

# *Metrics for measuring ideation effectiveness*

Jami J. Shah and Noe Vargas-Hernandez, Mechanical & Aerospace Eng., Arizona State University, Tempe, Arizona 85287-6106, USA

Steve M. Smith, Dept. of Psychology, Texas A & M University, College Station, Texas 77843-4235, USA

*Systematic methods for idea generation in engineering design have come about from a variety of sources. Do these methods really aid ideation? Some empirical studies have been conducted by researchers to answer this question. These studies include highly controlled lab experiments by cognitive psychologists, as well as experiments in simulated design environments carried out by engineering design theorists. A key factor in design and analysis of empirical studies is characterization and measurement of ideation effectiveness. This paper describes four objective measures of ideation effectiveness. The theoretical basis of each is discussed and procedures for application of each are outlined and illustrated with case studies.*

© 2002 Elsevier Science Ltd. All rights reserved.

*Keywords: design cognition, creative design, conceptual design, engineering design, ideation*

Cognitive psychologists consider fluency and novelty as the primary measures of ability to generate ideas<sup>1,2</sup>. However, this is not quite adequate for engineering designs. As shown in Fig. 1, an engineering design must not only be novel (unusual, unexpected) but it must also satisfy some intended function(s) to desired specifications (have desired utility). Thus, engineering designs must be evaluated by metrics somewhat different from those for non-utilitarian artefacts. Engineering designs do not happen by accident; they must satisfy a set of pre-defined set of specifications, even if these specs sometimes get modified as the designer and client both get a better understanding of the design problem and design space. Thus, design is goal oriented. A designer's success is judged by how well his/her design meets desired goals and how well he/she has identified the alternative ways of achieving the those goals.

A comprehensive study of design ideation must identify operating variables

**1 Torrance, E P** Role of Evaluation in Creative Thinking Bureau of Educational Research, University of Minnesota, MN (1964)

**2 Torrance, E P** Guiding Creative Talent Prentice Hall, Englewood Cliffs, NJ (1962)



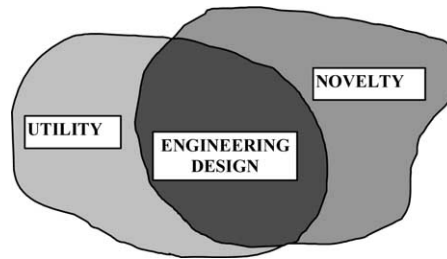


Figure 1 Engineering design at the intersection of creativity and utility.

**3 Shah, J** 'Experimental Investigation of Collaborative Techniques for Progressive Idea Generation' in Proceedings of ASME Design Theory and Methodology Conference, Atlanta, GA (1998)

**4 Shah, J, Kulkarni, S and Vargas-Hernandez, N** 'Guidelines for experimental evaluation of idea generation methods in conceptual design' *Journal of Mechanical Design* Vol 122 No 4 (2000) 377–384

**5 Zwicky, P** *Discovery, Invention, Research through Morphological Analysis* McMillan, New York (1969)

**6 Osborn, A** *Applied Imagination* Scribners, New York (1979).

**7 Hogarth, R** *Judgment and Choice—The Psychology of Decision* Wiley, New York (1980)

**8 DeBono, E** *Lateral Thinking: Creativity Step by Step* Harper and Row, New York (1970)

**9 Rhorbach, B** 'Creative nach Regeln: Methode 635, eine neue Technik zum Lösen von Problemen' *Absatzwirtschaft* Vol 12 (1969)

**10 Shah, J** Method 5-1-4 G—A variation on Method 635, MAE 541 Class Notes, Arizona State University (1993).

**11 VanGundy, A B** *Techniques of structured problem solving*, 2nd ed. Van Nostrand Reinhold Company, NY (1988)

**12 Quantum Improvement Consulting** Affinity diagrams, <http://shell.rmi.net/~lifestar/qiwizard/affinity.html> (1998).

**13 Fogler, H and Le Blanc, S** *Strategies for Creative Problem Solving* Prentice Hall (1995)

**14 Gordon, W J J** *Synectics, the Development of Creative Capacity* Harper, New York (1961)

and key ideation components of design methods, model them in terms of atomic cognitive processes and structures, relate them to outcome metrics, and model interaction effects among components. Developing such a model of design ideation will allow one to evaluate existing ideation methods and predict performance in various conditions, and may lead to new theoretically based methods. This paper looks at *only one aspect* of ideation research—the identification of metrics to measure ideation effectiveness of design idea generation methods. These measures are the result of collaboration between a mechanical design researcher and a cognitive psychologist.

Many ideation methods have been developed to aid designers generate alternative designs. Formal idea generation methods are broadly classified into two categories—*intuitive* and *logical*. Intuitive methods use mechanisms to break what are believed to be mental blocks. Logical methods involve systematic decomposition and analysis of the problem, relying heavily on technical databases and direct use of science and engineering principles and/or catalogues of solutions or procedures.

Intuitive Methods have been sub-classified into five categories<sup>3,4</sup>: *Germinal, Transformational, Progressive, Organizational, and Hybrid*. Germinal methods aim to produce ideas from scratch. Some examples are Morphological Analysis<sup>5</sup>, Brainstorming<sup>6</sup> and the K–J Method<sup>7</sup>. Transformational Methods generate ideas by modifying existing ones, and include methods like Checklists<sup>6</sup>, Random Stimuli<sup>8</sup>, and PMI Method<sup>8</sup>. Progressive Methods generate ideas by repeating the same steps many times, generating ideas in discrete progressive steps. Examples of Progressive Methods are Method 635<sup>9</sup>, C-Sketch<sup>10</sup>, and Gallery Method<sup>11</sup>. Organizational Methods help designers group generate ideas in some meaningful way. The Affinity Method<sup>12</sup>, Storyboarding<sup>11</sup>, and Fishbone Diagrams<sup>13</sup> belong to this class of methods. Hybrid methods like Synectics<sup>14</sup> combine different techniques to address varying needs at different phases of ideation.

Logical methods may be classified into two categories: *History Based* and *Analytical*. History Based Methods use past solutions catalogued in some

**15 Pahl, G and Beitz, W** *Engineering Design—A Systematic Approach*, 2nd ed. Springer, London (1996)

**16 Altshuller, G** *Creativity as an Exact Science* Gordon and Breach, New York (1984)

**17 Shigley, J and Uicker, J** *Theory of Machines & Mechanisms* McGraw Hill (1995)

**18 Sickafus, N** *Unified Structured Inventive Thinking: How to Invent*. NTELLEK (1998).

**19 Kulkarni, S V** *Idea Generation Techniques—a critical survey*, Tech. Report ASU/DAL/IG/98-2, Arizona State University (1998).

**20 Shah, J, Vargas-Hernandez, N, Summers, J D and Kulkarni, S** 'Collaborative sketching (S-Sketch)' *Journal of Creative Behavior* Vol 35 No 3 (2001) 168–198

**21 Kulkarni, S V** *Survey for Evidence of 'Components of Creativity'*, Tech. Report ASU/DAL/IG/99-7, Arizona State University, AZ (1999).

**22 Marples, D** *The Decisions of Engineering Design* Inst. of Eng. Designs, London (1960)

**23 Ericsson, K and Simon, H** *Protocol Analysis—Verbal Reports as Data* MIT Press (1984)

**24 Sobek, D and Ward, A** 'Principles from Toyota's Set-Based Concurrent Engineering Process' in *Proceedings of ASME Computers in Engineering Conference*, Irvine, CA (1996)

**25 Waldron, M and Brook, R** 'Analyzing Inter and Intra Group Information Exchanges in Conceptual Collaborative Design' in *Proceedings of ASME Design Theory and Methodology Conference* (1994)

**26 Ullman, D, Wood, S and Craig, D** 'Importance of Drawing in Mechanical Design Process' in *NSF DTM Conference* (1989)

**27 Ullman, D, Dieterich, T and Stauffer, L** 'A model of the mechanical design process based on empirical data A1' *EDAM* Vol 2 No 1 (1988)

**28 Ullman, D, Stauffer, L and Dieterich, T** 'Preliminary Results on an Experimental Study of Mechanical Design' in *NSF DTM Conference* (1987)

**29 Christiaans, H and Dorst, K** 'An Empirical Study Into Design Thinking' in *Research in Design Thinking*, Delft Press, (1991)

form of database. The German school produced catalogues of both physical effects and solutions corresponding to a variety of mechanical functions<sup>15</sup>.

The Russian school devised TRIZ<sup>16</sup> by collecting so-called invention principles by examining patents from different domains. Analytical methods develop ideas from first principles by systematically analysing basic relations, causal chains, and desirable/undesirable attributes. In *Forward Steps*<sup>15</sup> one analyses variations of initial solutions. Inversion is a standard method used in kinematics to create new types of mechanisms<sup>17</sup>. SIT began as a variation of TRIZ at FORD but became more general and independent of databases<sup>18</sup>. Surveys of idea generation methods can be found in Van Gundy<sup>11</sup>, Shah<sup>3</sup>, and Kulkarni<sup>19</sup>.

Despite several claims and much anecdotal evidence about the usefulness of these methods, little formal experimental evidence exists currently to indicate that these methods are effective in engineering design to generate concepts. Further, the rules and procedures for these methods seem to have been specified arbitrarily, regardless of the nature of the problem being solved. We need metrics to experimentally evaluate the effectiveness of these methods for different kinds of design problems.

## *1 Literature review*

There have been sporadic attempts at studying ideation processes both in cognitive science and design theory. The models in cognitive theories are derived from highly controlled lab experiments involving simple and isolated tasks. There is little similarity between conditions for these experiments and design concept generation in the real world. On the other hand, direct experiments using design idea generation methods in their entirety, simulate real world design better but cannot discriminate between necessary and superfluous components<sup>21</sup>, require a prohibitive number of experiments, and are unable to explain the performance of methods under different conditions. These two types of studies have opposite ecological and internal validity, but both types are needed to develop holistic models of design ideation. Regardless of the ecological level of the experimental studies, one needs metrics for analysing the effectiveness of ideation methods and processes. Ideally, the community should have some consensus on these measures so as to provide a basis for comparing the consistency of the results.

Several experimental methods have been used for studying the design process and/or its associated cognitive activities. These include case studies<sup>16,22,23</sup>, protocol studies<sup>24–31</sup>, and controlled tests<sup>32</sup>. There have been some studies in the past of designers working in teams<sup>33–35</sup>. Christiaans<sup>29,30</sup> studied individual industrial designers of varying creative levels and found

- 30 Christiaans, H and Venselaar, K** 'Practical Implications of a Knowledge Based Design Approach' in *Research in Design Thinking*, Delft Press, (1991)
- 31 Cross, N, Dorst, K, Roozeburg, N** *Research in Design Thinking*. Proceedings of Workshop Delft Univ. Press (1991).
- 32 Schön, D** 'Teaching and Learning as a Design Transaction' in *Research in Design Thinking*, Delft Press, (1991)
- 33 Hale, C** *Analysis of the Engineer Design Process in an Industrial Context* Grant Hill Pubs, Cambridge MA (1987)
- 34 Nagy, R, Ullman, D and Dietterich, T** 'A data representation for collaborative mechanical design' *Research in Eng. Design* Vol 3 (1993)
- 35 Leifer, L** Presentation at DARPA-RaDEO contractors meeting, Melbourne FL (1996).
- 36 Stauffer, L and Ullman, D** 'Fundamental processes of mechanical designers based on empirical data' *Journal of Engineering Design* Vol 2 No 3 (1991) 113–125
- 37 Koestler, A** *The Art of Creation* Hutchinson and Co., London (1964)
- 38 Wallas, G** *The Art of Thought* Harcourt, New York (1926)
- 39 Simonton, D K** *Scientific Genius: a Psychology of Science* Cambridge University Press, Cambridge MA (1988)
- 40 Finke, R A, Ward, T B and Smith, S M** *Creative Cognition: Theory, Research, and Applications* MIT Press, Cambridge MA (1992)
- 41 Smith, S M, Ward, T B and Finke, R A** *The Creative Cognition Approach* MIT Press, Cambridge MA (1995)
- 42 Ward, T B, Smith, S M and Vaid, J** 'Conceptual structures and processes in creative thought' in **T B Ward, S M Smith and J Vaid** (eds) *An Investigation of Conceptual Structures and Processes*, American Psych. Association, Washington DC (1997)
- 43 Smith, S M** 'Creative cognition: 'Demystifying Creativity'' in **C N Hedley** (ed.) *Thinking and Literacy—the Mind at Work*, Lawrence Erlbaum Associates, Hillsdale NJ (1995)

differences in frequency of data collection and information processing. There is not much reported on experimental studies of *specific* idea generation methods applied to engineering design, particularly groups engaged in conceptual design (except studies on Brainstorming or 'free form' idea generation). From analysis of 'think-aloud' video and audio protocols, Ullman et al.<sup>27</sup> proposed the task/episode accumulation model (TEAM), a set of ten operators to describe non-routine mechanical design. Further, Stauffer and Ullman<sup>36</sup> identified unique sequences of TEAM operators they called 'local methods'. From protocol studies of conceptual design, Shah et al.<sup>10</sup> also defined a set of generic cognitive activities and patterns in which they recurred.

Some cognitive models have been developed specifically for design team activities. Hale<sup>33</sup> studied industrial designers during design review sessions. Nagy et al.<sup>34</sup> proposed a data representation for collaborative mechanical design. Waldron and Brook<sup>25</sup> focused in the communication between and within groups. The activities in the group can be of three categories: Gaps, Helps, and Decisions. Leifer<sup>35</sup> measured the frequency of communication and its relation to success rate. Each of the above studies used ad-hoc measures of ideation effectiveness in their analysis.

There are several areas in cognitive psychology that may have relevance to the study of design ideation. Of these, two are of particular significance: study of technological creativity and perception. Early attempts at explaining technological creativity include Koestler's Bisociation Theory<sup>37</sup>, Wallas Model<sup>38</sup>, and Chance-Configuration Theory<sup>39</sup>. Finke, Ward and Smith's *Creative Cognition* approach developed at Texas A&M<sup>40–42</sup> served as a theoretical impetus that spurred several research initiatives, each relating cognitive structures and processes to various aspects of creative activities. The Roadmap Theory proposed by Smith<sup>43</sup> explains creativity through basic cognitive structures and processes, moving between mental states from an initial problem state to a goal or solution state. Fixation corresponds to dead-end branching; incubation allows escape from dead-ends. The Connectionist theory<sup>44</sup> explains four stages of the creative process in terms of focusing/defocusing attention, the number of nodes activated, and activation level. It is not clear how the theory can be used to study idea generation methods. The only measurable quantity described is cortical arousal, which is not appropriate for this study. The Computational Model of Scientific Insight<sup>45</sup> is based on two separate lines of research in cognitive science—reasoning by analogy and qualitative mental models. The Generate model<sup>40</sup> divides creative mental processes into Generative and Exploratory. Examples of Generative processes are memory retrieval, association, synthesis, transformation, analogical transfer, and categorical

**44 Martindale, C** 'Creativity and connectionism' in **S M Smith** (ed.) *The Creative Cognition Approach*, MIT Press, Cambridge MA (1995)

**45 Langley, P and Jones, R** 'Computational model of scientific insight' in **R J Sternberg** (ed.) *The Nature of Creativity—Contemporary Psychological Perspectives*, Cambridge University Press, Cambridge, NY (1988)

**46 Hoffmann, O and Kollingbaum, M** 'Creativity as transformation: multi-agent system and human cognition' in **L Candy and E Edmonds** (eds) *Creativity and Cognition 1996 Proceedings*, AVS Services, Loughborough University, Loughborough (1996) pp 10–26

**47 Schmid, K** 'Towards creative AI systems: a psychology-based approach' in **L Candy and E Edmonds** (eds) *Creativity and Cognition 1996 Proceedings*, AVS Services, Loughborough University, Loughborough (1996) pp 123–134

**48 Candy, L** 'Understanding creativity: an empirical approach' in **L Candy and E Edmonds** (eds) *Creativity and Cognition 1996 Proceedings*, Loughborough University, Loughborough (1996)

**49 Boden, M** *The Creative Mind: Myths and Mechanism* Weidenfeld & Nicolson, London (1990)

**50 Elton, M** 'Towards artificial creativity' in **L Candy and E Edmonds** (eds) *Creativity and Cognition 1993 Proceedings*, Loughborough University, Loughborough (1993) pp 76–83

**51 Cross, N** 'Creativity in design: not leaping but bridging' in **L Candy and E Edmonds** (eds) *Creativity and Cognition 1996 Proceedings*, AVS Services, Loughborough University, Loughborough (1996) pp 27–35

**52 Dorst, K and Cross, N** 'Protocol analysis as a research technique for analyzing design activity' *ASME Design Engineering Technical Conferences* Vol DE-83 No 2 (1995)

**53 Kulkarni, S V** Survey of creativity models. Technical Report ASU/DAL/IG/98-1, Arizona State University (1998).

reduction. Examples of Exploratory processes are attribute finding, conceptual interpretation, functional inference, contextual shifting, hypothesis testing, searching for limitations. These cognitive theories were derived either from pure conjecture or controlled experiments that use simple tasks. The suitability of these models for design problems that are much more complex has never been investigated.

Some AI researchers tend to model the ideation process from the information-processing point of view. For Hoffman and Kollingbaum<sup>46</sup> the individual starts with an original model (priori knowledge or structural information) and some fluctuating information. This process implies a mapping from the constraint space to the specification space. Schmid<sup>47</sup> proposes a model of the cognitive process that has three aspects: Operators, Memory, and Control. Elton<sup>50</sup> classifies ideation processes into Generation (subsuming preparation and incubation) and Evaluation (subsuming illumination and verification), stressing the role of evaluation and contending that generation without evaluation is not creativity. According to Cross<sup>51</sup>, the generation of creative thoughts can be described with four generalized models, these are: Analogy, Combination, First Principles, and Emergence. Many of these 'models' or 'theories' are pure conjectures not supported by experiments.

## 2 Ideation measures

In order to carry out empirical studies of design ideation at any level, one must specify how effectiveness of ideation is to be measured. Two issues need to be addressed for this purpose—*what* should be measured, and *how* should it be measured. Further, should the idea generation process be evaluated, or is it better to simply evaluate the outcome, i.e. the ideas generated? We will refer to these as *process based* and *outcome based* approaches.

To base the evaluation of ideation methods on the characteristics of the process rather than the outcome requires data collection via protocol studies and analysis using ideation cognitive models. Such models must contain a classification of cognitive process types, attributes that can be used to recognize them, and an understanding of their role in promoting ideation, the influence of the frequency of occurrence on the outcome, a threshold value of frequency to have any impact, and so on. We do not have this level of understanding of ideation today<sup>40,53</sup>. We did not arrive at this conclusion by speculation. We conducted through surveys of relevant areas of cognitive science<sup>1,2,37–53</sup>, in collaboration with specialists in that field. Prior to this realization, we also conducted process based studies at first. In an experiment conducted at ASU, *generative* and *exploratory* cognitive

processes were identified by videotaping designers who were asked to 'think aloud' while they were generating ideas using the C-Sketch method<sup>3</sup>. However, it was difficult to observe these cognitive processes. There are no commonly agreed upon techniques to analyse the data collected from such protocol studies<sup>52</sup>. More importantly, the cognitive processes described by psychologists in cognitive models such as the *Geneplore Model*<sup>40</sup>, the *Roadmap Theory*<sup>43</sup> and other models described in<sup>53</sup>, are based on controlled experiments that involve simple problems or tasks. The appropriateness of the use of these models in experiments involving engineering problems, that are far more complex by comparison, is subject to speculation. Further, it is difficult to directly relate the occurrence of cognitive processes to the effectiveness of an idea generation method<sup>54</sup>.

The aforementioned difficulties in process based measurement of ideation effectiveness led us to consider outcome based metrics. The premise of the outcome based approach is that an idea generation method is considered effective if its use results in 'good' ideas. While engineers are alien to evaluating cognitive processes they are quite used to evaluating design ideas. We can give any set of design ideas to a domain expert and ask them to evaluate them for a given design problem, and we will have no difficulty getting answers from the expert. Now the question is how to relate measures of goodness for design ideas to measures of goodness of idea generation methods. Two basic criteria are identified for this purpose:

- how well does the method expand the design space
- how well does the method explore the design space.

Design space is the count of all possible options for a given problem<sup>63</sup>. Design space is not fully known. A designer may start out with a tentative set of possibilities; some of these may need to be dropped because they do not meet requirements upon further analysis (reduction of space by quality measurement); removal of fictitious constraints, or discovery of a novel way of doing something would expand the design space. Thus, pushing the limits of design space is a good thing at the conceptual phase of design when divergent thinking is called for. Exploring the design space implies the number and variety of design alternatives discovered. Dylla has shown that there is a significant correlation between the percentage of the design space covered and quality of the final product<sup>64</sup>. Neither of these mentions creativity—term we avoid because of difficulty in defining this term (and agreeing on its meaning).

**54** Smith, S M, Ward, T B Discussion at Texas A&M University, June 25, 1999, College Station TX (1999).

Based on the above two criteria and the discussion above, four separate effectiveness measures are proposed: *novelty*, *variety*, *quality* and *quantity*

of the ideas generated using that method. *Novelty* is a measure of how unusual or unexpected an idea is as compared to other ideas. Not every new idea is novel since it may be considered usual or expected to some degree and this is only known after the idea is obtained and analysed. *Variety* is a measure of the explored solution space during the idea generation process. The generation of similar ideas indicates low variety and hence, less probability of finding better ideas in other areas of the solution space. *Quality*, in this context, is a measure of the feasibility of an idea and how close it comes to meet the design specifications. *Quantity* is the total number of ideas generated. The rationale for this measure is that generating more ideas increases the chance of better ideas<sup>6,55–59</sup>. Classical literature in psychometric psychology use fluency (how prolific one is in generating ideas) as a measure of creativity<sup>1,2</sup>. To calculate these measures, the design artefact is decomposed into its desired key functions. Weights can be assigned to each.

These four metrics measure different aspects of ideation effectiveness. Their independence is not an issue; they are all important in measuring one aspect of effectiveness. For example, weight and cost of a structure are not independent, since material cost is dependent on weight, but both may be used as separate measures of goodness in evaluating a structure. We will now give the details about each of these measures. Each measure will be defined and justified, its measurement procedure described and illustrated with the help of examples.

**55 Candy, L and Edmonds, E A** 'Creative design of the Lotus bicycle: implications for knowledge support systems research' *Design Studies* Vol 17 No 1 (1996) 71–90

**56 Cross, N and Cross, A C** 'Winning by design: the methods of Gordon Murray, racing car designer' *Design Studies* Vol 17 No 1 (1996) 91–107

**57 Kumar, V K, Holman, E R and Rudegeair, P** 'Creativity styles of freshmen students' *Journal of Creative Behavior* Vol 25 No 4 (1991) 275–303

**58 Basadur, M S and Thompson, R** 'Usefulness of the ideation principle of extended effort in real world professional and managerial creative problem solving' *Journal of Creative Behavior* Vol 20 No 1 (1986) 23–34

**59 Parnes, S J** 'Effects of extended effort in creative problem solving' *Journal of Educational Psychology* Vol 52 No 3 (1961)

## 2.1 Novelty

Two approaches may be taken to measure novelty. The universe of ideas for comparison can be obtained by defining what is not novel (what is usual or expected), preferably before analysing any data to avoid biasing. Alternatively, we can collect all ideas generated by all participants from all methods, identify key attributes such as motion type, control mechanism, propulsion, etc. Then find all the different ways in which each of those attributes is satisfied (example: motion=rotating, sliding, oscillating, etc.). Then we can count how many instances of each solution method exist in the entire collection of ideas. The lower the count (i.e., the less a characteristic is found) the higher the novelty.

### 2.1.1 Measurement procedure

The problem is first decomposed into its key functions or characteristics. Every idea produced is analysed by first identifying which functions it satisfies and describing how it fulfils these functions, at the conceptual and/or embodiment level. Each description is then graded for novelty according to one of two approaches. It is possible to compute a total score

for novelty *for each idea*, by applying the weights to each function and stage. Overall novelty of each idea can be computed from (1)

$$M_1 = \sum_{j=1}^m f_j \sum_{k=1}^n S_{1jk} p_k \quad (1)$$

$M_1$  is the overall novelty score for the idea with  $m$  functions or attributes and  $n$  stages. Weights ( $f_j$ ) are assigned according to the importance of each function or characteristic in order to compute an overall score. Further, each function may be addressed at the conceptual and/or embodiment stage and weights ( $p_k$ ) assigned according to the stage's importance.

The calculation of  $S_1$  depends on the approach chosen. For the first approach (*a priori* knowledge) a universe of ideas for comparison is subjectively defined for each function or attribute, and at each stage. A novelty score  $S_1$  is assigned at each idea in this universe. To evaluate the function and stage of an idea a closest match is found in the table and the score  $S_1$  noted. For the second approach,  $S_1$  is calculated from (2)

$$S_{1jk} = \frac{T_{jk} - C_{jk}}{T_{jk}} \times 10 \quad (2)$$

Where  $T_{jk}$  is the total number of ideas produced for function (or key attribute)  $j$  and stage  $k$ ; and  $C_{jk}$  is the count of the current solution for that function (or key attribute) and stage. Multiplying by 10 normalizes the expression. This metric has also been used by psychologists to measure creativity<sup>1,2,60</sup>.

### 2.1.2 Justification

The use of a measure of novelty in idea generation is of fundamental importance. In terms of design space, novel designs occupy points that are initially not perceived to be within the design space. Expanding the design space offers the opportunity to find better designs that are so far not known to exist. Many idea generation methods provide deliberate mechanisms to view the problem in a different way, to use analogies and metaphors, to play around by loosening the tight grip on goals that engineers generally have.

Novelty can be assessed at multiple levels, depending upon the scale. The simplest level is personal novelty, in which an individual discovers or creates ideas that are new to the individual. For example, even though a thousand others might have solved the same problem, when one first solves an



unfamiliar puzzle, the solution can be said to have personal novelty. A higher level is societal novelty, in which a product or idea is new to all people in a particular society, regardless of whether the product is commonplace in other societies. At the highest level is historical novelty, in which a product or idea is the first of its kind in the history of all societies and civilizations.

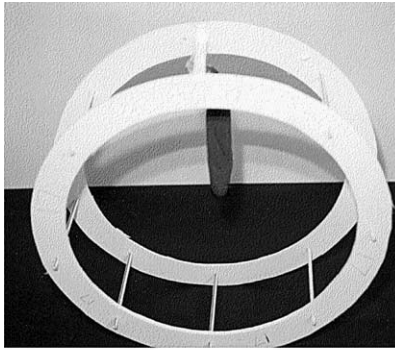
### *2.1.3 Example*

We illustrate the application of the above procedure with the help of a design problem used in a student design competition. The objective of the design was to build a device made from a fixed set of materials and powered by a fixed volume of pressurized air. The device that travelled the farthest from the starting position would be the winner. The group of designs that we are considering in this example (Fig. 2, Table 3) did not come from any particular ideation method, nor did they come from the same person or group. But this is irrelevant because we are merely using this set of designs to illustrate how novelty can be determined in an objective, auditable way.

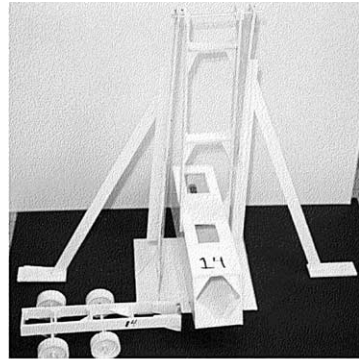
We will illustrate the evaluation of this group of ideas only at the conceptual stage. First, one must choose the attributes on which novelty assessment is to be based; this will depend on the specifics of the design problem and objectives. For this particular case, the key functions and characteristics were identified as:

- Propulsion/thrust method (jet, sail, paddlewheel, etc)
- Medium of travel (air, land, water)
- Motion of device (rolling, sliding, gliding, tumbling,..)

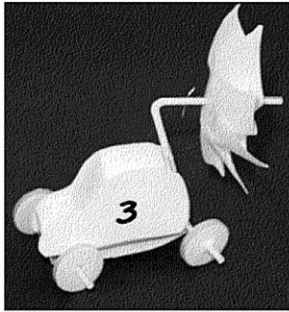
Additionally, the number of pieces into which the device separated in operation was found to be another distinguishing characteristic for evaluating novelty. Most people would normally assume that the device had to be a single piece, i.e. this is a fictitious constraint that limits the design space to monolithic devices. Table 3 shows the characteristics of 46 entries in this contest and Fig. 2 shows photos of just a few of them. As is evident from the table the most common solution was a ground based vehicle with rolling wheels with some sort of jet propulsion. A few devices travelled in air and only two in water—this latter choice of medium was unusual and unexpected. There were a small number (3) vehicles that used a tumbling motion; again, this can be regarded as unusual. Both a peer review and independent judges found #19 and #34 to be the most unusual. One used a burst of energy by instantaneous release of pressure to fling a projectile and the other used elastic energy of the container like a catapult. Both broke some perceptual blocks.



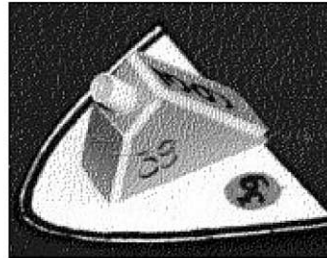
Entry #17



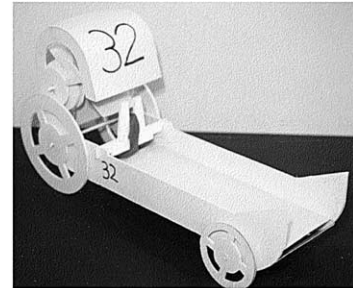
Entry #14



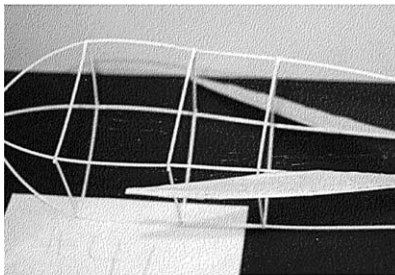
Entry #3



Entry #39



Entry #32



Entry #41



Entry #19

Figure 2 Partial set of design entries.

Using Eq. (1) for novelty, and the four attributes identified above ( $m = 4$ ) with weights as follows: Thrust ( $f_1 = 0.35$ ), Medium of travel ( $f_2 = 0.35$ ), Motion ( $f_3 = 0.20$ ), and #pieces in operation ( $f_4 = 0.1$ ). Only one design stage was used, so  $n = 1$ . For the first novelty calculation approach, the universe of solution for each of the three key attributes is pre-defined and values assigned to them *a priori* (Table 4).

There were only three levels of values chosen (3, 7, 10); thus, this set

**Table 3 Design data for Example 1**

Entry #	Novelty attributes		Medium <i>Wt=0.35</i>	Motion <i>Wt=0.20</i>	Paris <i>Wt=0.1</i>	Novelty scores			Contest performance	
	Thrust <i>Wt=0.35</i>					<i>M<sub>i</sub></i> priori	<i>M<sub>i</sub></i> post	<i>M<sub>i</sub></i> Peer	Distance (ft)	Speed ft/s
1	Jet		Ground	Rolling on wheels	1	3.00	3.28	6.00	0.25	2.00
2	Jet		Ground	Rolling on wheels	1	3.00	3.28	6.33	3.42	13.67
3	Jet on Sail		Ground	Rolling on wheels	2	3.40	5.81	7.33	3.67	0.57
4	Jet		Water	Gliding	1	6.25	6.45	8.43	5.50	0.11
5	Turbine		Ground	Rolling on wheels	1	4.40	4.81	6.43	14.00	1.75
6	Jet		Ground	Rolling on wheels	2	3.40	3.99	5.90	12.10	1.39
7	Jet		Ground	Rolling on wheels	1	3.00	3.28	6.87	6.50	12.33
8	Jet		Ground	Rolling on wheels	1	3.00	3.28	5.80	20.45	2.05
9	Turbine		Ground	Rolling on wheels	1	4.40	4.81	6.07	7.50	0.54
10	Turbine/PW		Ground	Rolling on Wheels	1	4.40	5.18	6.40	7.50	1.02
11	Jet		Ground	Rolling on Wheels	1	3.28	3.28	5.93	0.50	0.40
12	Jet		Water	Gliding	1	6.25	6.45	8.73	16.80	0.34
13	Turbine		Ground	Rolling on wheels	1	4.40	4.81	6.83	18.25	13.50
14	Impact		Air	Gliding	2	8.05	8.46	7.08	8.42	2.57
15	Jet on Sail		Ground	Rolling on wheels	1	3.00	5.11	7.17	4.67	0.90
16	Hovercraft		Air	Gliding	1	7.65	7.67	8.03	-1.67	-5.57
17	Jet		Ground	Tumbling Vehicle	1	4.40	4.46	6.73	18.00	1.19
18	Jet		Ground	Rolling on wheels	1	3.00	3.28	6.67	25.50	4.08
19	Explosion		Air	Gliding	2	8.05	8.46	7.97	3.92	1.31
20	Jet		Ground	Rolling on wheels	1	3.00	3.28	5.80	0	0
21	Turbine		Ground	Rolling on wheels	1	4.40	4.81	6.70	4.30	0.51
22	Paddlewheel		Ground	Rolling on wheels	1	5.45	4.88	7.00	11.20	1.85
23	Jet		Ground	Tumbling Vehicle	1	4.40	4.46	6.37	3.42	0.34
24	Jet		Ground	Tumbling Vehicle	2	4.80	5.16	6.53	3.30	0.33

(continued on next page)

**Table 3 Design data for Example 1 (Continued)**

Entry #	Novelty attributes		Medium Wt=0.35	Motion Wt=0.20	Paris Wt=0.1	Novelty scores		M <sub>1</sub> Peer	Contest performance	
	Thrust Wt=0.35					M <sub>1</sub> priori	M <sub>1</sub> post		Distance (ft)	Speed ft/s
25	Jet	Air		Gliding	1	5.20	5.85	7.63	15.87	20.10
26	Turbine	Ground		Rolling on wheels	1	4.40	4.81	7.23	8.00	0.80
27	Paddlewheel	Ground		Rolling on wheels	1	5.45	4.88	6.57	7.09	1.22
28	Jet	Ground		Rolling on wheels	1	3.00	3.28	5.90	0.00	0.00
29	Jet	Air		Gliding	2	5.60	6.56	6.50	11.16	4.09
30	Paddlewheel	Ground		Rolling on wheels	1	5.45	4.88	6.97	9.42	1.54
31	Jet	Ground		Rolling on wheels	1	3.00	3.28	5.40	0.00	0.00
32	Turbine	Ground		Rolling on wheels	1	4.40	4.81	7.00	0.67	0.96
33	Hovercraft	Air		Gliding	1	7.65	7.67	8.43	12.00	3.24
34	Catapult	Air		Gliding	2	8.05	8.46	5.43	200.00	51.33
35	Jet	Ground		Rolling on wheels	1	3.00	3.28	5.77	3.67	1.04
36	Jet	Ground		Rolling on wheels	1	3.00	3.28	6.37	20.00	4.25
37	Jet	Air		Gliding	1	5.20	5.85	7.60	11.00	5.24
38	Jet	Ground		Gliding	1	3.80	4.02	7.30	4.25	1.86
39	Hovercraft /jet	Air		Gliding	1	7.65	7.75	7.97	1.50	3.00
40	Jet	Ground		Rolling on wheels	1	3.00	3.28	6.53	21.00	16.47
41	Jet	Air		Gliding	1	5.20	5.85	8.03	0.87	0.26
42	Jet	Ground		Rolling on wheels	1	3.00	3.28	6.47	25.00	1.43
43	Paddlewheel	Ground		Rolling on wheels	1	5.45	4.88	6.63	6.50	0.97
44	Jet	Ground		Rolling on wheels	1	3.00	3.28	6.07	2.33	4.66
45	Jet	Ground		Rolling on wheels	1	3.00	3.28	6.17	12.58	3.84
46	Paddlewheel	Ground		Rolling on wheels	1	5.45	4.88	6.78	7.92	1.62

**Table 4 Novelty scores, a priori value assignment**

<i>J</i>	<i>Attribute</i>	<i>Novelty sub-score S<sub>j</sub></i>		
		<i>S<sub>j</sub>=3</i>	<i>S<sub>j</sub>=7</i>	<i>S<sub>j</sub>=10</i>
1	Propulsion	Jet	Turbine	Other
2	Medium	Ground	Air	Other
3	Motion	Wheels	Fly	Other
4	Parts	1	2	2+

was of rather coarse granularity. The most common/expected solution (e.g. travelling on the ground) was given the least value (3); unexpected solutions given the highest value (10). Using these pre-assigned values, the novelty scores are calculated as follows.

The attributes of each idea are scored according to the ones in the Table 4. For example, entry #1 uses a Jet for thrust ( $S_{11} = 3$ ), Ground as the travel medium ( $S_{12} = 3$ ), Rolling on wheels for motion ( $S_{13} = 3$ ), and operates as a single unit ( $S_{14} = 3$ ). The novelty  $M_1$  of entry #1 is calculated from (1) as follows. For single stage evaluation, Eq. (1) reduces to:

$$M_1 = \sum_{j=1}^4 f_j S_{1j}.$$

Then for entry #1

$$M_1 = \sum_{j=1}^4 f_j S_{1j} = f_1 S_{11} + f_2 S_{12} + f_3 S_{13} + f_4 S_{14} = (0.35)*3 + (0.35)*3 + (0.2)*3 + (0.1)*3 = 3.00.$$

Entry #19 uses an Explosion for thrust (other  $-S_{11} = 10$ ), Air as the travel medium ( $S_{12} = 7$ ), gliding motion ( $S_{13} = 7$ ) and operates in 2 units ( $S_{14} = 7$ ). The novelty  $M_1$  of entry #19 is calculated from (1) as follows:

$$M_1 = \sum_{j=1}^4 f_j S_{1j} = f_1 S_{11} + f_2 S_{12} + f_3 S_{13} + f_4 S_{14} = (0.35)*10 + (0.35)*7 + (0.2)*7 + (0.1)*7 = 8.05.$$

The second method for computing novelty rating relies on *a posteriori* classification and counting of distinct solutions ideas and use of Eq. (2).

Table 5 shows counts of each solution for each attribute and their corresponding score  $S_1$  from Eq. (2). For example, since only two designs out of 46 used water as the medium of travel (attribute 2), the individual value of this attribute is found as follows.

$$T_{12} = 46, C_{12} = 2 \quad \Rightarrow \quad S_{12} = (46-2)*10/46 = 9.56.$$

All  $S_1$  scores (for each attribute) have been calculated this way and are shown in Table 5. To get the novelty score for any idea, its attributes are compared to those in Table 5 and the scores assigned accordingly. For example, entry #1 uses a Jet for thrust ( $S_{11} = 4.35$ ), Ground as travel medium ( $S_{12} = 2.61$ ), rolling wheels for Motion ( $S_{13} = 3.48$ ), and operates as a single unit ( $S_{14} = 1.52$ ). The novelty  $M_1$  for entry #1 is found from (1):

$$M_1 = \sum_{j=1}^4 f_j S_{1j} = (0.35)*4.35 + (0.35)*2.61 + (0.2)*3.48 + (0.1)*1.52 = 3.28.$$

Entry #19 uses an Explosion for thrust ( $S_{11} = 9.78$ ), air as travel medium ( $S_{12} = 7.83$ ), Gliding Motion ( $S_{13} = 7.17$ ), and operates with 2 separate units ( $S_{14} = 8.58$ ). The novelty  $M_1$  of entry #1 is calculated from Eqs. (1) and (2) as follows:

$$M_1 = \sum_{j=1}^4 f_j S_{1j} = (0.35)*9.78 + (0.35)*7.83 + (0.2)*7.17 + (0.1)*8.58 = 8.46.$$

It is interesting to compare the two calculation procedures for  $M_1$  with

**Table 5 Novelty scores, a posteriori feature counting and values from Eq. (2)**

<i>Propulsion</i>	$C_j$	$S_{11}$	<i>Medium of travel</i>	$C_j$	$S_{12}$	<i>Motion</i>	$C_j$	$S_{13}$	<i>Parts</i>	$C_j$	$S_{14}$
Jet	26	4.35	Ground	34	2.61	Rolling on wheels	30	3.48	1	39	1.52
Jet on Sail	2	9.56	Water	2	9.56	Gliding	13	7.17	2	7	8.58
Turbine	6	8.70	Air	10	7.83	Tumbling Vehicle	3	9.35			
Turb/Paddlewheel	1	9.78									
Impact	1	9.78									
Hovercraft	2	9.56									
Explosion	1	9.78									
Paddlewheel	5	8.91									
Catapult	1	9.78									
Hover/Jet	1	9.78									

each other and with human judges. The group of competing students was asked to evaluate the designs of their peers. The averages of these scores are also shown in Table 3. A correlation study was also conducted (Fig. 3). The correlation between the pre and post calculation methods was excellent (0.944). The correlation between peer evaluation and the two calculation methods was fair (0.624 for the post method and 0.591 for the a priori values). It should be noted, however, that these were not evaluations done by expert panel of judges but by the students themselves.

## 2.2 Variety

To measure variety, one examines how each function is satisfied. A variety rating applies to an entire group of ideas, not an individual idea. Ideas are grouped based on how different two ideas are from each other. The use of a different physical principle to satisfy the same function makes two ideas very different. On the other hand, if two ideas differ only in some secondary construction level detail, say a dimension value, the ideas are only slightly different. We elaborate on this in the following section.

### 2.2.1 Measurement procedure

The conceptual origins of ideas are analysed through a genealogical categorization based on how ideas fulfil each design function. A genealogy tree is shown in Fig. 4. At the highest level, ideas are differentiated by the different physical principles used by each to satisfy the same function; this is the most significant extent of finding differences between ideas. At the second level ideas are differentiated based on different working principles but they share the same physical principle. At the third and fourth levels, ideas have different embodiment and different detail, respectively. The nodes in the tree carry the count of ideas in each category at each level.

The number of branches in the tree gives an indication of the variety of

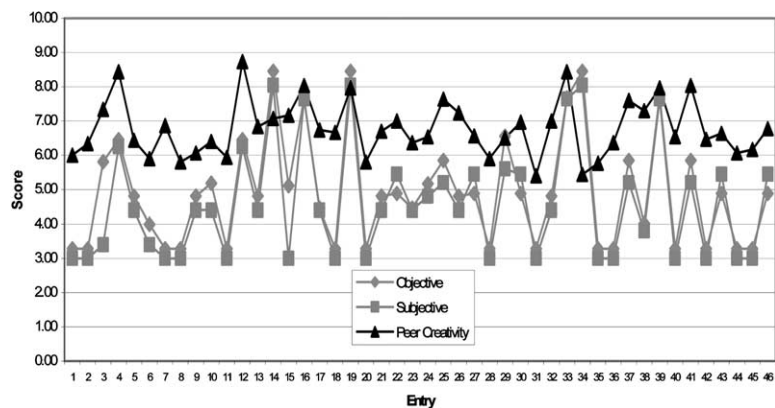


Figure 3 Correlation of the three novelty measures.

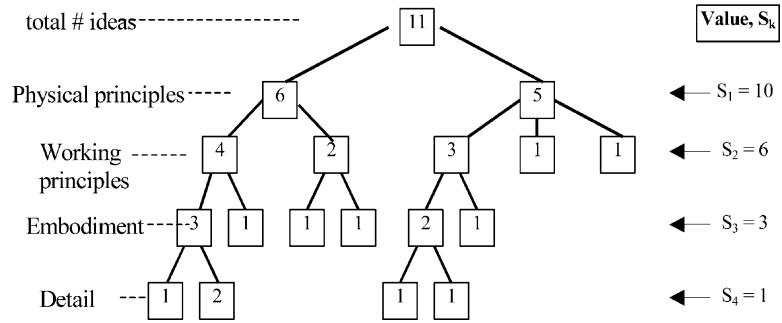


Figure 4 Genealogy of an idea set.

ideas. If greater variety is to be valued, branches at upper levels (physical principle differences) should get higher rating than the number of branches at lower levels. We have assigned values of 10, 6, 3, and 1 to physical principle, working principle, embodiment, and detail levels respectively. These values were chosen to ensure that separation at higher levels will always score a greater total. If there is only one branch at a given level, it shows no variety and the score should be zero; otherwise the score should be the number of branches times the value assigned to that level.

The genealogy tree needs to be constructed for each of the functions of a device. Not all the functions are equally important, so one can assign a weight  $f_j$  to account for the importance of each. Then the overall variety measure  $M_3$  takes the following form:

$$M_3 = 10 * \sum_{j=1}^m f_j \sum_{k=1}^4 S_k b_k / M_{3max} \quad (3)$$

where  $b_k$  is the no. of branches at level  $k$ ;  $S_k$  is the score for level  $k$  (suggested scores are 10, 6, 3, 1 for the four levels, respectively);  $m$  is the total number of functions;  $M_{3max}$  is the max possible variety score for the number of ideas in the set. The max score obviously would be obtained if all ideas used different physical principles. Thus,  $M_{3max}$  is total number of ideas times 10. Therefore, (3) reduces to

$$M_3 = \sum_{j=1}^m f_j \sum_{k=1}^4 S_k b_k / n.$$

For the numbers given in Fig. 4, the variety score for this function would be calculated as follows.  $m = 1$  (only one function being evaluated),  $f_j = 1$ ; total number of ideas,  $n = 11$ . There are two branches at level 1, five at level 2, six at level 3, four at level 4.



$$M_3 = ((2*10) + (5*6) + (6*3) + (4*1))/(11) = 72/11 = 6.54.$$

Of course, one does not need to conduct this analysis at these four levels for all designs; a subset of these may suffice. For example, if the ideas do not contain enough detail to go as far as the lowest level, and if it is hard to distinguish between physical and working principles, one could use just the working principles and embodiment levels.

### 2.2.2 *Justification*

Variety is an indication of how well one has explored the design space. Ideas that differ radically in physical principle are separated by large distances in design space. This measure is necessary to counterbalance the quantity measure. Producing a large number of ideas that differ from each other in minor or superficial ways does not prove effective idea generation. From a cognitive science point of view, variety in idea generation is a measure of the number of categories of ideas that one can imagine. The measure of variety is an indication of the multiple perspectives that one may use in solving a problem. Often, one finds that routine or hackneyed approaches to problems can lead to unsuccessful and uncreative ideas. In such cases, the original cognitive knowledge structures applied to a problem are inappropriate, and insight can be achieved only through what cognitive psychologists have called *cognitive restructuring*. The ability to generate a wide variety of ideas is directly related to the ability to restructure problems, and is therefore an important measure of creativity in design.

### 2.2.3 *Example*

To illustrate the application of the above measure and calculation procedure we look at two example sets of ideas. The two sets of ideas shown in Fig. 5 were generated by two different students. They were asked to design devices for a design competition in form of a game. The objective of the game was for a semi-autonomous device to collect golf balls from a playing field and bring them to a storage area. The score was based on the number and colour of golf balls collected in one minute. The two basic functions are pick/collect balls and transport/manoeuvre device. Each student generated several ideas; the sets of ideas generated by two of the students are shown in Figure 5. The set shown in (i) all designs appear to share the same basic principle to fulfil the functions: some form of shovel for pick/collect and wheels for transport. In set (ii), every idea uses a different principle for collecting. Note also that idea *b* in set (ii) is also quite novel in transport (no wheels).

The genealogy trees for each function and each of the two sets of Fig. 5 are shown in Tables 6 and 7. Note that the high variety sets have more

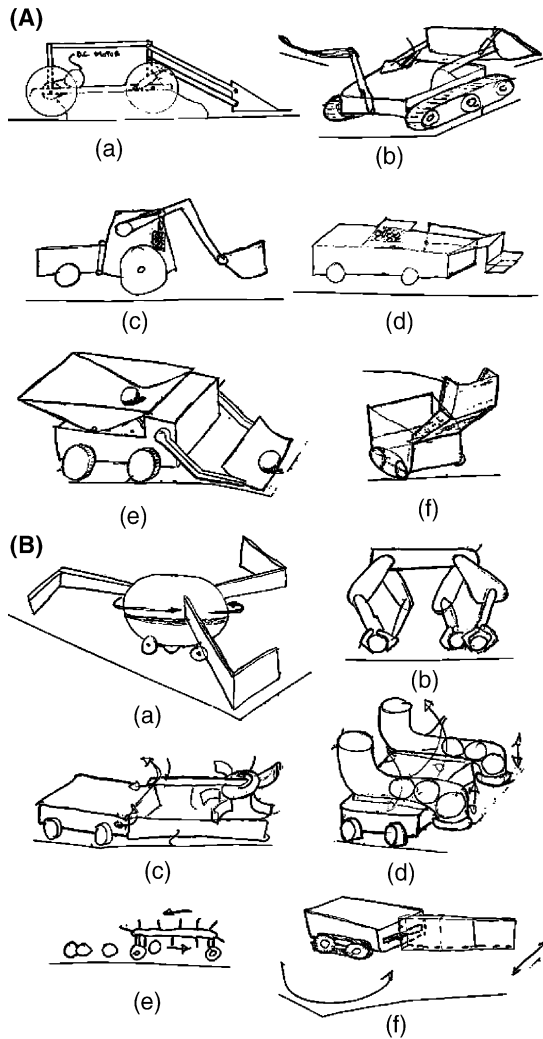


Figure 5 Comparing variety between two idea sets.

branches on the top, this will be reflected in the novelty scores. For this example, there are only two functions ( $j = 2$ ), Pick/Collect has a weight of  $f_1 = 0.6$  and Transport/Manoeuvre  $f_2 = 0.4$ . Eq. (3) is used to calculate the variety score  $M_3$ .

$$M_3 = \sum_{j=1}^2 f_j \sum_{k=1}^4 S_k b_k / n = [f_1(S_1 b_1 + S_2 b_2 + S_3 b_3 + S_4 b_4) + f_2(S_1 b_1 + S_2 b_2 + S_3 b_3 + S_4 b_4)] / n.$$

For the set in (i):

**Table 6 Genealogy tree for function: pick/collect**

Level	Score	Set (i)	Set (ii)
K	S <sub>k</sub>	Level	
1	10	Physical	
2	6	Working	
3	3	Embodiment	
4	1	Detail	

Level	Score	Set (i)	Set (ii)
6		Scoop	6
5		Shovel	Positive Engagement
4		2-Arm Wide Shovel	Rotating
3		a b c d e f	e. Translating
2			a c
1			b. Wedge f. Scoop d. Suction

**Table 7 Genealogy tree for function: transport/manoeuvre**

Level	Score	Low Variety Set	High Variety Set
K	S <sub>k</sub>	Level	
1	10	Physical	
2	6	Working	
3	3	Embodiment	
4	1	Detail	

Level	Score	Low Variety Set	High Variety Set
6		Rolling	6
5		Wheels	Rolling
4		Threads	b. Walking
3		a c d e b f	f. Threads
2			a c d e
1			

$$M_3 = 10*[0.6(10*0 + 6*0 + 3*0 + 1*6) + 0.4(10*0 + 6*2 + 3*6)]/6 = 15.6/6 = 2.6.$$

For the set in (ii):

$$M_3 = 10*[0.6(10*4 + 6*2 + 3*2 + 1*0) + 0.4(10*2 + 6*2 + 3*4)]/6 = 52.4/6 = 8.7.$$

### 2.3 Quality

The quality of an idea is an independent measure since it can be based on a physical property or ratio related to the performance of the artefact (time, weight, energy, etc.). At the conceptual stage, quality can usually be adequately estimated even though there is not enough quantitative information to do formal analysis. At the embodiment stage it may be possible to do some quantitative analysis and ratios of expected value to desired value of key attributes. These could be computed to quantify quality. It should be noted that the number of design phases evaluated (physical prin-

cept, concept, embodiment, detail, etc.) will depend on the type of idea generation method. For example brainstorming and 635 will have only one phase (concept) while C-Sketch and Gallery may have two or three.

### 2.3.1 Measurement procedure

Evaluation of technical feasibility and performance of design alternatives is the very essence of engineering. Questions, such as, “Can it get off the ground? How fast can it go? What is the probability of failure?” need to be answered. Evaluation uses both analytical and experiential knowledge. There are many domain specific procedures in each sub-discipline that are used first to determine key characteristics of a design, such as weight, fuel consumption, safety margin, acceleration and speed, etc. Then some general methods can be used to relate these characteristics to design objectives. QFD<sup>61</sup>, the Pugh matrix<sup>61</sup>, and Decision Tables<sup>15</sup> are popular general-purpose methods for selecting between design alternatives.

Any of the above methods can be used for determining the overall quality of a set of design alternatives generated. The only difference is that we add all the quality scores for all the alternatives to get the total score for the set. Thus, the quality rating  $M_2$  is found from (4):

$$M_2 = \frac{\sum_{j=1}^m f_j \sum_{k=1}^2 S_{jk} p_k}{n * \sum_{j=1}^m f_j} \quad (4)$$

$S_{jk}$  is the score for quality for function  $j$  at stage  $k$ ;  $m$  is the total number of functions;  $f_j$  is the weight for function  $j$ ;  $p_k$  is the weight for stage  $k$ . The denominator is for normalizing to a scale of 10.

### 2.3.2 Justification

The end game of engineering is to have a better (more marketable, profitable) product. Designs can be evaluated in objective ways. Regardless of how many designs and of a great a variety one generates, if there isn't one that would be physically feasible or competitive, all design effort will amount to naught. Therefore, evaluation of quality must also be included in the overall evaluation of effectiveness of ideation methods.

### 2.3.3 Example

Evaluation of designs at conceptual, embodiment, or detailed stage is such an integral part of engineering that it seems hardly necessary to illustrate this with an example. Also, evaluation procedures tend to be domain specific. Nevertheless, for the sake of completeness we illustrate the procedure with one example. We should clarify first that we are talking about *evalu-*

61 Huang, G Design for X: Concurrent Engineering Imperatives Chapman & Hall, London (1996)

ation of design ideas and not testing of an artefact after it has been built. So the contest results shown in Table 3 are *not* quality measures because these were obtained after the device was built. We have to evaluate quality of ideas to determine which ones should be developed further.

In this experiment, subjects were asked to design a link in a mechanism that could transmit intermittent tension/compression and torque through pin joints at either end. Four of the designs actually produced by the subjects are shown in Fig. 6. There were two quality criteria: minimum weight and manufacturability. The constraints were structural integrity (failure avoidance) and compatibility with adjoining links (allow relative rotation in one direction).

### 2.3.3.1 Minimum weight evaluation

Domain specific rules were used to evaluate each design. This includes the use of form synthesis rules for minimum weight structures. The annotations in Fig. 6 indicate the application and violation of form synthesis rules. These rules could be weighted differently to reflect their importance and

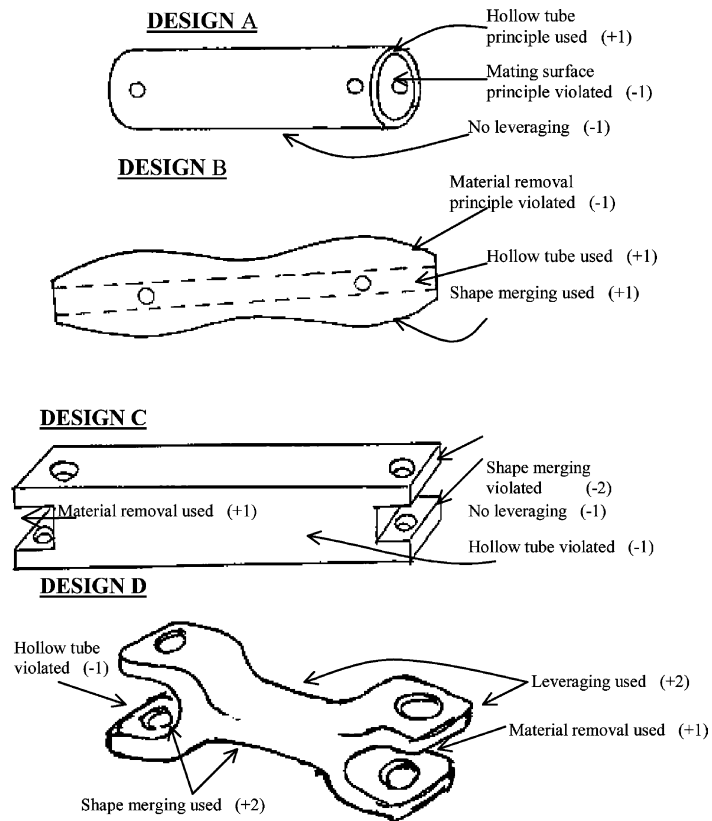


Figure 6 Four designs of connecting linkage.

severity of violations could carry higher penalty. But if we scored each application of a form synthesis rule as +1 and each violation as -1. Based on this, the scores for A, B, C, D are -1, +1, -3, and +4. (Refer to the figure for rules applied/violated and scoring).

### 2.3.3.2 Manufacturing analysis

Manufacturing analysis is also domain specific, using manufacturability rules or procedures to determine the number and type of operations, setups, standard/special tooling, fixturing, standard stock or cast/forged stock, etc. A is the simplest if made from standard stock with just a couple of drilling operations. C requires a little more to remove material from the ends. B would be difficult to manufacture, requiring non-standard turning tools. D is the hardest to manufacture; it requires cast/forged stock and special fixtures. One can use relative costing methods to estimate the ratings for each. If A is assumed to be the baseline at 10, C will be 8, B will be 4 and D will be 1. Regarding the constraints, any of the four can be sized for a given material to carry the loads. C and D are easily compatible; A only if the other links are much smaller or larger; B will be the most difficult to interface with other components.

Up to this point the evaluation is domain specific. Now let us apply a general procedure to get a quality rating for each design. If we apply the evaluation only at one level, Eq. (4) becomes:

$$M_2 = \frac{\sum_{j=1}^2 f_j S_j}{n * \sum_{j=1}^2 f_j},$$

where the two design objectives are minimum weight and maximum manufacturability. Suppose we make  $f_1 = 2*f_2$ . We cannot directly add the scores for each objective because they must be normalized to be added (be on the same scale). Manufacturability was already scaled to 1-10. Using the same scale for min. weight (min. score=-3, max. score=4), designs A-D are scored as 3.57, 6.14, 1.0, and 10. So this group evaluates as:

$$M_2 = ((2*3.57 + 1*10) + (2*6.14 + 1*4) + (2*1.0 + 1*8) + (2*10 + 1*1))/4*(2 + 1) = 64.42/12 = 5.37.$$

## 2.4 Quantity

*Quantity* is the total number of ideas generated by a group or individual during a designated amount of time or over the entire course of running through all the steps in a given design procedure. Some methods prescribe the number of ideas generated. For example, 6-3-5 has six ideas per par-

ticipant per cycle. In this case, it may not be too meaningful to use this measure unless comparing it to other methods.

Measuring the number of ideas generated raises the question, “when does one consider two ideas as different enough to count separately?” Since the extent of differences between ideas is already accounted for by measuring variety, we count all ideas that a participant submits or documents separately.

The justification for using quantity is that many believe that generating several ideas increases the chances of occurrence of better ideas<sup>6,48,51,57–59</sup>. Also classical literature in psychometric psychology uses fluency (how prolific one is in generating ideas) as a measure of an individual’s creativity<sup>1,2</sup>.

### 3 Closure

To date we do not have any comprehensive models of design creativity developed on scientific foundations. Such studies must systematically identify operating variables and key ideation components of design methods, model them in terms of atomic cognitive processes and structures, relate them to outcome metrics, and model interaction effects among components. Developing such a *model of design ideation* will allow one to evaluate existing ideation methods and predict performance in various conditions, and may lead to new theoretically based methods. This paper presented one important aspect of research that will some day lead to models of design ideation. We identified four types of outcome based metrics. Quantity, quality, novelty, and variety. We developed objective procedures for evaluating each. We gave the rationale for including each of these measures, both from a design point of view and cognitive psychology. By means of case studies, we demonstrated how to apply these procedures. Quantity and variety scores apply to the entire idea generation session, while novelty and quality scores are computed for each idea. The total quality and novelty scores are found by multiplying each idea by its respective score in that category and summing them to get the overall score for novelty or quality.

One question that now occurs is: Does it make sense to consolidate the scores for all four measures into an *overall* effectiveness measure. Since each of them measures something different, we feel that adding them directly makes no sense. Even if we were to normalize them in order to add, it is difficult to understand the meaning of such a measure. Besides, we may be interested in knowing how one method compares to another in terms of quantity vs novelty, etc. We can also argue that a method is worth using if it helps us with *any* of the measures.

These measures and methodology should be applicable in all domains of engineering design. We have refined these measures through their use in several studies in the past 5 years. The results of experiments are published elsewhere<sup>20,62</sup>. Our next task is to use these measures at multiple levels while aligning the experiments.

### *Acknowledgements*

This research was supported, in part, by US National Science Foundation (DMI9812646, DMI0115447) and Ford Motor Company.

**62** McKoy, F M, Vargas-Hernandez, N, Summers, J D and Shah, J 'Influence of Design Representation on Effectiveness of Idea Generation' in Proceedings of ASME Design Theory and Methodology Conference, Pittsburgh PA (2001)

**63** Ullman, D *The Mechanical Design Process*, 2nd ed. McGraw-Hill (1997)

**64** Dylla, N Thinking Methods and Procedures in Mechanical Design, Dissertation, Technical University of Munich, in German (1991).