An Experimental Study of Group Idea Generation Techniques: Understanding the Roles of Idea Representation and Viewing Methods

J. S. Linsey[1]
Department of Mechanical Engineering,
Texas A&M University,
3123 TAMU,
College Station, TX 77843
e-mail: jlinsey@tamu.edu

E. F. Clauss
e-mail: emily.clauss@gmail.com

T. Kurtoglu
e-mail: tolga@mail.utexas.edu

J. T. Murphy
e-mail: jmurphy101@hotmail.com

K. L. Wood
e-mail: wood@mail.utexas.edu

Department of Mechanical Engineering,
The University of Texas at Austin,
1 University Station,
M/C C2200,
Austin, TX 78712

A. B. Markman
Department of Psychology,
The University of Texas at Austin,
1 University Station,
M/C A8000,
Austin, TX 78712
e-mail: markman@psy.utexas.edu

Advances in innovation processes are critically important as economic and business landscapes evolve. There are many concept generation techniques that can assist a designer in the initial phases of design. Unfortunately, few studies have examined these techniques that can provide evidence to suggest which techniques should be preferred or how to implement them in an optimal way. This study systematically investigates the underlying factors of four common and well-documented techniques: brainsketching, gallery, 6-3-5, and C-sketch. These techniques are resolved into their key parameters, and a rigorous factorial experiment is performed to understand how the key parameters affect the outcomes of the techniques. The factors chosen for this study with undergraduate mechanical engineers include how concepts are displayed to participants (all are viewed at once or subsets are exchanged between participants, i.e., "rotational viewing") and the mode used to communicate ideas (written words only, sketches only, or a combination of written words and sketches). Four metrics are used to evaluate the data: quantity, quality, novelty, and variety. The data suggest that rotational viewing of sets of concepts described using sketches combined with words produces more ideas than having all concepts displayed in a "gallery view" form, but a gallery view results in more high quality concepts. These results suggest that a hybrid of methods should be used to maximize the quality and number of ideas. The study also shows that individuals gain a significant number of ideas from their teammates. Ideas, when shared, can foster new idea tracks, more complete layouts, and a diverse synthesis. Finally, as teams develop more concepts, the quality of the concepts improves. This result is a consequence of the team-sharing environment and, in conjunction with the quantity of ideas, validates the effectiveness of group idea generation. This finding suggests a way to go beyond the observation that some forms of brainstorming can actually hurt productivity.
[DOI: 10.1115/1.4003498]

Keywords: idea generation, empirical study, brainsketching, 6-3-5, C-sketch, gallery method, ideation, innovation

1 Introduction

The ability to invent, create, and innovate is at the very core of engineering and product development. The initial design phases, including conceptual design, have the most significant impact on product cost [1]. Numerous idea generation techniques are available to assist the engineer in this process. Over 100 formal idea generation techniques have been developed in areas such as psychology, business, and engineering [2–4]. Techniques range from the well-known brainstorming method developed by Osborn [5] to methods specific to engineering, such as the theory of inventive problem solving (TIPS) [6]. Some of these techniques are meant to be implemented in a group setting, and others are intended for solitary work. Unfortunately, little empirical data exist to guide the use of methods for engineering design. This paper focuses on group idea generation techniques and provides guidance on their use through data provided by a controlled experiment with

engineering design teams.

One of the first studies using Osborn's brainstorming method in engineering design included engineering professionals working on a realistic engineering problem and showed that groups using brainstorming produced fewer ideas than the combined efforts of an equivalent number of individuals working alone [7]. This result, called *productivity loss*, is consistent with the vast majority of studies on variations in Osborn's brainstorming [8].

One might think that groups would produce more ideas. Associative memory models from cognitive psychology [9–11] suggest both why there is productivity loss and ways around it. In these models, memory is treated as a network of concepts. Closely related concepts are connected to each other more strongly than less related concepts. When a concept is brought to mind, it activates (and hence makes more accessible) other concepts that are connected to it. In group contexts, the first concept presented by a group activates a common set of concepts in all group members and thus tends to lead groups to converge quickly on a small set of concepts.

However, if individuals are first able to generate ideas individually and are then exposed to the ideas of other group members, these additional ideas serve as additional ways to access associative memory. So, after group members are allowed to generate

---

Fig. 1 Illustration of gallery method

Fig. 2 Illustration of 6-3-5 and C-sketch. Six people silently describe three ideas on a sheet of paper and then exchange papers.
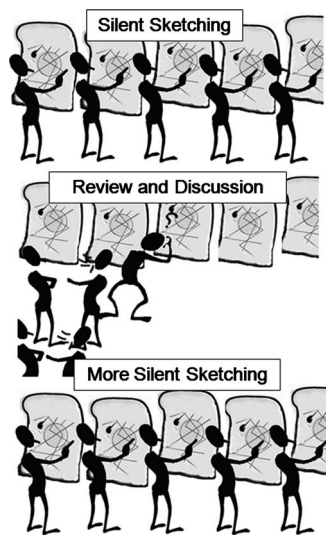
their own ideas, ideas provided by other group members can activate and facilitate the retrieval of additional, and sometimes disparate, ideas. Recent studies on engineering design support the potential of idea exchange for promoting new ideas [12–14]. Associative memory models are one reason why groups may produce fewer ideas, but there are many other possibilities such as fear of evaluation by teammates or production blocking; only one person can speak at a time, and this blocks other people [8].

For other group idea generation techniques, groups are more effective than the combined individual efforts [15,16]. One such example is brainwriting, which uses written communication rather than spoken communication. Recent studies have focused on the development and evaluation of more effective idea generation methods in engineering and design related fields, including industrial design and architecture [17–21]. These studies have used a mixture of sketches, verbal descriptions of ideas, and physical models in the idea generation process. The vast majority of idea generation techniques focus on the sentential expression of ideas despite the fact that designers rely heavily on sketches to express their ideas during the conceptual phase of design [1].

An exhaustive comparison of idea generation techniques is beyond the scope of this paper due to the vast number of techniques in the literature. Instead, we focus here on a generalization of idea generation techniques by systematically dissecting these methods into key factors or parameters. A design of experiment (DOE) approach is utilized to quickly identify the effects of these key factors. By understanding the factors that influence the success of idea generation methods, new and better methods may be developed. This type of approach was recommended by Shah et al. [19], and we believe that it facilitates a scientific understanding of the methods and their effectiveness.

Our study focuses on four group idea generation methods: brainsketching, C-sketch, 6-3-5, and the first phase of the gallery method (Figs. 1 and 2). These methods are gaining popularity and exposure in the engineering research community, in addition to industrial application [22]. They also form a diverse set of group idea generation techniques, which vary in how ideas are exchanged and in the types of representations used (written words, sketches, etc.). To understand the theoretical basis of these methods, we dissect them into two key factors: (1) how a group's ideas are displayed to other members ("rotational view" or all are posted in "gallery view") and (2) the form of communication between group members (written words only, sketches only, or a combination of words and sketches). All other method parameters are kept constant for all experimental conditions, including the
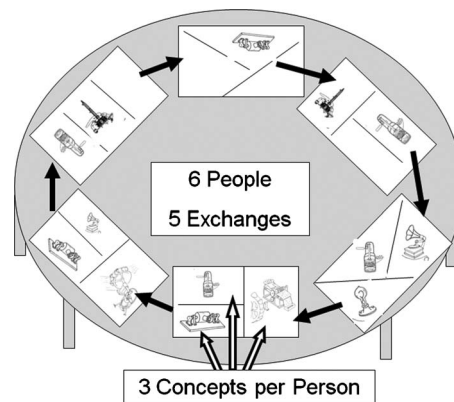
quantity of time for idea generation, whether the originator of ideas is identifiable to other participants, and the suspension of judgment.

2 Overview and Previous Work

To place our study in perspective, we begin by reviewing relevant methods of idea generation as well as the set of methods that form the basis of our systematic experiment.

2.1 Osborn's Brainstorming. The term "brainstorming" is frequently applied to idea generation techniques in general and not just to the technique developed and named by Osborn. Osborn's brainstorming begins with a facilitator explaining the problem. A group then verbally exchanges ideas following four basic rules: (1) Criticism is not allowed, (2) "wild ideas" are welcomed, (3) building off each others' ideas is encouraged, and (4) a large quantity of ideas is sought. Despite the face validity of these rules, much research demonstrates productivity loss in brainstorming compared with an equal number of individuals working alone [8]. Osborn's brainstorming has been studied extensively, and its shortcomings are well-known. Thus, we focus our research on techniques that incorporate other modes of communication, such as sketching, into the process. We now turn to a discussion of these methods

2.2 Brainsketching. In brainsketching, individuals begin by silently sketching their ideas on large sheets of paper including brief annotations. Group members exchange drawings, and silent sketching continues for another period of time [3]. This technique allows for a visual means of expression, and so it is well suited for product design. Van Der Lugt used teams of advanced product design students to compare brainstorming to a variant of brainsketching (which included the explanation of ideas between exchanges) [17]. The brainsketching variant led to more cases in which group members built on previously generated ideas than did brainstorming.

2.3 Gallery. In the gallery method, individuals begin by sketching their ideas silently on large sheets of paper. After a set amount of time, participants discuss their ideas and move about the room studying others' ideas. This review phase is followed by a second stage of silent sketching [3,23,24]. The review phase allows team members to clarify their ideas, and it provides social interaction (Fig. 1).

2.4 C-Sketch/6-3-5. For 6-3-5 [18,23,22] and C-sketch [18], six participants are seated around a table, and each silently describes three ideas on a large sheet of paper (Fig. 2). The ideas are then passed to another participant. This exchange goes on for five rounds. For the original 6-3-5 method, ideas are described using

Table 1 Summary of formal idea generation techniques

Formal technique	Form of external representation	How ideas are exchanged
Osborn's brainstorming	Spoken word	All are viewed at the same time
6-3-5	Written word	Rotational view
C-sketch	Sketches only	Rotational view
Gallery	Sketches and written word followed by spoken word	All are viewed at the same time
Brainsketching	Sketches and written word	Rotational view

only words. In contrast, the C-sketch method permits only sketches. One advantage of C-sketch over 6-3-5 is that sketches are typically ambiguous, and so one person may misinterpret aspects of someone else's sketch, which may lead to new ideas [23]. Other variations of 6-3-5 have also been proposed [3,22]. One variation permits annotated sketches [22]. In experimental comparisons with conditions different from those reported in this paper, C-sketch and gallery outperformed 6-3-5 (words only) for variety, quality, and novelty of ideas [23]. Novelty is how unique a particular idea is, and variety is how much of the design space is captured by a set of ideas. This previous study used groups of mechanical engineering undergraduates, mechanical engineering graduate students, and professional designers. Each group was evaluated on all three techniques, and a different design problem was solved for each of the techniques. This design eliminated individual differences as a noise variable but caused the technique results to be confounded with the design problem.

2.5 Potential Influence of the Components of Idea Generation Methods. Much of the variation in formal group idea generation methods is likely attributable to two main parameters: the representation used for communication and the method for exchanging ideas among participants. These dimensions, as summarized in Table 1, in many ways differentiate the methods and may be studied in the context of cognitive science findings. Relevant prior work is discussed here with respect to these dimensions.

In these idea generation techniques, group members communicate using some form of external representation. The choice of representation is important because it makes some information easier to convey than others. The representation is called "external" because it is outside of the head, as opposed to internal representations that are inside the head [25]. Embodied cognition theories suggest that external representations are particularly useful for difficult tasks such as design because they reduce the cognitive load by reducing the amount of information that an individual needs to represent internally [26].

Designers use many different kinds of external representations because the design process is cognitively taxing [1,27]. These representations include sketches, various forms of diagrams, and sentential annotations. Some diagrammatic representations are very specific in application, such as force flow diagrams for reducing system components, while others such as the technique known as the "house of quality" are more generally applicable [22,28]. Designers know that these varying forms of representation affect their thinking and therefore the final product [27].

In addition, the *theory of perceptual symbols* posits that internal mental representations are often based on perception (e.g., vision or audition) rather than being languagelike amodal representations whose connection to perceptual information is arbitrary [29]. Sketches, as external representations, are processed by people's visual systems, and thus they have mental representations that closely resemble internally generated perceptual representations. Thus, techniques that involve sketches may allow perceptual representations to have a greater influence on idea generation than

techniques that involve only verbal descriptions.

In contrast, many idea generation techniques currently available emphasize communication through sentential description because they were developed for less visually oriented applications such as business. It is fundamental that we understand the utility of these methods in engineering and the potential for combining sketching representations as an integral part of the methods. The importance of sketches in design is clear [1,27], highlighted by a recent issue of Design Studies that focused on sketching (Sept. 2006). Sketches support the transformation of ideas and help prevent premature fixation [27]. Designers also use their sketches to perceive and mentally simulate the function of their design, thereby supporting revision and refinement [30,31].

It is thus clear, based on cognitive science, that the communication of ideas among designers is likely to be significantly influenced by the modality in which the idea is presented (e.g., written, sketched, or combination). This statement is supported by Shah's research. Shah [18] contended that a potential benefit of limiting individuals in a group to sketches without verbal annotations is the increased potential for misinterpretations, which can lead to greater novelty and variety of solutions. A participant may misunderstand a teammate's sketch, leading the individual to an alternative idea not intended by the original sketch, thus producing a different solution.

The method of exchanging ideas (i.e., "gallery viewing" or "rotational viewing") is also likely to affect the communication of ideas among designers. The viewing conditions influence the amount of visual stimulus available, the evaluation of ideas, and how teams provide feedback to the individual members. Prior research shows that available visual stimuli affect the ideas generated [32]. Rotational viewing allows for only a subset of ideas to be viewed by each team member at a given time, whereas in gallery view a team's entire set of ideas is visible. In rotational viewing, there is no immediate feedback, whereas in gallery viewing, the individuals can see how their ideas are added to and changed at all stages of the idea generation process.

2.6 Potential for Verbally Based Techniques to Suppress Perceptual Memory. Many cognitive models of memory theorize that there are both perceptual and verbal representations [33–36]. There are many forms of evidence for this distinction. We briefly describe one. The verbalization of perceptual information can interfere with the retrieval of perceptual information from memory [37]. This effect is known as *verbal overshadowing*. Prior studies have evaluated an individual's ability to recall a number of different types of complex perceptual information, including memories of faces. In experiments examining this phenomenon, participants study a series of faces. Some give verbal descriptions of the faces, and others do not. Later, they are all given a face recognition task. In these studies, memory for faces is consistently better when participants did not describe the faces verbally than when they did. These data suggest that the verbal description interferes with the perceptual representation of the faces because the visual information that people use to recognize faces is not the same as the features that people describe when they give a verbal description of a face. Analogously, verbal idea generation techniques may suppress some of the perceptual information in memory, thus giving sketching based techniques a possible advantage.

2.7 Experimental Approach and Research Questions. Engineers seek a robust idea generation method for predictably producing a large quantity of high quality, novel product solutions. For the purposes of this study, an idea unit is defined as a solution to a single function, and a product solution is all ideas grouped together to solve the design problem. At present, there is no single approach to idea generation that meets all of these criteria, nor is it clear which idea-generation-method parameters are responsible for improving outcomes. Using a factorial design of experiments,

Table 2 Summary of experimental conditions

Factor 1: display	Factor 2: representation		
	Words only	Sketches only	Words and sketches
Gallery view	1	3	5
Rotational view	2	4	6

our study explores the influence of the representation used to communicate ideas and how ideas are displayed to individuals. We seek to answer the following research questions:

- Question 1: How do the techniques being tested influence the quantity, novelty, and variety of ideas? Which idea generation method produces the largest quantity, highest quality, largest variety, and greatest novelty?
- Question 2: Does the representation method of ideas interplay with the display method, or are they independent?
- Question 3: Are certain representations better for producing or improving the quality of solutions? Do certain representations cause bias toward certain types of ideas?

These three research questions are addressed systematically in the following sections. We discuss our experimental method, metrics for evaluation, data analysis approach, and a summary of the results. In addition, we discuss secondary issues such as the following: Does building off teammate's ideas improve the quality of the idea? These secondary issues are corollaries of the primary research questions and may be investigated directly from the data produced from the primary experiment.

3 Experimental Method

We conducted a factorial experiment in order to explore the effects of two key factors on the outcome of group idea generation. The first factor controls how participants view the ideas; either all ideas are posted via a gallery (on the wall), or sets of ideas are rotated between participants. The second factor controls how participants represent their ideas. Participants either use written words only, sketches only, or a combination of written words and sketches to communicate ideas to their teammates. A 2(display of ideas:gallery or rotational view)X3 (representation:words only, sketches only, or words combined with sketches) factorial experimental design is used (Table 2). No oral discussions are allowed during the session; all communication is written. This approach produces methods similar to 6-3-5 [24], C-sketch [18], brainsketching [3], or gallery method [24], as shown in Table 3.

The group factorial experiment was conducted over a 2 week period. Participants were asked to sign a confidentiality agreement to minimize the likelihood that other participants would hear about the problem and spoil the experiment. Additionally, a post-experiment survey asked participants if they had heard about the problem and if they had tried to generate ideas prior to the session.

3.1 Participants. The participants were students from a mechanical engineering senior design methods course at the Univer-

Table 3 Experimental conditions and similar formal method

Experimental condition	Similar formal idea generation method
1	Electronic gallery [38]
2	6-3-5
3	
4	C-sketch
5	Gallery
6	Brainsketching

sity of Texas at Austin. Participants' age ranged from 20 to 35, with 21–24 being typical. Participant teams were chosen because they provided a large sample of equally sized groups with experience working as a team. Participants at this level have exposure to significant mechanical engineering theory and have some experience in the design process through class work, internships, and work experience. More experienced designers have a greater database of knowledge to draw from, and therefore they may be more likely to generate a larger number of ideas than less-experienced engineers. However, choosing participants at the same level of education minimizes the variability in prior experience across groups. We asked a number of questions of participants to find out their previous internships and additional related experience. The amount of experience is expected to be independent of the parameters under study in this experiment. To verify this expectation, the participants were polled regarding their knowledge of other ideation techniques and, in particular, the techniques used as part of our study. No students in the sample size had previous knowledge of these techniques, and, thus, no bias was introduced based on the student backgrounds with ideation techniques.

The participants were told that they would receive extra credit for their participation based on the number, quality, novelty, and variety of solutions they develop. In the design methods course, students work throughout the semester in teams of four to six members. The course assigns teams based on a strategy for improved team dynamics based on Myers–Briggs personality types, 6-hats, and analytical/fabrication skills, in addition to participants' skills and experience level [39–43]. These team-formation strategies do not assess the level of creativity of the participants but instead try to spread the range of team skills and personality preferences across the class. Thus, the variance in the experimental results due to these factors should be reduced through the team-formation approach. This method of assignment is not expected to influence the results of the study, except for the reduction of variance and uncontrolled effects. For the factorial experiment, 14 of the possible 15 teams chose to take part in the experiment and participated with their assigned teams. Participants were required to sign up as a complete assigned team, and each team participated in the experiment only once.

As part of the design course, students were taught a series of idea generation techniques including brainstorming, TIPS, information gathering, patent searching, use of analogy, and a hybrid version of 6-3-5/C-sketch that emphasizes sketching with short annotations [22]. These methods were taught in a series of five 1 h lectures where mandatory class attendance was required on the part of the students. None of the students in the experiment missed more than one lecture. The experiment took place after these methods had been presented in class. Students have minimal practice with any one of these techniques, and so they should not be biased to prefer any one particular technique. An advantage of this prior exposure to ideation techniques though is that participants are aware of any technical jargon that may arise in describing the techniques used in the experiment to them.

3.2 Description of the Design Problem. This work draws from both mechanical engineering and cognitive psychology. The typical psychology experiment uses participants with no domain knowledge, and so the problems that are being solved in research on creativity do not require specific domain knowledge to solve. Because we use domain knowledgeable participants in this research, we can use real-world problems. Thus, it is crucial that we select a design problem appropriate for the study. Our goals for choosing a design problem are making it need-based, so that participants are motivated to solve it, making it real-world and currently relevant, and choosing a scope that is difficult, without obvious solutions, but not so complex that participants cannot find solutions for it within the time limits of the study.

Based on the above criteria, the problem given to groups is to design a device to quickly shell peanuts for use in places like Haiti

Device to shell peanuts
Problem Description
In places like Haiti and certain West African countries, peanuts are a significant crop. Most peanut farmers shell their peanuts by hand, an inefficient and labor-intensive process. The goal is to build a low-cost, easy to manufacture peanut shelling machine that will increase the productivity of the peanut farmers. The target throughput is approximately 50 kg (100 lbs) per hour.
Customer Needs:
• Must remove the shell with minimal damage to the peanuts
• Electrical outlets are not available as a power source
• A large amount of peanuts must be quickly shelled
Functions:
• Import energy to the system
• Break peanut shell
Separate peanut shell from nut

Fig. 3   Design problem description

Fig. 4   Options for building from others' ideas

and West African countries. This problem is based on a real-world problem posted on the website ThinkCycle [44]. Participants are told that no electrical energy sources are available. They are given a description of customer needs along with corresponding functions (see Fig. 3). The problem is read to the participants with no further clarification. This problem was chosen because it is a real-world problem that a mechanical engineer should be able to solve and has a diverse set of available solutions. We did not expect any of the participants to have extensive prior experience in solving this problem, yet shelling a peanut is a task all of the participants should have experienced.

3.3   Procedure for All Group Conditions. Teams were randomly assigned to conditions. For teams with six members, one person was randomly assigned to work alone (as a control), and their ideas were not included in the team totals. The four-member team also worked individually as a control. During one session, only four of the five members were present; thus, their results were not considered further. In total, 12 teams of five are included in the results. Sessions were scheduled at the team's convenience throughout the day and week. All sessions took place in the same room, a windowed conference room in the mechanical engineering building.

Participants were each given a unique set of five colored markers and were seated next to each other facing the same direction. The variety of colors makes it difficult for other participants to identify the originator of an idea while at the same time allowing identification by the experimenter. Previous work suggests that more ideas are generated when people believe that their responses will be anonymous, but ideas must be identifiable to the experimenter to reduce social loafing [45]. Participants were told they could use the various colors any way they desired, but three examples of how color could be useful were given to encourage the use of multiple colors. These examples include using color to show different components of a design and variations on an idea and to help explain ideas such as coloring water blue. The examples were intentionally selected to be unrelated to the design problem and included a sketch of a box with two different styles of holes and a facet. An additional effect of the markers is the equalization of drawing abilities. Sketches from participants with greater drawing ability look essentially the same as sketches from participants with less ability when using these markers. There is no drastic difference in sketch quality across the participants.

The same experimenter ran all experimental groups. The experimenter read a set of scripted instructions and posted the ideas on the wall or rotationally exchanged ideas between participants. The instructions included a description of the problem, the basic idea generation rules [5] of seeking a large quantity of ideas along with encouraging diversity (wild, eccentric, or nonstandard ideas), a reminder that criticism is not allowed, and a statement that the session is to test a new idea generation method. The experimenter told the participants how to represent their ideas (words only, sketches only, or a combination of words and sketches) and then described the viewing method (gallery view or rotational view).
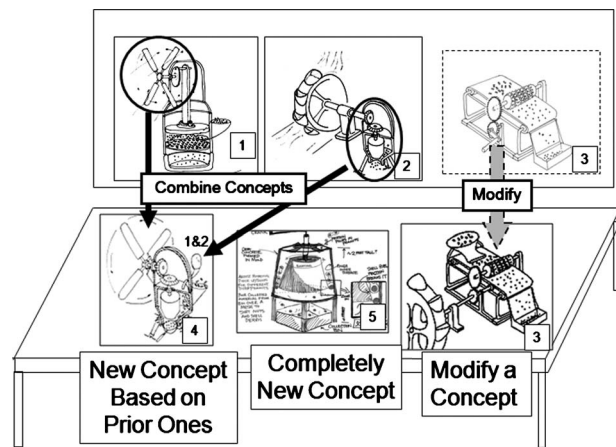
The session lasted approximately 50 min with 40 min for idea generation, followed by a post-session questionnaire.

3.4   Factor 1: Display of Ideas. One key factor in this study is whether ideas are displayed all at once or whether participants see only a subset at any given moment. In the gallery view condition, all ideas generated by the team are posted on the wall, so all participants can see all of the ideas at the same time. This approach results in a method similar to the gallery method or brainsketching [3,24]. In the rotational view condition, ideas are passed around the table, so that each participant sees only a subset of the ideas at any given moment. This condition is similar to 6-3-5 or C-sketch [18,24,22].

3.4.1   Gallery View Condition Similar to Brainsketching or Gallery Method. For the first 10 min period, each student is given a number of paper sheets and told to write down at least two ideas on separate sheets of paper. Sheets are collected as participants finish but are not displayed until the end of the period. The time period length is based on the available time and recommendations from the literature, which vary from 5 min to 15 min [3,19,46]. The ideal time period for the methods under evaluation is not explicitly known and is not one of the experimental parameters. At the end of the first period, all sheets are numbered and posted gallery style on the wall. In the four subsequent 7.5 min periods, ideas are posted as they occur, and participants are told to execute one of the following options (Fig. 4).

1. Add new ideas to one of the posted drawings. Participants can request a drawing by writing down its number on a small sheet of paper.
2. Make a separate drawing that is related to the ideas that are already posted, and write the number of the linked idea on the new sheet.
3. Start a completely new sheet after reviewing the posted ideas.

3.4.2   Rotational View Condition Similar to 6-3-5 or C-Sketch. For the first 10 min period, each participant is given a number of paper sheets and told to write down at least two ideas on separate sheets of paper similar to the gallery view condition. At the end of the period, the experimenter collects all sheets and systematically redistributes them such that each participant views each set of papers once. Participants cannot identify which one of their teammates had the sheets previously. In the four subsequent periods, lasting 7.5 min each, participants have the same options as in the gallery view condition: to add ideas to an existing sheet, to create a new product solution linked to another sheet, or to start a completely new product solution. The exception here is that participants focus on the specific set of papers given to them at a particular instance in time.

3.5 Factor 2: Representation. The second experimental factor prescribes how the participants communicate their ideas to other participants (words only, sketches only with no words, or a combination of words and sketches). At the end of the sessions and after completion of the surveys, participants in either of the sketches only conditions labeled their sketches with brief descriptions to facilitate evaluation. American mechanical engineers are typically not taught to draw free-hand, and therefore their sketches are usually difficult to interpret without annotations.

4 Metrics for Evaluation

It is crucial to have good metrics for evaluating the outcomes from idea generation techniques. Unfortunately, these techniques are not yet well-developed in general. Previous studies have used different methods for measuring outcomes, including quantity of ideas, number of good ideas, practicality, novelty, and variety [15,17,23,47,48]. Commonly used metrics are the quantity of non-redundant ideas and quality rating [8]. Shah et al. [19] developed a set of metrics specifically for the evaluation of engineering idea generation techniques, including quantity, quality, novelty, and variety.

Our study also measures the quantity, quality, novelty, and variety of product solutions and the quantity of ideas. However, we extended beyond the current measurement methods by developing algorithms for the measurements that increase reliability. For example, three existing solutions for shelling peanuts were sketched and added to the participants' results to benchmark and add additional validity to the metrics. Two solutions, which we believed were good solutions to the problem, were aimed at third world countries [49,50], and the third solution, which we felt was a poor solution to this problem, was aimed at a large-scale industrial application [51]. Sections 4.1–4.3 describe our extended measurement methods in more detail.

A product solution is defined as all ideas sketched on a single page, and a single idea unit is defined as a single solution to one of the device's functions. Metrics are measured at two different levels, the product solution and the individual idea. Quality, novelty, and variety are only measured at the level of a product solution because many very novel products are unique combinations of common components, and it is the product solution level that is of most interest.

4.1 Method for Measuring the Quantity of Product Solutions and Ideas Within the Product Solutions. Previous research suggests that the number of unique (or nonredundant) ideas is important for ensuring the successful development of a product [52]. A single product solution is defined as all the ideas contained on a single page unless participants made a clear indication that the product solution is continued onto another page. Figures 5 and 6 show examples of a single product solution for "words only" and "sketches only" conditions. Figure 6 also lists, in the form of a table, the ideas that are contained within the product solution sketches.

A single idea is more difficult to define. A critical element for measuring the number of ideas is a precise definition of what constitutes a single idea. Is a single idea an off-the-shelf component or piece-part, a single noun phrase, an item that meets any function, or something else? This question is particularly difficult to answer when the data are in the form of sketches because sketches frequently contain many vague details.

Building from the procedure developed by Shah et al. [19], a set of procedural rules is defined for what constitutes a single idea; see Linsey et al. [52] for more detailed examples. Our basic definition of an idea is something that solves one or more of the functions of the design as defined by the functional basis, a clearly defined and tested language for expressing design functions [53,54]. For example, in Fig. 6, if a second participant had proposed the idea of using a grate to separate the peanuts and the shells, it would have been counted as a different idea, whereas if

Fig. 5 An example of one product solution from the words only condition

they had proposed a flat plate with holes rather than the holes in the drum, this would have been counted as the same idea; see Linsey et al. [52] for more detailed examples.

However, we must define a second type of idea for cases in which participants reframe the problem more abstractly. This situation occurs more frequently when participants use only words for their descriptions. These solutions are clearly ideas, but they do not fit defined functions of the functional basis for the stated problem. For example, ideas in this category range from genetically engineered peanuts to training squirrels to shell the peanuts. A third refinement is made for function sharing ideas, that is, features that perform two or more functions. Function sharing ideas count as a single idea. This choice is made because it provides greater consistency between judges, leaving less room for interpretation of the intended function. Clearly, function sharing is good design practice, but even if it were possible to create a good definition of function sharing, the importance of innovations that serve multiple functions will be captured by our quality metrics, not our quantity metrics. Our quantity metric is biased toward a functional view, but this definition combined with the definition of "reframing" ideas covers virtually every solution encountered.
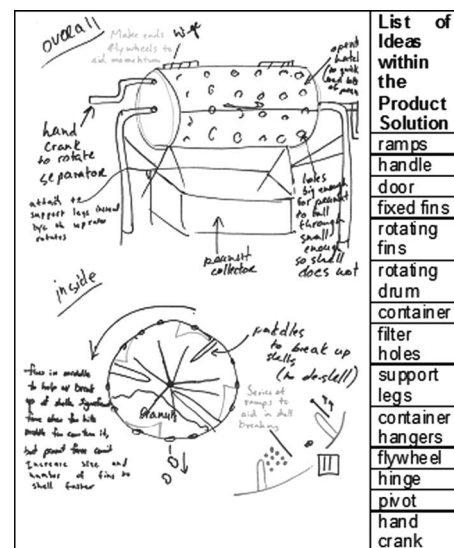

Fig. 6 Example of one sketched product solution and the list of ideas contained within the product solution

This method for measuring the quantity of ideas allows for a high degree of inter-rater agreement and a robust metric, as demonstrated by tests during the experiment.

Three judges independently counted the number of ideas based on the guidelines given above. Two judges were blind to the conditions of the experiment and the hypothesis, one of whom counted all of the data. The other two judges each counted a nonoverlapping subset. Pearson's correlation coefficients [55] for the two sets of judges were 0.99 and 0.95, demonstrating high reliability in the results. Because the counting rules are being applied consistently, the analysis of the quantity data is completed using only the comprehensive judge's results to minimize the variance due to using different judges.

Once ideas for each team are counted, the score for each individual is found by identifying the originator of each idea based on the marker color used to write or sketch it. Similarly, the time period the idea was conceived is determined. For gallery view conditions, the time each sketch is completed and the time of additions to it are recorded. The credit is split evenly between the participants when multiple participants think of the same idea during the same time period. This situation occurs frequently for the rotational viewing condition and occasionally in the gallery view condition when two participants are writing down the same idea at the same time.

4.2 Method for Measuring the Variety and Novelty of Product Solutions. To measure the variety and novelty of product solutions generated, two independent raters sort the sheets into groups or bins of similar product solutions, where a given rater chooses what constitutes "similar." One rater was blind to the conditions of the experiment and the hypotheses. The first evaluator formed 34 bins, and the second created 28 bins. The variety score for a team is measured by the percentage of total bins that the team's product solutions occupy. For example, if a team produces product solutions that are sorted into six bins by rater 1, that group would receive a variety rating of 6/34 or 17.6%. Pearson's correlation [55] between the raters was high, r=0.82. This value of correlation is in the acceptable range based on the literature [55].

The novelty score for each product solution is a function of the number of similar product solutions (i.e., number of product solutions in that particular bin) relative to the total number of product solutions. Specifically, novelty is equal to 1 − the frequency an idea occurs and has been used previously [56] (Eq. (1)). This metric fails if a team develops one very novel product solution and then creates numerous variations on it. None of the teams in this study demonstrated such a result. A high correlation between raters, again, was observed (r=0.80),

$$\text{novelty} \equiv 1 - \text{frequency of idea}$$

$$= 1 - \frac{\text{number of very similar concepts}}{\text{total number of concepts}} \quad (1)$$

4.3 Method for Measuring Product Solution Quality. Quality, as defined by Shah et al. [20], is a measure of a product solution's feasibility and how well it meets design specifications. In this paper, quality is measured on a three-point rating scale (Fig. 7) independently by two judges, one of whom was blind to the conditions of the experiment and the hypothesis. Each product solution generated by a team received a quality score. After initial data evaluation by both judges, Cohen's kappa [55] showed a fair level of inter-rater agreement (0.42). All differences were readily resolved through discussion. As expected, the benchmark solutions designed for third world countries scores a 2. and the industrial benchmark solution is rated at a 1 (Fig. 8).

It should be noted that a coarse (three-point) highly defined rating scale is used rather than an unanchored rating scale for our quality measurements. Our previous work suggests that raters have difficulty applying an unanchored scale, consistently leading
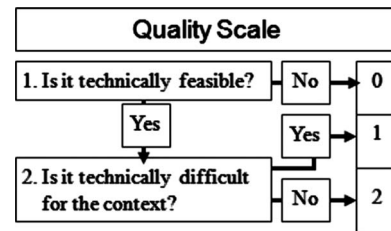

Fig. 7 Quality scale

to a low correlation between raters [57]. An unanchored rating scale, or rubric, has an expert evaluator rate a product solution on a spectrum, for example, 1–7 with "1" corresponding to lowest quality and "7" corresponding to highest quality, without specifically defining each point on the scale. Figure 7, alternatively, is a highly defined rating scale, or rubric, since each point on the scale has a specific definition.

5 Results and Discussion

Sections 5.1-5.6 present results that address the research questions posed previously and describe the significance and implications of these results according to each experimental metric. Table 4 highlights the aggregate results for the experimental conditions, and Fig. 9 presents sample product solutions that contain numerous ideas. As shown in Table 4, the participants generated many ideas, and on average these ideas were technically feasible. Participants obviously committed their time seriously to this real-world problem. In the post-experiment survey, one participant in experimental condition 4 did note that they had heard about the experimental problem ahead of time and thought about the problem. This team's data were reviewed, and no significant or noticeable bias existed in the team's or individual's results. Additional survey results may be found in Linsey et al. [58].

5.1 Quantity. We measured the number of ideas generated both by teams and by individuals. Recall that each participant used a unique set of markers so that we could identify who generated each idea. Figure 10 shows the mean number of ideas generated by the teams in each condition, and Figs. 11 and 12 illustrate the individual results. Since each team contained an equal number of individuals, the individual averages map very closely onto the team averages. Since the individual results contain a higher sample size, the statistical significance is higher for the individual results compared with the group.

An analysis of these data shows that the quantity of ideas increases by 50% due to the variation in experimental conditions (comparing the first bar in Fig. 10, words only with gallery view, to the last bar, words and sketches in rotational viewing). The number of ideas generated depends on both the representation implemented and the viewing conditions, as shown by the strong
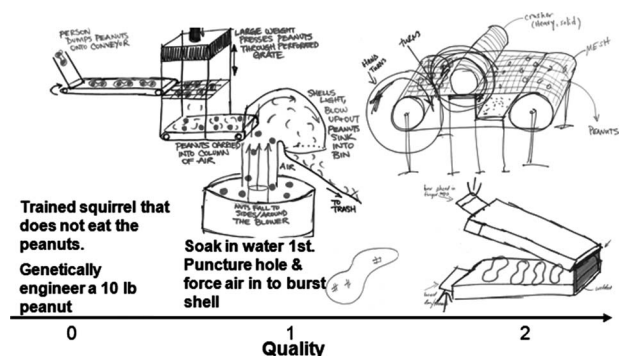

Fig. 8 Examples of product solutions at each quality level

Table 4 Averages for each metric by condition

Condition	Representation	Viewing approach	Quantity of ideas per person (SD)	Number of product solutions (SD)	Ave. quality (SD)	Number of high quality product solutions (SD)	Variety (SD)	Ave. novelty (SD)
1	Words only	Gallery	12.1 (5.5)	25.5 (10.6)	1.36 (0.82)	11 (7.1)	43% (0.20)	0.94 (0.03)
2	Words only	Rotational	17.5 (4)	18.5 (0.7)	1.19 (0.11)	6.5 (0.7)	38% (0.02)	0.93 (0.03)
3	Sketches only	Gallery	17.6 (5.4)	23.0 (0.0)	1.74 (0.07)	17.5 (2.1)	35% (0.03)	0.92 (0.006)
4	Sketches only	Rotational	14.4 (4.9)	19.0 (7.1)	1.86 (0.01)	16 (5.7)	33% (0.04)	0.92 (0.01)
5	Words and sketches	Gallery	16.7 (4.1)	22.0 (2.8)	1.60 (0.14)	14 (5.7)	32% (0.05)	0.92 (0.02)
6	Words and sketches	Rotational	19.3 (4.9)	15.0 (2.8)	1.42 (0.16)	6.5 (2.1)	33% (0.04)	0.94 (0.01)
Average			16.3	20.5	1.53	11.9	36%	0.93

interaction effects between the representation and the viewing conditions as highlighted by the nonparallel interaction lines in Fig. 11 and the analysis of variance (ANOVA) analysis (display: $F_{(1,54)}=1.65$, $p>0.2$; representation: $F_{(2,54)}=2.24$, $p>0.1$; and interaction: $F_{(2,54)}=4.12$, $p<0.03$, and $MS_{error}=23.32$). The information in brackets gives sufficient information to produce the ANOVA (Table 5). The mean square for each component is the F-value multiplied with the (MS_{error}). The sum of squares is the mean square multiplied with the degrees of freedom; see Ref. [55] for more details. Evaluating the ANOVA for the total group scores shows a similar pattern of results but with lower significance levels due to the smaller sample size (Fig. 10) (display: $F_{(1,6)}=1.21$; representation: $F_{(2,6)}=1.71$; and interaction: $F_{(2,6)}$

$=3.13$, $p=0.12$, and $MS_{error}=158.33$).

Figure 11 more clearly illustrates the interaction effect. In the words only and "words combined with sketches" conditions, the interaction follows the same pattern such that rotational viewing increases the number of ideas. For the sketches only condition, rotational viewing decreases the number of ideas. Three 2×2 ANOVAs are compared to further understand the source of the interactions and highlight hidden effects. ANOVAs for the words only and sketches only conditions (display: $F_{(1,36)}=0.48$, $p>0.5$; representation: $F_{(1,36)}=0.58$, $p>0.5$; and interaction: $F_{(1,36)}=0.7.46$, $p<0.01$, and $MS_{error}=24.80$) and the "words and sketches" and sketches only conditions (display: $F_{(1,36)}$
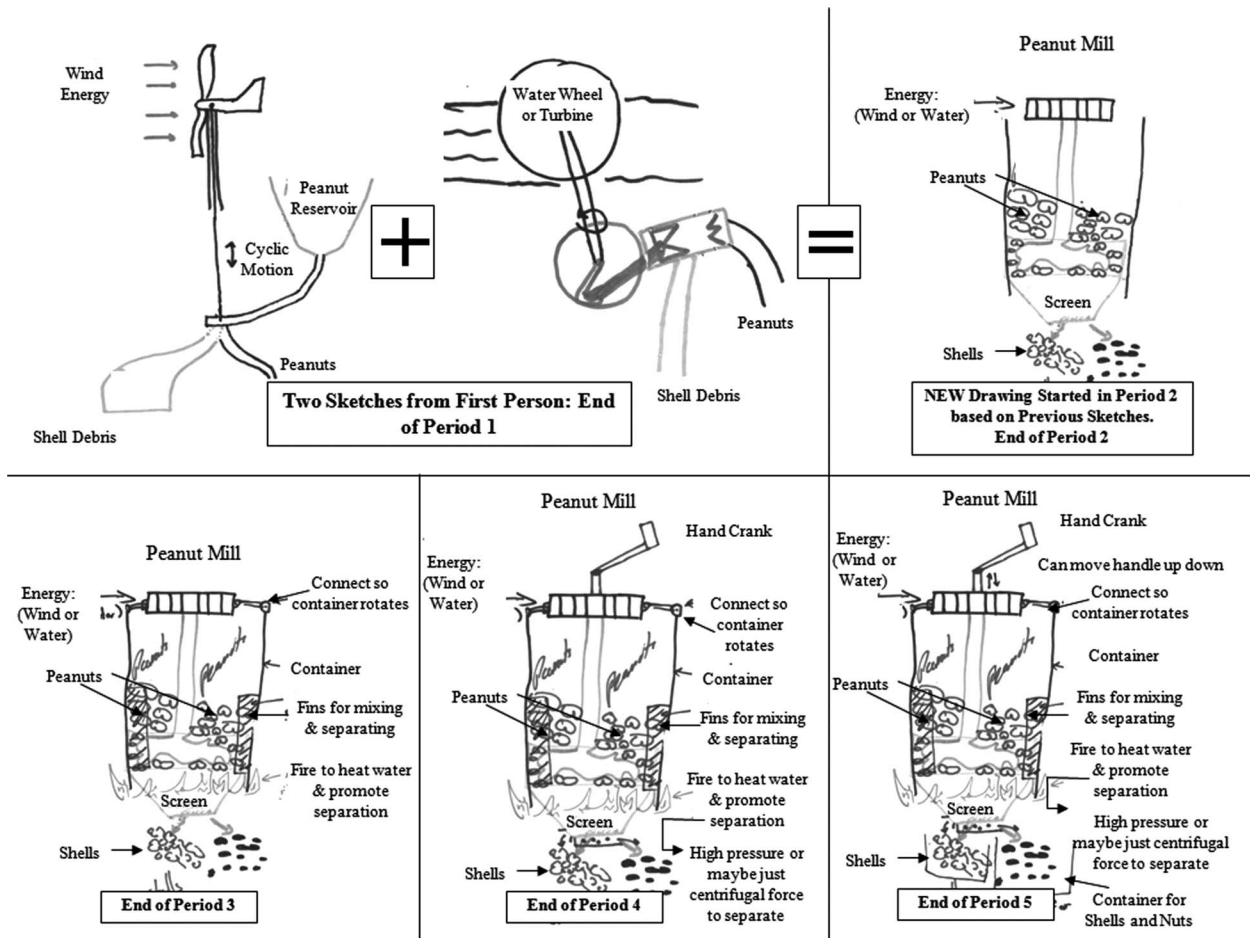

Fig. 9 Example of interesting results from a words and sketches combined with rotational viewing session. The quality of the product solution increases as ideas are added.
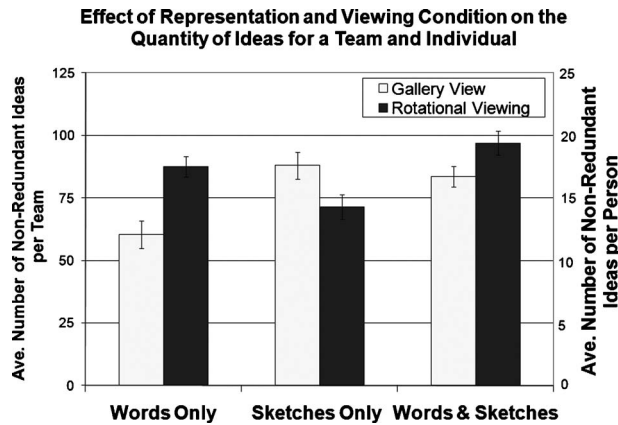
Fig. 10 Average results for each individual match team results since team sizes were equal. A significant interaction between viewing condition and representation exists. Error bars are ±1 standard deviation for the individual results.

$=0.038$, $p > 0.5$; representation: $F(1,36) = 1.71$, $p > 0.15$; and interaction: $F(2,6) = 3.59$, $p < 0.0007$, and $MS_{error} = 23.40$) show that an interaction effect is due only to the sketches only conditions. This result suggests that other factors may be influencing the sketches only condition. Participants' sketching ability is generally poor, and they may be frustrated by the difficulty in communication when limited to sketches.
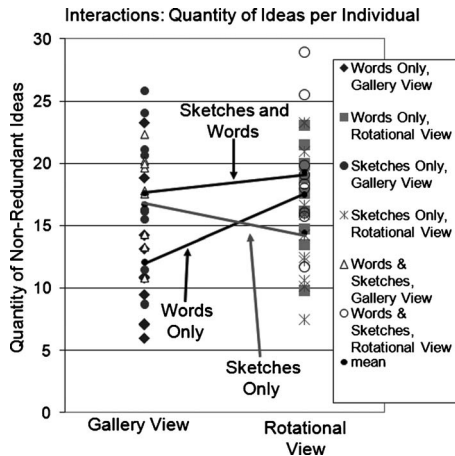

Fig. 11 There is an interaction between the representation and how the ideas are displayed
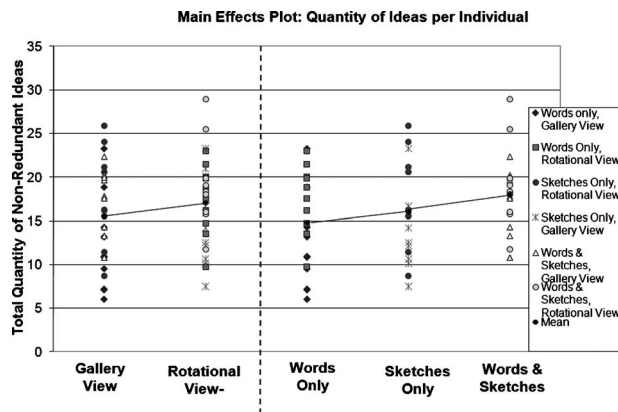

Fig. 12 Main effects for representation and how the ideas are displayed. Number of ideas for each individual is shown.

Table 5 Quantity 3×2 ANOVA results and quantity of ideas per individual

	Sum of squares	Degrees of freedom	Mean square	F
Display	38.4	1	38.40	1.65
Representation	104.5	2	52.27	2.24
Interaction	192.4	2	96.20	4.12[a]
Error	1259.4	54	23.32	
Total	1594.7	59		

[a]Statistically significant results at the 0.03 level.

The sketches only data follow a different pattern of results than the other two representation conditions. A comparison of the rotational view condition versus the gallery view condition shows statistical significance at the 0.01 level after removal of the sketches only effect (Table 6), (display: $F(1,36) = 7.35$, $p < 0.01$; representation: $F(1,36) = 4.70$, $p < 0.05$; interaction: $F(1,36) = 0.9$, $p > 0.5$, and $MS_{error} = 21.77$). This trend corresponds to an approximately 30% increase in the number of ideas for the five person team. A 30% increase in the number of ideas over a 40 min time period is important. This result is significant and intriguing but also needs to be validated with additional experiments over greater time periods. The use of sketches and words increases the total number of ideas by about 20% compared with only words for the five person team and is statistically significant.

These results highlight a few critical aspects of group idea generation. First, an effective group idea generation process can have a dramatic effect on the quantity of ideas generated. Second, the communication representation interacts with the method of idea exchange. This interaction is caused by the sketches only conditions. When the sketches only conditions are removed, there is no longer an interaction effect, only two main effects due to the representation and the viewing condition. This indicates that the interaction effect is due to the sketches only conditions behaving in a different pattern than the other two representations. From this study, it is unclear why the sketches only condition behaves in a distinctly different manner, and further research needs to be carried out. This result was unexpected and warrants further exploration particularly because many areas of design stress sketching and the prior experimental results [23] suggest that there are benefits of strictly sketch representations.

5.1.1 Quantity of Ideas Over Time. Figure 13 shows the quantity of ideas as they are incrementally developed during the idea generation process, and this pattern is typical of all of the teams. Groups develop more nonredundant ideas during the initial time period, but a virtually equal and substantial number of ideas continue to be generated throughout the session (Fig. 13). This result is significant and supports the concept that team members are piggy-backing on other members' ideas or leap-frogging as inspired by viewing others' ideas. About the same number of ideas

Table 6 Quantity 2×2 ANOVA results; words only and sketches and words conditions

	Sum of squares	Degrees of freedom	Mean square	F
Display	160	1	160	7.35[a]
Representation	102	1	102	4.70[b]
Interaction	19	1	20	0.90
Error	784	36	22	
Total	1066	39		

[a]Statistically significant results at the 0.01 level.
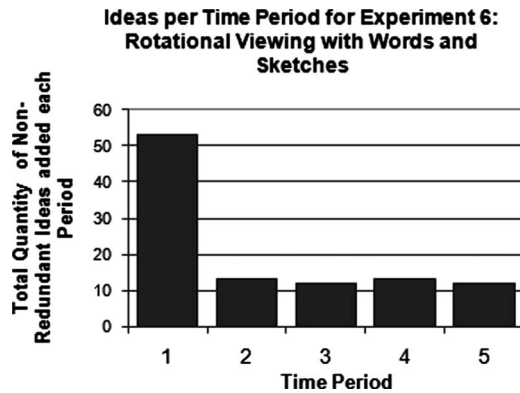[b]Statistically significant results at the 0.05 level.

Fig. 13 Quantity of nonredundant ideas added each time period. Results are from a single team in condition 6.

Fig. 15 Most groups evaluated only a small fraction of the total design space

were added in time periods 2–5, indicating that the teams did not exhaust their ideas. This finding suggests that future research needs to investigate longer time periods for idea generation.

As shown in Fig. 14, a large number of ideas are gained during the first time period when individuals work alone and also through collaboration with other team members. Building from others' ideas produces a nearly equal number of ideas as the individuals working alone. This result is a clear contribution and insight from this experiment and is consistent throughout the experimental conditions.

5.2 Novelty. The novelty of the results did not vary across the conditions (Table 4). Participants found a number of unusual solutions to the design problem. For one rater, 19 of the product solution bins contained only one or two product solutions, indicating that the product solutions are unique. All conditions produce essentially equal levels of novel product solutions, with ANOVA results showing no main effects for either representation or viewing conditions nor any interaction (display: $F(1,6)=0.37$, $p>0.5$; representation: $F(2,6)=0.27$, $p>0.5$; and interaction: $F(2,6)=0.71$, $p>0.5$, and $MS_{error}=0.00032$). The results indicate that this set of methods produces an equal level of novelty.

5.3 Variety. The variety for each team's results is not influenced by the experimental conditions, as shown in Table 4, but the teams explored only a segment of the total design space (Fig. 15). For this experiment, the total design space is defined as all solutions found by the participants. Most teams evaluated possible solutions in less than half of the design space. One team did explore significantly more of the design space showing that a greater breadth, i.e., greater variety, of solutions is possible in the time allowed. This result is of particular concern since the variety metric is a measure relative to all the solutions generated by the 12 groups of participants and not all theoretically possible solu-
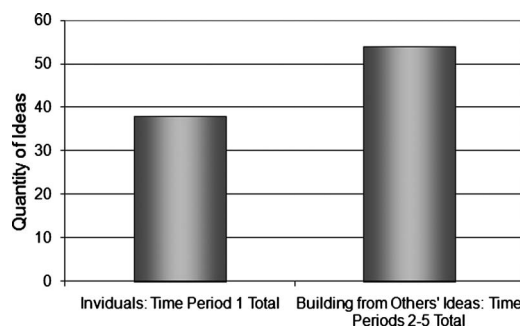
tions to this problem. ANOVA results show no significant differences for either the main or interaction effects (display: $F(1,6)=0.08$, $p>0.5$; representation: $F(2,6)=0.75$, $p>0.5$; and interaction: $F(2,6)=0.23$, $p>0.5$, and $MS_{error}=0.008$).

There is no difference in the total percentage of the design space evaluated by each method, but there are differences in the average amount of the design space per product solution (total percentage of design space divided by the number of product solutions). The viewing condition does affect the average amount of the design space covered by each product solution (Fig. 16). There is a significant main effect due to the viewing condition and no effect for the representation or interaction (display: $F(1,6)=6.06$, $p<0.05$; representation: $F(2,6)=0.22$, $p>0.5$; and interaction: $F(2,6)=0.58$, $p>0.5$, and $MS_{error}=1.4768\times10^{-05}$). Teams in the rotational conditions tended to use fewer product solutions while covering the same total percentage of the design space. This means that the product solutions from rotational viewing tend to span a great extent or a basis set of the design space. Rotational viewing tends to produce product solutions with greater average diversity and fewer product solutions but with greater breadth. While there is some variation in the amount of the design space covered by each product solution, overall teams search only a fraction of the design space, and additional idea generation methods are required to more fully explore the extent of the entire design space.

5.4 Quality. Each product solution generated by a team receives a quality score. The quality scores are evaluated with respect to two different approaches, the average for each team and the number of high quality product solutions for each team. ANOVAs were done on the average scores per team. As shown in Fig.

Fig. 14 Teams gain a significant number of ideas through collaboration. Results are from a single team in condition 6.

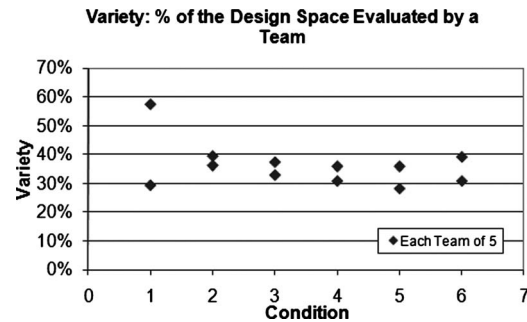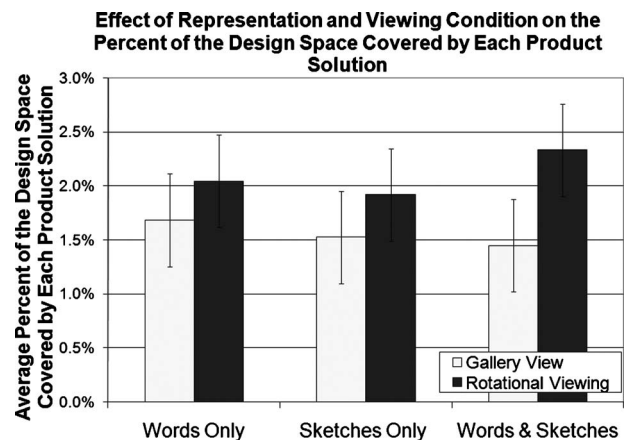**Fig. 16 The average amount of the design space covered by each product solution varies. Each error bar is ±1 standard deviation across all teams.**
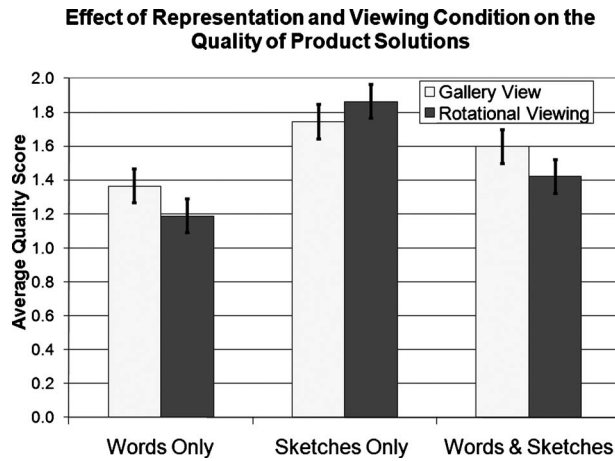
Effect of Representation and Viewing Condition on the Quality of Product Solutions

Fig. 17 Average quality of a product solution. Each error bar is ±1 standard error.

Effect of Representation and Viewing Condition on the Number of High Quality Product Solutions

Fig. 19 Number of high quality product solutions for each condition. Each error bar is ±1 standard error calculated across all conditions.

17, the average quality of a product solution is not influenced by the experimental factors. Overall, the quality of the product solutions is very high with an average of 1.5/2. Thus, the average product solution is deemed to be technically feasible, though not quite practical for the context. It is important to note that this is a fairly high quality of product solutions overall, particularly given that participants did not have specific experience with this domain. The average quality of product solution did not vary significantly across the factors; ANOVA results show no significance for either the main or interaction effects (Fig. 17) (display: $F_{(1,6)}=0.15$, $p>0.5$; representation: $F_{(2,6)}=2.31$, $p>0.15$; and interaction: $F_{(2,6)}=0.24$, $p>0.5$, and $MS_{error}=0.12$). The set of idea generation methods evaluated in this set of experiments is clearly useful for finding high quality, practical solutions to a given design problem. From the quality results, it is also apparent that the participants seriously committed themselves to solving the experimental problem, exerting a high level of effort in the process.

The distribution of the scores and the number of high quality product solutions (a score of 2) provides additional insights into the various techniques being evaluated in this study (Figs. 18 and 19). One group in condition 1 (words only, gallery view) generated an unusually high number of technically infeasible solutions (Fig. 18). In contrast, almost no technically infeasible solutions are generated by groups communicating with either only sketches or a combination of sketches and words (conditions 3–6). Groups in the sketches only condition created significantly more high quality product solutions than the other two representations (Fig. 19). ANOVA results show the main effect for representation and a close to significant result for the viewing condition (display: $F_{(1,6)}=2.95$, $p=0.14$; representation: $F_{(2,6)}=3.51$, $p<0.1$; and interaction: $F_{(2,6)}=0.44$, $p>0.5$, and $MS_{error}=20.58$). Compar-
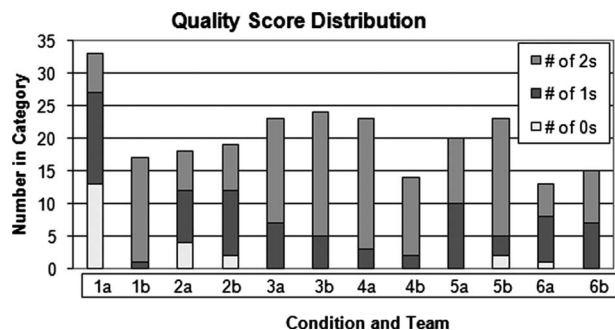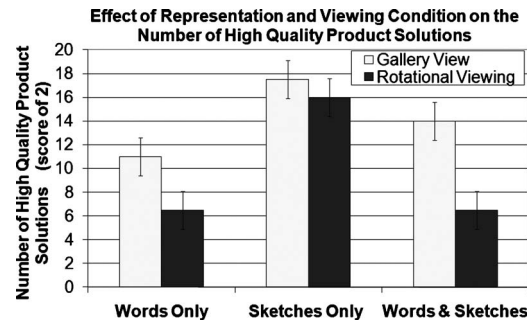
ing the means of the words only and words combined with sketches conditions to the sketches only condition using a linear contrast [59] shows that the sketches only condition results in more high quality product solutions ($F_{(1,6)}=6.8$, $p<0.05$). This comparison indicates that more high quality product solutions are being produced by the sketches only condition, but in combination with the quantity results, fewer ideas are being generated. This study does not address why this occurred, but one possibility is that it is very difficult to draw many of the solutions that were considered low quality such as genetically engineered solutions or chemicals to dissolve the peanut shell.

5.5 Change in Quality Over Time. The quality of a product solution frequently changes as team members add their ideas (for example, in Fig. 9). The product solution begins as a large-scale machine powered by a wind or water mill. As ideas are added, it became a hand-powered, more complete, and feasible system. In this experiment, two methods are available for the teams to build off each others' ideas and change the quality of the product solution. They can add their ideas directly onto the same sheet (embellish) or start a new sheet and include a cross-reference or "link" to a previous product solution. Prior studies have evaluated the design process in terms of how one idea is linked to other ideas [17,60]. Product solutions that are embellished tend to be higher quality product solutions (Tables 7 and 8).

As individuals add ideas, the overall product solution can drastically improve. For example, Fig. 9 shows the successive additions made to one sketch. The initial product solution is interesting, but it is probably impractical to import energy into this system. During the fourth time period, importing human energy

Fig. 18 Distribution of quality scores (quality scores: 1 =technically feasible and 2=feasible for the context)

Table 7 Quality of product solution added or linked to

	Percent of product solutions for a given quality score (%)		
	0	1	2
Linked to	14	7	10
Added to at least once	23	44	43
Not linked or added to	64	47	54

Table 8 High quality product solutions are linked to most often

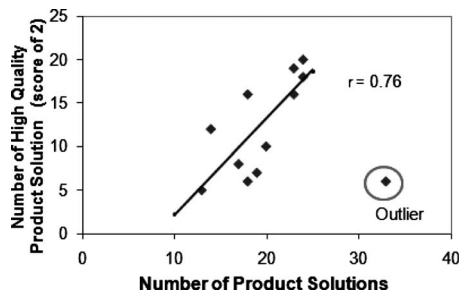|	Initial linked product solution quality score		
	0	1	2
Number of product solutions	3	8	29

Fig. 20 A greater number of product solutions results in more high quality ones

through a hand crank is added. This change is intriguing because the physical size of the system using a wind or water turbine is drastically different from the size of the system using a hand crank. The lack of dimensions in the sketch promotes improvements to the product solution. Each time period results in more solutions to the required functions and overall a more complete product solution. This high quality product solution is very similar to a solution currently used for shelling peanuts in third world countries [50].

5.6 Additional Analysis: Correlation Between Quality and Quantity. We find a high correlation between the number of product solutions generated by a team and the number of high quality product solutions they produced (quality score of 2) (Fig. 20). This result is consistent with the anecdotal evidence stating that quantity increases quality [61,62] and consistent with related brainstorming studies [48,63]. Pearson's correlation coefficient is 0.76 ($p < 0.01$) when one outlier is not included in the analysis (Fig. 20). The outlier team produced a large number of product solutions but few high quality ones. The quantity of ideas does not correlate significantly with the number of high quality product solutions ($r = -0.2$).

6 Addressing the Research Questions

The data provide significant insights into the effects of the two factors on the idea generation process and gives guidance for the approaches engineering design teams should use. The following discussion provides further insights based on the results.

6.1 Question 1: Do the Techniques Being Tested in This Experiment Vary in the Quantity of Ideas Generated or the Novelty or Variety of Product Solutions? Words combined with sketches and rotational viewing produce the largest number of ideas. A combination of words and sketches produces about 20% more ideas than words alone. Also, sketches only produces more ideas than words only. Ward's path of least resistance model, for how new ideas are structured by information in memory [64], applies to this question. As people begin to categorize a problem in a particular way (for example, by seeing other people's ideas), the more their memory of existing products or physical systems will affect the new designs and fewer ideas will be generated. Gallery viewing provides more product solutions to be viewed simultaneously, and therefore teams may more quickly categorize the problem in a similar manner. This result could explain why gallery viewing produces fewer ideas than rotational viewing.

The effects of adding sketches may be underestimated because the current method does not take into account the geometry, layout, or overall configuration of the sketch. This type of information is frequently included in sketches but rarely in verbal descriptions. In addition, as engineers sketch, they tend to add many details that they may not include if they give a verbal description of a device. For example, if an individual describes the use of a motor and gear train to power a system, the drawing will fre-

quently include parts such as shafts, bearings, and supports, but when only a verbal description is used, these supporting details are not included.

All conditions in this experiment are virtually equal in terms of variety and novelty. Some very novel product solutions were produced, but there is more potential in this area. Overall, teams explored only a fraction of the design space. Other idea generation methods need to be sought in order to improve the variety and novelty of the solutions.

In this experiment, all members of the team generated ideas without communicating during the first time period. This independent step may have led the teams to develop a similar variety of product solutions regardless of the later communication modes resulting in an equal level of variety and novelty for all conditions. If this idea generation session had been compared with other idea generation techniques, such as Osborn's brainstorming, the results may be different.

Gallery viewing produces more global product solutions and more high quality product solutions but overall fewer single functional ideas. McKoy et al. [65] also found that sketches result in higher quality product solutions than sentential descriptions. In contrast, rotational viewing produces fewer global product solutions and fewer high quality product solutions but an overall greater number of ideas per function. The difference in number of high quality approaches statistical significance ($p = 0.14$). The rotational conditions use a smaller number of product solutions to span the same fraction of the design space as the gallery conditions, where the average diversity for each product solution is greater. This result suggests that an improved process for idea generation consists of first using a gallery communication method to generate a large number of high quality product solutions and then moving to a rotational viewing method using words and sketches to develop the details of the product solutions and a large number of functional ideas.

6.2 Question 2: Does the Modality of Representation Interact With the Display Method or Are They Virtually Independent? Representation and the viewing method interact for the quantity of ideas but not for the other metrics. This interaction reflects that the sketches only conditions follow a different pattern of results compared with the other conditions. Rotational viewing produced more ideas when there were words in the representation but fewer ideas when there were only sketches. One hypothesis is that regardless of representation, rotational viewing encourages the participants to spend more time understanding other teammate's ideas. Unfortunately, the results of this effort vary based on representation. In the sketches only conditions, this process has a detrimental effect because the sketches without any verbal descriptions may be difficult to interpret, and therefore more effort is applied to understanding, to the detriment of generating more ideas. In general, engineers are not taught to draw, and their skill in sketching may be lacking. In contrast, the other two representations are relatively straightforward to interpret, and less time is spent in attempting to interpret the ideas, and thus more ideas are produced.

One reason why the sketches only conditions are different from those conditions in which subjects use words is that sketches tend to contain fewer general, abstract product solutions. Sketches may be viewed as being more detailed and concrete, therefore placing constraints on what the participants perceived to be allowable changes, or the physical extent of the product solution. If so, then sketches may bias participants toward making less radical changes to the product solutions than are made to product solutions described in words. Another possible hypothesis for the differences is that sketches are vaguer than words, and so they produce more unintended interpretations. In linguistics, vagueness reflects the degree to which a given statement is open to different interpretations. These hypotheses, and particularly the role of vagueness in sketches, should be the subject of future research.

6.3 Question 3: Are Certain Representations Better for Producing or Improving the Quality of Solutions? Do Certain Representations Cause Bias Toward Certain Types of Product Solutions?

The average quality is similar across conditions, but the sketches only conditions result in a greater number of high quality product solutions. One possible reason for this result is that there are certain categories of product solutions that are difficult to draw. Participants may have thought of these product solutions then did not attempt to draw them, refocusing their attention on product solutions that they could sketch. For this particular problem, these difficult-to-draw product solutions tend to be of lower quality, including product solutions such as using chemicals to dissolve the shells and genetically modifying the peanuts. The "sketching only" condition acts to filter out these low quality solutions by virtue of the difficulty in embodying them.

Some information, particularly abstract product solutions, is easier to convey in words, whereas other information, such as geometry and configuration, tends to be easier to convey with drawings. Most design problems involve a combination of these two types of information. In addition, verbal overshadowing suggests that conditions using written words may have a disadvantage compared with using only sketches. The verbalization of perceptual information can interfere with the retrieval of perceptual information from memory. The verbal overshadowing effect has the potential to make it difficult to retrieve information on how a device moves and related pieces of highly visual information that is difficult to verbalize.

7 Discussion of Additional Results

7.1 Does Building off Teammate's Product Solution Improve the Quality of the Product Solution? How Does Adding Modifications to a Design Compare With More Drastic Links From One Product Solution to the Next?

Product solutions that are added to tend to be better product solutions. As ideas are added, there is potential for significant improvement in a product solution's overall completeness and quality (Fig. 9). When a completely new product solution is built from a previous one, this process usually produces new product solutions that are equal in quality to the old product solution. High quality product solutions are linked much more frequently than lower quality ones, showing that as teams build from others' ideas, there is an implicit evaluation taking place.

7.2 How Do the Contributions of the Individuals Before the Product Solutions Are Shared With the Group Compare With the Number of Ideas the Group Generates by Building From These Initial Product Solutions?

A large number of ideas are developed when individuals work alone. Teams are likewise able to develop a significant number of ideas by sharing product solutions (Fig. 14). Therefore, both individual and group works are important in the idea generation process. The number of ideas teams produced during time period 5 is equal to that during time periods 2–4, indicating that the participants' ideas are not exhausted by time period 5 (Fig. 13). Open questions include the following: What are the reasons engineering teams choose to stop, at what point and for what reasons would these student teams have chosen to end the idea generation session? What criteria do engineering teams use when deciding to stop developing ideas? Is it when a feasible product solution is found or due to time constraints? These questions are the topics for future investigations. Further work needs to be completed to determine how long individuals should work alone prior to sharing their ideas.

7.3 Seeking a Large Quantity of Product Solutions to Achieve High Quality.

Teams with a greater number of ideas tended to produce more high quality product solutions. The guideline to produce a large quantity of ideas in order to achieve high quality is supported by these data. Anecdotal evidence in engineering design also strongly suggests that the two are correlated [61,62]. Early brainstorming studies also find this to be true [48].

More recent work with electronic brainstorming agrees with this but cautions that quality may not always track quantity [63].

8 Conclusion

Creativity and innovation are significant driving factors in engineering design and are what draws many individuals to the profession. To support innovation, we must seek improved methods for idea generation. This paper addresses important elements of the idea generation, or ideation, process. While past research in psychology, engineering design, and other fields has included human studies in idea creation, there is still much to be learned about the underlying factors of many of the popular idea generation techniques. A number of anecdotes exist about the advantages and disadvantages of the techniques, and some quantitative results also exist that address the aggregate methods in a group setting. By using a systematic approach, we identify key factors that differentiate the methods and that may be exploited to create more effective techniques.

We have uncovered a number of important insights. The choice of group idea generation method significantly affects the total quantity of ideas generated as well as the number of high quality concepts generated, but no significant effects for either novelty or variety. Over the 40 min session, 50% more ideas are generated using rotational viewing combined with ideas being described with words and sketches compared with using only words and displaying them "gallery style." This experimental condition corresponds to a hybrid 6-3-5/C-sketch method and should be taught to engineers. In contrast, more high quality concepts result when all product solutions are displayed on the wall, gallery viewing, and represented using only sketches. These results suggest an improved process for idea generation consisting of first using a gallery communication method to generate a large number of high quality product solutions and then moving to a rotational viewing method using words and sketches to develop the details of the product solutions and a large number of ideas.

This study also shows that both individual and group interactions are important in the idea generation process. As group members add ideas, the overall product solution becomes more complete and improves in quality. Participants do not simply create their own product solutions in isolation. An equal or greater number of new ideas are developed that build upon or are directly influenced by other group members. Engineers need to be taught the importance of both individual and group idea generation. That visualizing others' ideas produces even more ideas is not just an anecdote. Our data show that group member's ideas "spark" other members to a greater level of productivity.

9 Future Work

A design team must have confidence that it has found very good and innovative solutions to its design problem. An encompassing exploration of the design space increases the probability of this outcome. In general, teams explore only a segment of the total design space, and the methods tested do not differ in the breadth of search that they encourage. One team did evaluate a greater diversity of product solutions, showing that it is possible given the short time period of the experiment. Future work will focus on approaches to encourage design teams to more fully explore the design space and find more novel solutions. New analogy-based design methods have the potential to greatly enhance novelty and variety [58,66–72]. These new methods will be based on a deep understanding of the semantic and visual representations' influence on the design-by-analogy process. Systematic studies of analogy retrieval and use in the design process will be undertaken [58,70]. The use of functional models and other approaches for problem definition and representation in the design-by-analogy process will also be investigated. The final product will be a robust design-by-analogy search methodology capable of identifying nonobvious conceptual analogies resulting in novel design solutions.

Additional work to more fully understand the various available idea generation methods and their effects over longer periods of time is also required. Two factors are explored in this study, but many more are present in the methods. Further evaluation is needed to understand the influence of limiting communication to sketches and to more fully explain the pattern of results for the sketches only conditions.

Acknowledgment

References

[1] Römer, A., Weißhahn, G., and Hacker, W., 2001, "Effort-Saving Product Representations in Design—Results of a Questionnaire Survey," Des. Stud., **22**(6), pp. 473–490.

[2] Adams, J. L., 1986, *Conceptual Blockbusting*, Perseus Books, Cambridge, MA.

[3] Vangundy, A. B., 1988, *Techniques of Structured Problem Solving*, Van Nostrand Reinhold, New York.

[4] Higgins, J. M., 1994, *101 Creative Problem Solving Techniques*, The New Management, Winter Park, FL.

[5] Osborn, A., 1957, *Applied Imagination*, Scribner, New York.

[6] Altshuller, G. S., 1984, *Creativity as an Exact Science*, Gordon and Breach, Luxembourg.

[7] Lewis, A. C., Sadosky, T. L., and Connolly, T., 1975, "Effectiveness of Group Brainstorming in Engineering Problem Solving," IEEE Trans. Eng. Manage., **EM-22**(3), pp. 119–124.

[8] Mullen, B., Johnson, C., and Salas, E., 1991, "Productivity Loss in Brainstorming Groups: A Meta-Analytic Integration," Basic Appl. Soc. Psych., **12**(1), pp. 3–23.

[9] Collins, A. M., and Loftus, E. F., 1975, "A Spreading-Activation Theory of Semantic Priming," Psychol. Rev., **82**, pp. 407–428.

[10] Anderson, J. R., 1983, "A Spreading Activation Theory of Memory," J. Verbal Learn. Verbal Behav., **22**, pp. 261–295.

[11] Raaijmakers, J. G., and Shiffrin, R. M., 1981, "Search of Associative Memory," Psychol. Rev., **88**(2), pp. 93–134.

[12] Perttula, M. K., and Liikkanen, L. A., 2006, "Structural Tendencies and Exposure Effects in Design Idea Generation," Proceedings of the ASME Design Engineering Technical Conference.

[13] Perttula, M., Krause, C., and Sipila, P., 2006, "Does Idea Exchange Promote Productivity in Design Idea Generation?" CoDesign: International Journal of CoCreation in Design and the Arts, **2**(3), pp. 125–138.

[14] Perttula, M., and Sipila, P., 2007, "The Idea Exposure Paradigm in Design Idea Generation," J. Eng. Design, **18**(1), pp. 93–102.

[15] Gryskiewicz, S. S., 1988, "Trial by Fire in an Industrial Setting: A Practical Evaluation of Three Creative Problem-Solving Techniques," *Innovation: A Cross-Disciplinary Perspective*, Norwegian University Press, Oslo, Norway,.

[16] Paulus, P. B., and Yang, H. C., 2000, "Idea Generation in Groups: A Basis for Creativity in Organizations," Org. Behav. Hum. Decis. Process, **82**(1), pp. 76–87.

[17] Van Der Lugt, R., 2002, "Brainsketching and How It Differs From Brainstorming," Creat. Innov. Manag., **11**(1), pp. 43–54.

[18] Shah, J. J., 1998, "Experimental Investigation of Progressive Idea Generation Techniques in Engineering Design," DETC '98, 1998 ASME Design Engineering Technical Conferences.

[19] Shah, J. J., Kulkarni, S. V., and Vargas-Hernández, N., 2000, "Evaluation of Idea Generation Methods for Conceptual Design: Effectiveness Metrics and Design of Experiments," ASME J. Mech. Des., **122**, pp. 377–384.

[20] Shah, J. J., Smith, S. M., and Vargas-Hernandez, N., 2003, "Metrics for Measuring Ideation Effectiveness," Des. Stud., **24**(2), pp. 111–134.

[21] Vidal, R., Mulet, E., and Gómez-Senent, E., 2004, "Effectiveness of the Means of Expression in Creative Problem-Solving in Design Groups," J. Eng. Design, **15**(3), pp. 285–298.

[22] Otto, K., and Wood, K., 2001, *Product Design: Techniques in Reverse Engineering and New Product Development*, Prentice-Hall, Upper Saddle River, NJ.

[23] Shah, J. J., Vargas-Hernández, N., Summers, J. S., and Kulkarni, S., 2001, "Collaborative Sketching (C-Sketch)—An Idea Generation Technique for Engineering Design," J. Creat. Behav., **35**(3), pp. 168–198.

[24] Pahl, G., and Beitz, W., 1996, *Engineering Design—A Systematic Approach*, Springer, New York.

[25] Markman, A., 1999, *Knowledge Representation*, Lawrence Erlbaum Associates, Mahwah, NJ.

[26] Wilson, M., 2002, "Six Views of Embodied Cognition," Psychon. Bull. Rev., **9**(4), pp. 625–636.

[27] Goel, V., 1995, *Sketches of Thought*, MIT, Cambridge, MA.

[28] Greer, J. L., Jensen, D. D., and Wood, K. L., 2004, "Effort Flow Analysis: A Methodology for Directed Product Evolution," Des. Stud., **25**(2), pp. 193–214.

[29] Prinz, J. J., and Barsolou, L. W., 2002, "Acquisition and Productivity in Perceptual Symbol Systems: An Account of Mundane Creativity," *Creativity, Cognition, and Knowledge: An Interaction*, Praeger, Westport, CT.

[30] Suwa, M., and Tversky, B., 1996, "What Architects See in Their Sketches: Implications for Design Tools," *Human Factors in Computing Systems: Conference Companion*, Association for Computing Machinery (ACM).

[31] Suwa, M., and Tversky, B., 1997, "What Do Architects and Students Perceive in Their Design Sketches? A Protocol Analysis," Des. Stud., **18**(4), pp. 385–403.

[32] Goldschmidt, G., and Smolkov, M., 2006, "Variances in the Impact of Visual Stimuli on Design Problem Solving Performance," Des. Stud., **27**(5), pp. 549–569.

[33] Paivio, A., 1986, *Mental Representations: A Dual Coding Approach*, Oxford University Press, New York.

[34] Mandler, G., 1980, "Recognizing: The Judgment of Previous Occurrence," Psychol. Rev., **87**, pp. 252–271.

[35] Loftus, G. R., and Kallman, H. J., 1979, "Encoding and Use of Detail Information in Picture Recognition," J. Exp. Psychol. [Hum Learn], **5**, pp. 197–211.

[36] Bartlett, J. C., Till, R. E., and Leavy, J. C., 1980, "Retrieval Characteristics of Complex Pictures: Effects of Verbal Encoding," J. Verbal Learn. Verbal Behav., **19**, pp. 430–449.

[37] Schooler, J. W., Fiore, S. M., and Brandimonte, M. A., 1997, "At a Loss From Words: Verbal Overshadowing of Perceptual Memories," *The Psychology of Learning and Motivation*, Academic, New York.

[38] Aiken, M., Vanjani, M., and Paolillo, J., 1996, "A Comparison of Two Electronic Idea Generation Techniques," Inf. Manage., **30**, pp. 91–99.

[39] Jensen, D. D., Murphy, M. D., and Wood, K. L., 1998, "Evaluation and Refinement of a Restructured Introduction to Engineering Design Course Using Student Surveys and MBTI Data," ASEE Annual Conference.

[40] Jensen, D., and Bowe, M., 1999, "Hands-on Experiences to Enhance Learning of Design: Effectiveness in a Reverse Engineering/Redesign Context When Correlated With MBTI and Vark Types," ASEE Annual Conference.

[41] Jensen, D., Feland, J., Bowe, M., and Self, B., 2000, "A 6-Hats Based Team Formation Strategy: Development and Comparison With an MBTI Based Approach," ASEE Annual Conference.

[42] Jensen, D., and Wood, K., 2000, "Incorporating Learning Styles to Enhance Mechanical Engineering Curricula by Restructuring Courses, Increasing Hands-on Activities, and Improving Team Dynamics," ASME Annual Conference.

[43] Wood, K., Jensen, D., Bezdek, J., and Otto, K., 2001, "Reverse Engineering and Redesign: Courses to Incrementally and Systematically Teach Design," J. Eng. Educ., **90**(3), pp. 363–374.

[44] 2004, Thinkspace: Peanut Sheller, March, http://www.thinkcycle.org/tc-space/tspace?tspace_id=41963

[45] Diehl, M., and Stroebe, W., 1987, "Productivity Loss in Brainstorming Groups: Toward the Solution of a Riddle," J. Pers. Soc. Psychol., **53**(3), pp. 497–509.

[46] Baxter, M., 1995, *Product Design: A Practical Guide to Systematic Methods of New Product Development*, Chapman and Hall, London.

[47] Bouchard, T. J., Jr., 1969, "Personality, Problem-Solving Procedure, and Performance in Small Groups," J. Appl. Psychol., **53**(1), pp. 1–29.

[48] Parnes, S. J., and Meadow, A., 1959, "Effects of 'Brainstorming' Instructions on Creative Problem Solving by Trained and Untrained Subjects," J. Educ. Psychol., **50**(4), pp. 171–176.

[49] Delhagen, W., Hussam, S., Mohdramli, R., and Yi, A., 2003, A Low-Cost Peanut Sheller for Use in Developing Nations: Final Application for Ideas, August 2006, http://web.mit.edu/ideas/www/past%20proposals/peanut.pdf

[50] 2006, Full Belly Project, August, http://www.fullbellyproject.org/

[51] 2006, American Peanut Council, August, http://www.peanutsusa.com/index.cfm?fuseaction=home.page&pid=12#Shelling_and_Grading

[52] Linsey, J., Green, M. G., Murphy, J. T., Wood, K. L., and Markman, A. B., 2005, "Collaborating to Success: An Experimental Study of Group Idea Generation Techniques," ASME Design Theory and Methodology Conference.

[53] Hirtz, J., Stone, R. B., and Mcadams, D. A., 2002, "A Functional Basis for Engineering Design: Reconciling and Evolving Previous Efforts," Res. Eng. Des., **13**, pp. 65–82.

[54] Stone, R., and Wood, K., 2000, "Development of a Functional Basis for Design," ASME J. Mech. Des., **122**(4), pp. 359–370.

[55] Clark-Carter, D., 1997, *Doing Quantitative Psychological Research: From Design to Report*, Psychology, Taylor & Francis, Inc., UK.

[56] Jansson, D., and Smith, S., 1991, "Design Fixation," Des. Stud., **12**(1), pp. 3–11.

[57] Kurtoglu, T., Campbell, M., and Linsey, J., 2009, "An Experimental Study on the Effects of a Computational Design Tool on Concept Generation," Des. Stud., **30**(6), pp. 676–703.

[58] Linsey, J., Laux, J., Clauss, E. F., Wood, K., and Markman, A., 2007, "Increasing Innovation: A Trilogy of Experiments Towards a Design-by-Analogy Method," ASME Design Theory and Methodology Conference.

[59] Howell, D. C., 2002, *Statistical Methods for Psychology*, Duxbury, Pacific Grove, CA.

[60] Goldschmidt, G., and Weil, M., 1998, "Contents and Structure in Design Reasoning," Design Issues, **14**(3), pp. 85–100.

[61] Ulrich, K., and Eppinger, S., 2004, *Product Design and Development*, McGraw-Hill, New York.
[62] Ullman, D., 1997, *The Mechanical Design Process*, McGraw-Hill, New York.
[63] Briggs, R. O., Reinig, B. A., Shepherd, M. M., Yen, J., and Nunameker, J. F., 1997, "Quality as a Function of Quantity in Electronic Brainstorming," 30th Hawaii International Conference on System Sciences.
[64] Ward, T. B., 1998, "Analogical Distance and Purpose in Creative Thought: Mental Leaps Versus Mental Hops," *Advances in Analogy Research: Integration of Theory and Data From the Cognitive, Computational, and Neural Sciences*, New Bulgarian University, Sofia, Bulgaria.
[65] McKoy, F. L., Vargas-Hernández, N., Summers, J. D., and Shah, J. J., 2001, "Influence of Design Representation on Effectiveness of Idea Generation," DETC '01, 2001 ASME Design Engineering Technical Conferences.
[66] Christensen, B. T., and Schunn, C., 2007, "The Relationship of Analogical Distance to Analogical Function and Pre-Inventive Structures: The Case of Engineering Design," Mem. Cognit., **35**(1), pp. 29–38.
[67] Gentner, D., and Markman, A. B., 1997, "Structure Mapping in Analogy and Similarity," Am. Psychol., **52**, pp. 45–56.
[68] Casakin, H., and Goldschmidt, G., 1999, "Expertise and the Use of Visual Analogy: Implications for Design Education," Des. Stud., **20**(2), pp. 153–175.
[69] Ball, L. J., Ormerod, T. C., and Morley, N. J., 2004, "Spontaneous Analogising in Engineering Design: A Comparative Analysis of Experts and Novices," Des. Stud., **25**(5), pp. 495–508.
[70] Linsey, J., Murphy, J. T., Wood, K. L., Markman, A. B., and Kurtoglu, T., 2006, "Representing Analogies: Increasing the Probability of Success," ASME Design Theory and Methodology Conference.
[71] McAdams, D. A., and Wood, K. L., 2002, "A Quantitative Similarity Metric for Design-by-Analogy," ASME J. Mech. Des., **124**(2), pp. 173–182.
[72] Linsey, J., Tseng, I., Fu, K., Cagan, J., Wood, K., and Schunn, C., 2010, "A Study of Design Fixation, Its Mitigation and Perception in Engineering Design Faculty," ASME J. Mech. Des., **132**(4), p. 041003.