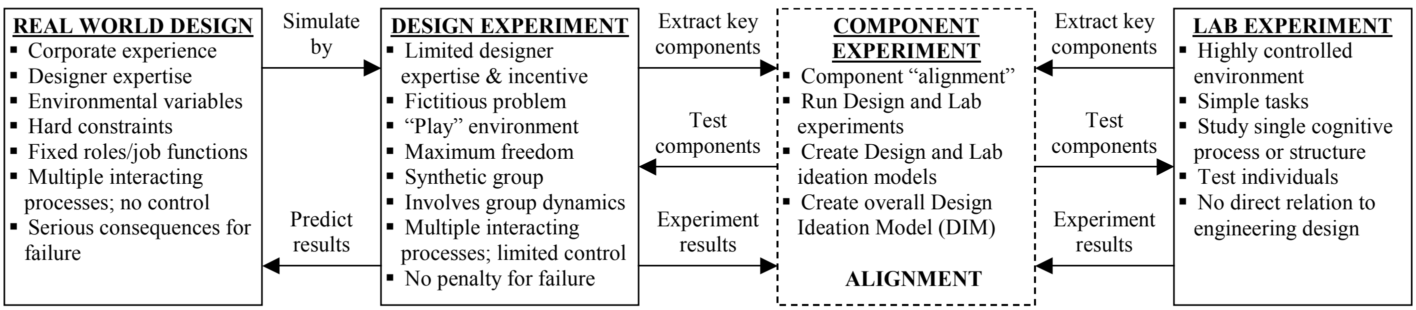


Figure 1 - Levels of Ecological Validity (DMI-0115447, 2001)

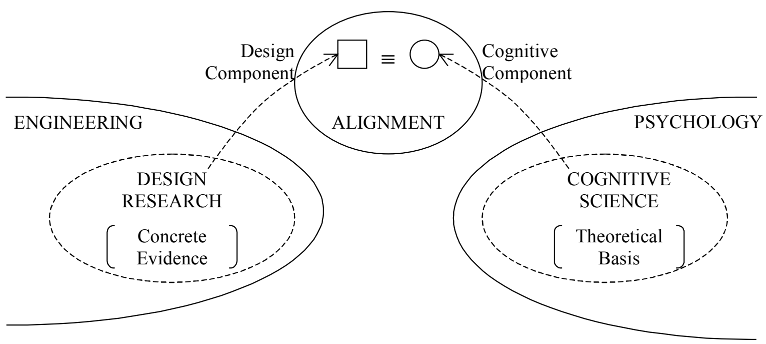


Figure 2 - Overcoming the Science-Engineering Dichotomy

4. RESEARCH APPROACH

Cognitive Science provides the theoretical basis to understand the ideation components, but this understanding is general and simplified. Design Research, on the other hand, provides evidence of the use of ideation components, but this evidence is specific and complex. This situation is pictured in Figure 2. The novel strategy proposed is to have a continuum of experiments across cognitive science (Lab Experiments) and Design Research (Design Experiments). This can only be possible if there is an "alignment" (i.e. agreement) between these two areas. The first key aligning concept are the cognitive components. Based in Cognitive Psychology and recognized in Design Research, they allow both levels to have an agreement through (conceptually) equivalent components. The second key aligning concept are the outcome metrics: quantity, quality, novelty and variety. These metrics characterize the effectiveness of each component. By using the same metrics there is a basis for comparison across components and levels.

The general procedure is to break Design IG methods into key components (macro-processes and structures), develop component models, generalize the effect of each component in different environments, and model interactions between components. Instead of conducting experiments by using a Design IG method in its entirety, each method is decomposed into its key components and its overall effectiveness is predicted by experimentally studying the effectiveness of its components and their mutual interactions. Therefore, if designers are allowed to use a few selected components at a time in a given experiment, it would be possible to assess how effective these components are individually in promoting IG and whether they are influenced by the presence of other components. During experimentation, designers are subject to one or more components at a time; this combination of components makes a design IG method that may resemble existing IG methods. The effectiveness of specific IG methods may then be predicted in terms of the key components that are built into the procedure and whose chances of occurrence are promoted by the use of the method. To predict the effectiveness of an IG method in terms of the effectiveness of individual components that are built into it, it is necessary to first conduct experiments on the effectiveness of each component and to identify possible interactions between different components. Lab Experiments are directed at studying key components individually, while the Design Experiments examine the interactions between components. Therefore, two levels of experiments are defined: Lab Experiments in controlled artificial environment settings and Design Experiments in simulated design setting. A model of Design Ideation will be built in four stages, as shown in Figure 3.

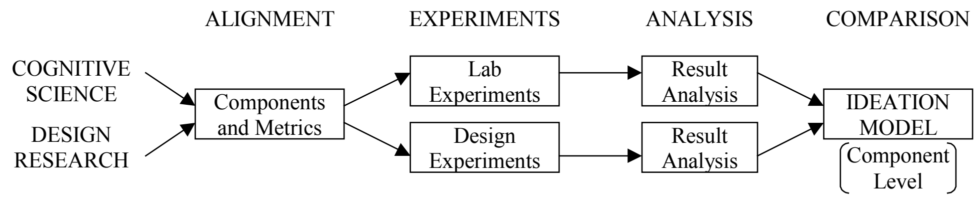


Figure 3 – Research Phase I – Modeling of Ideation Components

Table 1 - Involved Components

COMPONENT	DESCRIPTION	TREATMENT	EXAMPLE METHOD
PROVOCATIVE STIMULI	Excite ideas by exposing the subject to a concept idea	Expose subjects to conceptual (i.e. not detailed) ideas	C-Sketch, 635
SUSPENDED JUDGMENT	Postpone reaching decisions or making conclusions of an idea	Focus on quantity or quality of ideas	PMI, Brainstorming
FLEXIBLE REPRESENTATION	Unconstrain the manner in which ideas are represented	Constraint subjects to text or sketch and then allow freedom of representation	C-Sketch more flexible than 635
FRAME OF REFERENCE SHIFTING	Change in the basic set of ideas on which other ideas are interpreted	Change the frame of reference of the problem	Inversion, Synectics
INCUBATION	Period of time that elapses generation of ideas for a problem	Stop thinking about the problem for a period of time	Can be added to a method
EXAMPLE EXPOSURE	Excite ideas by exposing the subject to a model idea	Expose subjects to complete (i.e. detailed) ideas	Gallery

Table 2 - Component Levels

	EXPERIMENT MANIPULATION	
COMPONENT	LEVEL 0	LEVEL 1
PROVOCATIVE STIMULI	No-exposure	Exposure
SUSPNDED JUDGMENT	Quantity	Quality
FLEXIBLE REPRESENTATION	Sketch only or Text only	Free
FRAME OF REF. SHIFTING	No-change	Change
INCUBATION	No-interruption	Interruption
EXAMPLE EXPOSURE	No-exposure	Exposure

4.1. Component Alignment

The number of IG components identified in past studies (Shah et al., 2000) is more than a dozen. Because of limited research resources and the current NSF project being limited to three years, a subset of ideation components were selected for study. Table 1 lists the components together with a brief description, experimental treatment, and an example IG method that uses it.

Two levels are considered for each component. Although more levels could be defined, it is recommended (Montgomery, 2001) to run experiments initially with few levels and, if needed, increase the levels in additional experiments. Table 2 shows how these levels will be manipulated at least in principle. The Design Research side (ASU) and the Cognitive Science side (TAMU) agreed on these components to be tested at the two levels, Design Experiments and Lab Experiments, respectively.

4.2. Metrics Alignment

The experimental measures or effectiveness metrics used in the Design and Lab Experiments have been described at length by Shah et al. (2000). Four classes of operating variables were considered to characterize the design problem and the environment. Two fundamental values were used in judging the worth of a conceptual design method:

- How effective is it in expanding the design space.
- How well does it explore this space.

Based on that, four independent effectiveness measures were proposed (Shah et al., 2003): *quantity*, *quality*, *novelty*, and *variety* of the ideas generated using that method. *Quantity* is the total number of ideas generated by a group when it uses a certain IG method. *Quality* is a measure of how close it comes to meeting the design specifications. *Novelty* is a measure of how unusual or unexpected an idea is as compared to other ideas. *Variety* is a measure of the explored solution space during the IG process. For Lab Experiments, because of the simplicity of the ideas, these measures can be straightforwardly derived from the results. For Design Experiments, the process is more elaborate. Depending on the level of detail, evaluation is done in one or two stages: at conceptual level and embodiment level and the design artifact is decomposed into its desired key functions, and weights assigned to each.

Quantity and variety scores apply to the entire IG session, while novelty and quality scores are computed for each idea. The total quality and novelty scores are found by multiplying each idea by its respective score in that category and summing all of them to get the overall score for that category (novelty or quality). It does not make sense to consolidate the scores for all four measures into an overall effectiveness measure. Each of the four is a very different type of value and adding them directly makes no sense. Even if we were to normalize them in order to add, it is difficult to understand the meaning of such a measure. It is not always the case that all four are equally important. Besides, we may be interested in knowing how one method (i.e. a combinations of components) compares to another in terms of quantity vs. novelty, etc. It can also be argued that a method is worth using if it helps us with any of the measures.

4.3. Design Of Experiments

Full factorial design of experiments (DOE) explores all possible combinations among variables and their levels (Montgomery, 2001; Dean and Voss, 1999). For example, for three variables (A, B, and C) with two levels each (1 and 0 for high and low levels), the full factorial (i.e. 2^3 factorial design) required runs are listed in Table 3.

Note that embedded in this experiments are three main effect experiments, three two-factor experiments, and two three-factor experiments. If one were to study all possible interactions among the six selected components shown in Table 2, a full factorial design of experiments would be required; this means $2^6 = 64$ experiments (this assuming 2 levels for each component). A preliminary analysis suggested that the first three components (i.e. provocative stimuli, suspend judgment, and flexible representation) shown in Table 2 could be first studied individually as simple comparative experiments; this means one experiment (with two runs) per component. Interaction experiments could be run afterwards if needed. With respect to the other three components (frame of reference shifting, incubation, and exposure to examples), there is a special interest from the cognitive psychology point of view (Janson and Smith, 1991) to study their interactions; this means a 2^3 factorial design (i.e. 8 runs). Table 4 summarizes the proposed experiments for each component. Design Experiments run at ASU and Lab Experiments at Texas A&M University (TAMU).

4.4. Analysis and Comparison

As previously explained, the components to test were identified, and the outcome metrics established. The actual modeling occurs in two steps, first, the interaction effects are examined from the results for the Lab and Design Experiments separately using Analysis of Variance (ANOVA). The ANOVA permits an analysis of each variable averaged across all other variables (i.e. main effects), as well as interactive effects among variables (i.e. two-factor, three-factor interactions, etc.). Second, the interaction models from Lab Experiments and Design Experiments are compared. This critical step is possible because the tested components and the metrics used are conceptually equivalent (i.e. aligned).

Table 3 - Experiments in a 2^3 Full-Factorial Design

	FACTORS				INTERACTION	
RUN	A	B	C	MAIN EFFECT	TWO-FACTOR	THREE FACTOR
1	1	1	1			✓
2	1	1	0		✓	
3	1	0	1		✓	
4	1	0	0	✓		
5	0	1	1		✓	
6	0	1	0	✓		
7	0	0	1	✓		
8	0	0	0			✓

Table 4 - Suggested Experiments

	ASU DESIGN EXPERIMENTS		TAMU LAB EXPERIMENTS	
COMPONENT	MAIN FACTOR	INTER-ACTIONS	MAIN FACTOR	INTER-ACTIONS
PROVOCATIVE STIMULI	2 runs		2 runs	
SUSPEND JUDGMENT	2 runs		2 runs	
FLEXIBLE REPRESENTATION	2 runs		2 runs	
FRAME OF REFERENCE SHIFTING	2^3 Factorial Design: 8 Runs		2^3 Factorial Design: 8 Runs	
INCUBATION*				
EXAMPLE EXPOSURE				

*Main factor results for incubation are presented as an example in this paper.

5. EXPERIMENT EXAMPLE

This section presents a simplified demonstration of how the alignment works; this experimental example tests only one component: Incubation. It is not possible to describe all experiments that were conducted. Lab Experiments and Design Experiments were designed for a predefined set of components (see Table 4). Some of these components were tested individually for main effects and others in combinations for main effects and interactions. The experimental example presented here focuses on the individual test of one component: Incubation (i.e. a simple comparative experiment: incubation vs. no-incubation). Details of the DOE at the Lab and Design Experiment levels are presented and the results analyzed and compared.

5.1. Hypothesis

Many IG methods provide interruptions in work on a given problem, allowing incubation time. Theoretically, incubation allows one's mental set (i.e. the arrangement of processes and definitions in one’s mind) to change, so that when work resumes following incubation, new aspects of one's task may become more apparent. Thus, incubation was predicted to enhance measures of divergent thinking in both the Engineering Design and Laboratory contexts.

5.2. Experimental Method

5.2.1 Participants

Lab Experiment. A total of 177 undergraduate student volunteers participated in this Lab Experiment. From these, 82 participated in the incubation run and 95 participated in the no-incubation run. Participants could choose from a variety of experiments to take part in, or they could choose to write a paper in order to fulfill a research requirement for their introductory psychology course.

Design Experiment. Approximately 60 undergraduate students participated in the various Design Experiment runs, from these, 22 participated in the incubation simple comparative experiment presented here. The subjects are mechanical engineering undergraduate students with basic engineering design knowledge.

In both, lab and design experiment cases, the participants are assumed to have similar level of design expertise. This is appropriate for the experiments since the focus is on the component (i.e. incubation), not on the expertise level.

5.2.2 Materials

Lab Experiment. The task included written instructions, shown in Figure 3, and blank paper provided in a packet for each participant.

Design Experiment. Materials used in the Design Experiment included the written statement for the problem, shown in Figure 4, and blank paper to record the ideas. These materials were provided to the participants as explained in the procedure.

PROBLEM STATEMENT

Imagine a planet (not Earth) on which intelligent life has evolved, including a species that used many tools. Create novel tools for an alien race on this imaginary planet; do not create motorized or electronic devices.

Draw, label, and describe in writing as many tools as you can. Your drawings should include front and side views.

Figure 3 - Problem Statement Used for the Lab Experiment

PROBLEM STATEMENT

Design a device to transport a ping-pong ball. The device should be powered by only a spring. The objective is to travel the farthest horizontal distance, measured perpendicular to the starting line.

OPERATION

You are not allowed to have any contact with the device when in operation. Time is not a factor; only the distance will be measured after the device comes to stop.

REGULATIONS

- You can only use the materials listed below. No other material may be used or substituted outside what has been specified.
- You can cut and deform any of these materials in any way you like.
- You can use adhesives, staples, scotch tape, and solder to make the joints of the structure.

MATERIALS

- Spring
- PVC tubing and pipe, rigid and flexible.
- Copper tubing
- Steel wire
- Wood, plywood and balsa.
- Cardboard
- Styrofoam
- Aluminum sheet
- Bolts, nuts, washers
- Nails

Figure 4 - Problem Statement Used for the Design Experiment

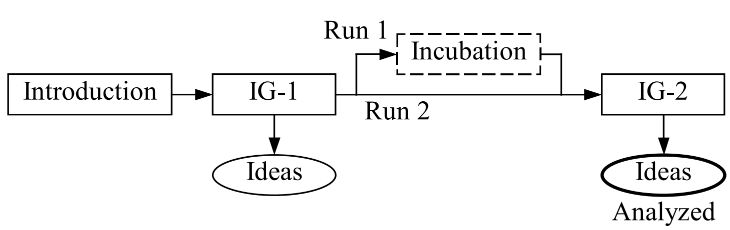


Figure 5 – Lab and Design Experiment Procedure – Incubation/No-Incubation

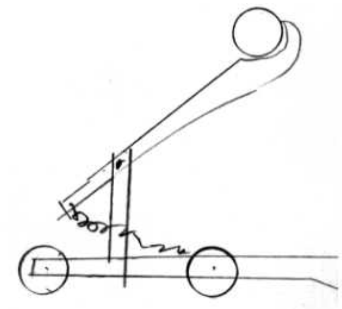
5.2.3 Procedure

The procedure for both the Lab and Design Experiment was similar; this is shown in Figure 5. Of particular interest were the ideas generated during the second IG session (IG-2), when the effect of incubation could be measured. This set of ideas were analyzed and compared between runs (1 and 2 for incubation or no-incubation), and between levels (Lab and Design Experiments).

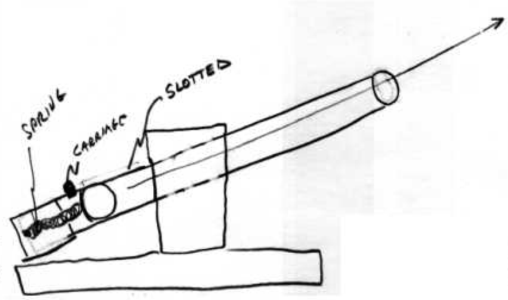
Lab Experiment. Participants were run in small groups (5 – 12 at a time) in a laboratory in the psychology building. An experimenter read the instructions (also provided to participants in writing) aloud to the group. The instructions explained that participants were to create novel tools for an alien race on some unknown planet. The instructions specified that it was an intelligent race that used many tools, but not to create motorized or electronic devices. Participants were to draw, label, and describe in writing as many tools as possible in the time given. Drawings were to include both a front view and a side view. The experiment was divided into two 20 minute sessions (IG-1 and IG-2). For half the groups there was a 10 minute incubation period in which participants completed mazes between the sessions, for the other half session two immediately followed session one.

Design Experiment. Two groups (i.e. the incubation group and the control group) of approximately 22 students each participated in the Design Experiment. For the incubation group, the participants were distributed in a classroom arranged for individual work. First, the material was distributed (i.e. pencils, problem statement and blank sheets), an introduction to the exercise was given, the problem statement read and idea recording guidelines explained. Second, participants were given 20 minutes to generate as many ideas as they could (IG-1). The sketches were collected and the participants excused; after three days of incubation, students returned to continue for a second 20-minute IG session (IG-2). The only difference between the incubation group and the control group is that the control group did not have an incubation period; they completed both IG-1 and IG-2 sessions one after the other.

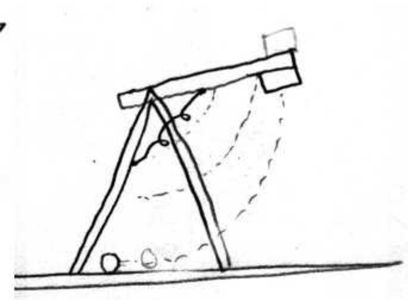
LOW NOVELTY SET



Catapult

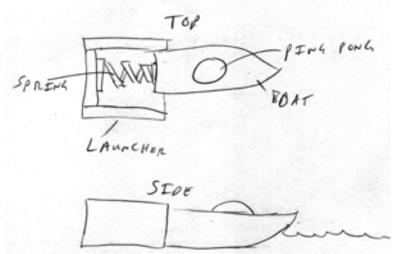


Cannon

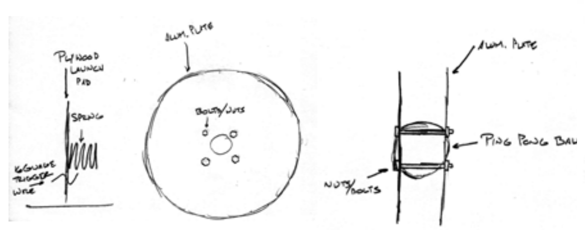


Hammer

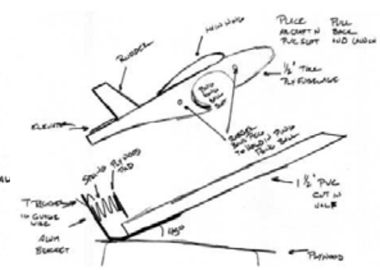
HIGH NOVELTY SET



Boat



Wheels Attached



Airplane

Figure 6 - Sample Sketches Produced in the Design Experiment

5.3. Data Collected

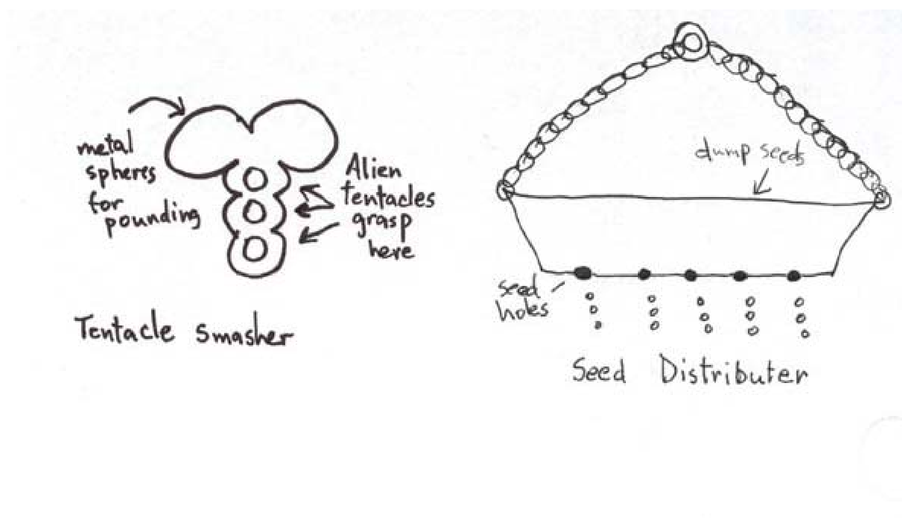
Figure 6 shows some of the sketches produced by the participants of the Design Experiment. These sample sketches are grouped into low and high novelty sets.

The measurement method followed is described at length in Shah et al. (2000; 2003) and is summarized here. Every idea was first characterized (i.e. solution method for each attribute is described); four attributes were identified from the Design Experiment problem statement (see Figure 4): Propulsion (i.e. impulse mechanism), Medium (e.g. fly, roll, float), Motion (e.g. sliding, rolling), and Number of parts. For Novelty scoring, the instances of each solution method were counted. The more a particular solution method was used the lower the novelty score assigned. For example, the catapult, cannon and hammer, shown in Figure 6, were more common than the boat, wheel attachment method, and airplane. Each idea's novelty score is computed by multiplying the novelty scores of each attribute by it's corresponding weight (e.g. Propulsion = 0.35, Medium = 0.35, Motion = 0.20, and Number of Parts = 0.10). For Variety scoring, the ideas are organized in a genealogy-like tree. Instead of using the four attributes, it was decided to use only the overall function: Ball Throwing. At the highest level ideas are branched according to the physical principle used. Subsequent levels branch the ideas according to working principle, embodiment, and detail differences. The nodes of this tree carry the number of ideas for that category and level. Upper levels have higher variety scores than lower levels. For example, the catapult and hammer shown in Figure 6 use the same working principle (i.e. linear spring potential energy with a lever mechanism) but have different embodiments. The variety score, which applies to the entire group of ideas, is calculated by multiplying each level's score by the number of corresponding branches. For Quality scoring, each idea was assessed with respect to four characteristics: Distance (i.e. estimated achievable distance), Operation (i.e. violation of operation rules defined in the problem statement), Manufacturing (i.e. how difficult it is to construct), and

Materials (i.e. comply to the given material list). Because the early state of the concept sketches, judges were employed to score the four characteristics. Each idea's quality score was computed by multiplying the judges' average score for each characteristic by the corresponding weight (e.g. Distance = 0.35, Operation = 0.15, Manufacturing = 0.20, Materials = 0.30). For Quantity scoring, the average number of ideas produced by each individual was calculated.

The sketches shown in Figure 7 are example responses from the tool-generation task used in the Lab Experiment. The Low Novelty Set are tools from common categories of generated tools (i.e., "hand" tools and farm tools), resemble existing tools (i.e., hammer and seed spreader), and use commonly used mechanical principles, so are rated low in novelty. The High Novelty Set, involve less commonly given categories of tools, have no direct existing counterparts, and utilize less commonly used principles. Quality scores were assigned by independent judges who were instructed to use the same standards and the same scale to assess novelty. Inter-judge reliability scores were high, indicating that the independent judges usually gave the same quality score. Novelty was assessed by constructing a master list of all tools generated by all subjects, and then tabulating the frequency of each idea. An idea's novelty score was the frequency divided by the total number of subjects.

LOW NOVELTY SET



HIGH NOVELTY SET

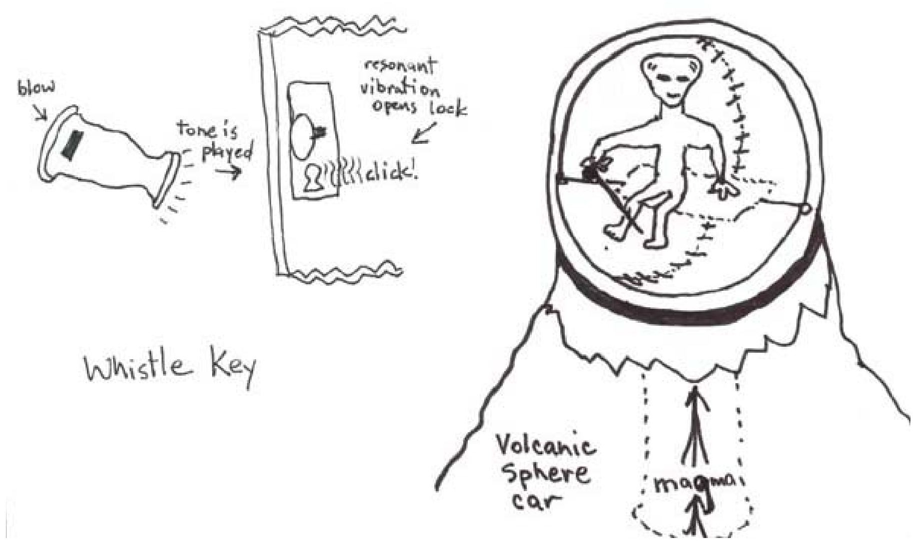


Figure 7 - Sample Sketches Produced in the Lab Experiment

5.4. Results Analysis

The ideas generated in the Lab and Design Experiments were scored for quantity, variety, quality, and novelty, and are shown in Tables 5 and 6 as a function of the experimental condition, either control or incubation. Scores range from 0 to 5 and 0 to 10 respectively for Lab and Design Experiments; higher scores mean superior quantity, quality, novelty, or variety. Scores can only be numerically compared for the same metric and experiment level. Scores for the same metric from two experiment levels can be compared through correlation. Correlation determines whether two data ranges move together.

Scores from different metrics cannot be compared numerically (e.g. quality vs. novelty).

Table 5 - Laboratory Experiment: Mean Divergent Thinking Scores as a Function of Incubation vs. Continuous Work Conditions

CONDITION	QUANTITY	VARIETY	QUALITY	NOVELTY
Control Group	1.51	1.20	1.47	0.00282
Incubation Group	1.93	1.60	2.37	0.00373

Table 6 - Design Experiment: Mean Ideation Effectiveness Scores as a Function of Incubation vs. Continuous Work Conditions

CONDITION	QUANTITY	VARIETY	QUALITY	NOVELTY
Control Group	4.86	2.81	6.15	4.71
Incubation Group	5.11	6.24	7.31	6.76

Table 7 – Two-Sample t-Test

	METRIC	t_0	*P*-value
LAB	QUANTITY	1.31	0.19
	QUALITY	1.85	0.07
	NOVELTY	3.16	0.002
DESIGN	QUANTITY	1.10	0.14
	QUALITY	5.86	0
	NOVELTY	12.38	0

Table 8 - Correlation Between Lab and Design Experiments

METRIC	CONDITION	LEVEL	
		LAB	DESIGN
QUANTITY	Control	1.51	4.86
	Incubation	1.93	5.11
	Correlation	1.00	
VARIETY	Control	1.20	2.81
	Incubation	1.60	6.24
	Correlation	1.00	
QUALITY	Control	1.47	6.15
	Incubation	2.37	7.31
	Correlation	1.00	
NOVELTY	Control	.00282	4.71
	Incubation	.00373	6.76
	Correlation	1.00	

5.4.1 Level of Confidence

Data in Table 5 and 6 show that, for every score, the incubation group had higher scores than the control group. Hence, the null hypothesis (i.e. Incubation has no effect) is rejected; the alternative hypothesis is true: Incubation enhances measures of divergent thinking in both the Engineering Design and Laboratory contexts. This is true for every metric considered and for both experiment levels. The significance level α (i.e. probability that the null hypothesis is rejected when it is true) can be calculated with the two-sample *t*-test. A normal distribution is assumed; this is acceptable since the experiment is assumed randomized. Table 7 summarizes these results. Based on the test statistic t_0, the *P*-value can be calculated. The *P*-value is the smallest level of α at which the data are significant (Montgomery, 2001). Variety is not included in this test since this metric applies only to the entire IG session.

For every metric except quantity, the P-values were less than 0.07, where 0.05 is generally considered low. For quantity, the highest P-value was 0.19, which means an 81% of confidence on the results. In each of the metrics used for Lab and Design Experiments, the incubation group always scored higher. This means that there was a benefit from incubation, compared to continuous work.

5.4.2 Correlation

Table 8 compares Lab and Design Experiments scores and shows the statistical correlation of both levels. It can be seen that all correlations are positive (i.e. equal to one) for all four metrics. This means that the results change simultaneously for both, the Lab and Design Experiments.

6. DISCUSSION

An issue might be raised regarding the subjectivity of the measuring method. The measuring method used here attempts to reduce subjectivity by defining and evaluating functions, attributes, or features, and by employing judges that must follow predefined guidelines. It has been found that judges or raters evaluating the same idea group and using the same method produce similar scores. This has been observed at both, Design and Lab Experiment levels in the incubation example presented here. The results presented here show that based on the correlation at both the Lab and Design Experiments; the selected component has a positive impact (i.e. promotes) on Design Ideation. This statement can be made with a reasonable level of statistical confidence. It must be clarified that the results presented here are for a single component (i.e. a single comparative experiment), and that ongoing work includes various components and their interactions and hence, will produce results that are far more complex. These results will be abstracted in a Design Ideation model and will be reported in the future. A second clarification is that the Design Experiments are a simulation from real world design, and as any prediction, its validity depends on the simulation model. Various issues exist in this area, such as design problem characterization, designer profiling, control of variables, and outcome metrics. Although these are current issues, the current paper is based on several years of experience in Design Research (Shah, 1998; Shah et al., 2000; Shah et al., 2001; Shah et al., 2003).

7. CONCLUSIONS

No model exists to date that comprehensively explains design ideation. The alignment approach described here provides the necessary framework for the creation of such a model. Two key concepts were identified for the alignment of Lab and Design Experiments: components and metrics. The alignment of components ensured that Lab and Design Experiments test the same concept (e.g. incubation, judgment, provocative stimuli, etc.). The alignment of metrics allowed the comparison of outcomes, and hence, the effectiveness of each component or combination (i.e. interaction) of components could be measured and compared at both levels. The incubation experiment presented in this paper exemplified the first steps of the alignment approach (i.e. individual modeling of ideation components). It can be concluded, based on the results from both Lab and Design Experiments, that the incubation component increases the effectiveness of ideas generated; further, these results correlate at both levels and show a satisfactory confidence level. What these result means in the overall alignment context is that the incubation's positive impact on Design Ideation is substantiated by concrete engineering evidence (from the Design Experiment results) and has a theoretical basis (from the Lab Experiment results). Finally, various issues were identified throughout this paper and are considered work in progress, but it is believed that what has been presented here is a right step towards a model of Design Ideation.

ACKNOWLEDGMENTS

The authors would like to acknowledge the support received through NSF Grant DMI-0115447 and Ford Research Labs. The views expressed are those of the authors and not endorsed by NSF. In addition, the authors thank the Chinese Scholarship Council for sponsoring Dr. Muqi Wulan as a visiting professor at the Design Automation Lab in ASU. Last but not least, thanks to the experiments' participants at ASU and TAMU.

REFERENCES

Altshuller, G., 1984, Creativity as an Exact Science, Gordon and Breach, New York

Christiaans H., and Dorst K., 1991, "An Empirical Study into Design Thinking", in Research in Design Thinking, Delft Press.

Christiaans H., and Venselaar K., 1991, "Practical Implications of a Knowledge Based Design Approach", in Research in Design Thinking, Delft Press.

Dean, A., and Voss, D., 1999, Design and Analysis of Experiments, Ed. 1, Springer Verlag, New York.

DMI-0115447, 2001, Development and Validation of Design Ideation Models for Engineering Design, Project Proposal Funded by NSF.

Ericsson, K., and Simon, H., 1984, "Protocol Analysis - Verbal Reports as Data", MIT Press.

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

Ideation, 2003, http://www.ideationtriz.com/

Koestler, A., 1964, "The art of Creation", Hutchinson and Co., London.

Kulkarni, S. V., 1998, Idea Generation Techniques – A Critical Survey, Tech. Report ASU/DAL/IG/98-2, Arizona State University.

Kulkarni, S. V., 2000, "A Framework for the Experimental Evaluation of Idea Generation Techniques", MS Thesis, Arizona State University, Tempe, AZ.

Kulkarni, S. V., Shah, J. J., 1999, Survey for Evidence of Components of Creativity, Technical Report ASU/DAL/IG/99-7, Design Automation Lab, Arizona State University.

Langley, P., and Jones, R., 1988, "Computational Model of Scientific Insight," in R. J. Sternberg, ed., The Nature of Creativity – Contemporary Psychological Perspectives, Cambridge University Press, NY.

Marples, D., 1960, "The Decisions of Engineering Design", Inst. of Eng. Designs, London.

Martindale, C., 1995, "Creativity and Connectionism," in S. M. Smith et al., eds., The Creative Cognition Approach, MIT Press, Cambridge, MA.

Montgomery, D. C., 2001, Design and Analysis of Experiments, Ed. 5, John Wiley and Sons, New York.

Pahl, G., and Beitz, W., 1996, Engineering Design – A Systematic Approach, Ed.2, Springer, London.

Schön, D., 1991,"Teaching and learning as a design transaction", in Research in Design Thinking, Delft Press.

Shah, J. J., Kulkarni, S. V., Vargas-Hernandez, N., 2000, "Guidelines for Experimental Evaluation of Idea Generation Methods in Conceptual Design", *Journal of Mechanical Design*, vol. 122, no. 4, pp. 377-384.

Shah, J. J., Smith, S. M., Vargas-Hernandez, N., 2003, "Metrics for Measuring Ideation Effectiveness", *Design Studies*, vol. 24, no. 2, pp. 111-134.

Shah, J. J., Vargas-Hernandez, N., Summers, J. D., Kulkarni, S., 2001, Collaborative Sketching (C-Sketch) – An Idea Generation Technique for Engineering Design, *Journal of Creative Behavior*, vol. 35, no. 3, pp. 1-31.

Shah, J., 1998, "Experimental Investigation of Collaborative Techniques for Progressive Idea Generation," Proceedings, ASME DTM Conference, Atlanta, GA.

Sickafus, N., 1998, "Unified Structured Inventive Thinking : How to Invent", NTELLEK.

Simonton, D. K., 1988, "Scientific Genius: a Psychology of Science", Cambridge University Press.

Smith, S. M., 1995, "Creative Cognition: Demystifying Creativity," in C. N. Hedley et al., eds., Thinking and Literacy – The Mind at Work, Lawrence Erlbaum Associates, Hillsdale, NJ.

Smith, S. M., Ward, T. B., and Finke, R. A., 1995, The Creative Cognition Approach, MIT Press. Cambridge, MA.

TRIZ Consulting, 2003, http://www.trizconsulting.com/

Ullman, D., Dietterich, T., and Stauffer L., 1988, "A Model of the Mechanical Design Process Based on Empirical Data", *A1 EDAM* V2 (1), 1988.

Ullman, D., Wood, S., and Craig, D., 1989, "The Importance of Drawing in Mechanical Design Process" NSF DTM conference.

VanGundy, A. B., 1988, Techniques of Structured Problem Solving, Ed.2, Van Nostrand Reinhold Company, NY.

Waldron M., and Brook R., 1994, "Analyzing Inter and Intra Group Information Exchanges in Conceptual Collaborative Design", ASME DTM Conference.

Wallas, G., 1926, "The Art of Thought", Harcourt, New York.

Ward, A., and Sobek D., 1996, "Principles of Toyota's Set Based Concurrent Engineering Process", ASME DTM Conference.

Ward, T. B., Smith, S. M. and 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. American Psychological Assoc. Books. Washington, DC:

