Journal of Engineering Design
Vol. 23, No. 4, April 2012, 271–296

Investigating effects of oppositely related semantic stimuli on design concept creativity

I. Chiu[a] and L.H. Shu[b]*

[a]*Department of Mechanical and Industrial Engineering, Ryerson University, 350 Victoria Street, Toronto, ON, Canada M5B 2K3;* [b]*Department of Mechanical and Industrial Engineering, University of Toronto, 5 King's College Road, Toronto, ON, Canada M5S 3G8*

(*Received 25 February 2011; final version received 4 July 2011*)

We are motivated to investigate methods to increase creativity in conceptual design since creativity is essential to design success, and no other stage influences final design success as much as conceptual design. Existing work supports that design stimuli may encourage creative concept generation, but does not give details on how to systematically generate stimuli. The established relationship between language and cognition, and the systematic nature of language prompt us to examine its use as design stimuli. Language relationships such as opposition provide a systematic method of generating non-obvious semantic stimuli for design problems. In this paper, we present two experiments, a pen-and-paper and a verbal protocol study, where participants used oppositely related and similarly related word stimuli in conceptual design. We found that designers using oppositely related word stimuli developed more creative concepts. Language analysis revealed how opposite stimuli elicited designer behaviours that may encourage and support creative concept generation. Our empirical results combined with linguistic theory lead us to propose a model explaining the interactions and effects of opposite-stimulus words on concept creativity. This knowledge can be used to facilitate more creative, and ultimately, more successful design.

Keywords: language; creativity; design stimuli; concept generation; engineering design

1. Introduction

We have been studying and quantifying the application of natural language, i.e. human language, not artificial language, to the process of stimulating creative conceptual design. We are motivated to study language in relation to design because language appears ubiquitously, is inherent in people, and is connected to cognition. Researchers have established a link between language and cognition, although the exact relationship is disputed. Some have shown that language affects cognition (Levinson 1996), while others have shown that language reflects cognition (Pinker 2007). Design is a cognitively intensive activity requiring such tasks as information gathering, spatial manipulation, searching, decision-making, etc. (Simon 1969, Gero *et al.* 1994, Coley *et al.* 2007). The cognition required of design suggests that we can take advantage of the relationship between language and cognition to facilitate and understand design.

*Corresponding author. Email: shu@mie.utoronto.ca

ISSN 0954-4828 print/ISSN 1466-1837 online

http://dx.doi.org/10.1080/09544828.2011.603298
http://www.tandfonline.com

In this paper, we present and describe work on stimulating creative concept generation using language. Creativity is an important measure of design effectiveness (Kan and Gero 2007). While customers seek creative designs, they may not explicitly indicate creativity as a requirement (Cross 2006). There is general agreement that creativity is essential to design (Gordon 1961, de Bono 1992, Altshuller and Shulyek 1996, Hubka and Eder 1996, Adams 2001, Cross 2006, Kan and Gero 2007, Shai *et al.* 2009, Yang 2009, Brown 2008, etc.), and some assert that design necessarily entails creativity (Holt 1993, Hubka and Eder 1996, Hatchuel and Weil 2009). We are specifically interested in the concept generation stage of design because the early stages of design have the largest influence on the final design (Ullman 2003, Keller *et al.* 2009), and creative concepts result in creative and successful designs.

We investigate the effects of language stimuli on concept creativity in experiments, where we provided participants with stimulus words that were either oppositely or similarly related to the desired functions of the solution. Stimulus words were systematically generated, using a thesaurus and WordNet. Words were chosen as stimuli because words are the smallest unit of language that carry meaning, and thus a natural starting point for our investigations. Verbs were the chosen part of speech for the stimuli because verbs denote abstract actions or functions and not specific forms. Word stimuli provided were verbs in root form, e.g. ‘remove’. Many agree that verbs should be used to model functions for design (cf. Pahl and Beitz 1996, Stone and Wood 2000). The opposite/similar relationship was used to generate stimulus words because it is a systematic and well-understood relationship and one of only two valid verb relationships. While we may never fully understand the effects of language on design cognition, insights gained from our empirical results combined with linguistic theory enable us to propose a model of the language and design cognition interactions required to generate more creative concepts.

In this paper, we will first review related work in language, design and creativity. Then we will describe and discuss our experiments with respect to the specific effects of different types of word stimuli on concept creativity (Experiments 1 and 2) and on designer behaviour (Experiment 2). Finally, we will discuss our experiments within an empirical and theoretical context and propose a model explaining the effects of opposite/similar word stimuli on creative concept generation.

2. Background

2.1. *Language and design*

The connection between language and cognition was first observed by the ancient Greeks, who used the same word, *logos*, to denote both reasoning and language (Kalmar and Davidson 1997). The Sapir-Whorf hypothesis dating from the early twentieth century argues that the language of a person determines how he or she understands the world (Ratner and Gleason 1993). While there is debate over whether cognition influences language (Levinson 1996, Li and Gleitman 2002, etc.), or language influences cognition, many agree that language and cognition are closely related (e.g. Pinker, Fodor, Boaz from Saeed 2003). Researchers such as Chomsky (1968), Jackendoff (1983) and Pinker (2007) argue that language provides insight into human cognitive processes. In our work, we explore the use of language to stimulate creative design.

Language, specifically at the word level, appears ideal as design stimuli, and fits well within early stages of the design process, where the problem is likely to be ill-defined, and exploration of the solution space encouraged. Words are the smallest unit of language that carry meaning, and thus an appropriate starting point for investigating language as design stimuli. While words impose a pre-existing symbol system on the user that can be shared and manipulated (Bruner 1964), they are not necessarily fixed to a particular form (Segers 2004), and ambiguity enables freedom of interpretation (Miller *et al.* 1993). Not only is language connected to reasoning, words also appear connected to our knowledge of the world. In one model of lexical memory, when a specific word is found in the internal lexicon of permanent memory, simultaneously retrieved are properties associated with that word, e.g. meaning, spelling, pronunciation, etc. At the same time, properties not strictly linguistic are also retrieved, e.g. if the word 'elephant' is retrieved, it might also trigger the common knowledge that elephants never forget (Carroll 1999). The knowledge retrieved is not necessarily just part of the word-meaning, but may be related to the conceptual knowledge of the world in general. While words seem simple and familiar, word meaning can be complex and ambiguous (Carroll 1999). A designer's familiarity with the complexities of language may make language a useful design tool.

While natural language is familiar and ubiquitous to most, it is not generally considered a conventional engineering tool. However, many researchers are investigating the use of natural language as a formal tool to support the engineering design process. For example, language has been used as input in requirements gathering (Nuseibeh and Easterbrook 2000), concept generation and synthesis (Thomas and Carroll 1984, Hacco and Shu 2002, Nagai and Taura 2006, Chiu and Shu 2007, Tseng *et al.* 2008), design modelling and representation (Stone and Wood 2000), and design analysis (Dong *et al.* 2003, Dong 2006).

Some concept generation and creativity methods explicitly use language as stimuli in an attempt to increase creativity. Thomas and Carroll (1984) demonstrated that participants who were given 20 pages of semi-random stimulus words generated more creative concepts compared with participants not given stimulus words. Nagai and Taura (2006) investigated the interpretation of noun–noun combinations for promoting creativity in concept synthesis. Some work in biomimetic design uses functional keywords to systematically retrieve analogies from biological corpora for use as stimuli in engineering design (Hacco and Shu 2002, Chiu and Shu 2007, Cheong *et al.* 2011).

Methods implicitly using language to stimulate creative design include synectics and random input. In synectics, Gordon (1961) proposes the use of metaphors and similes, which are figures of speech, to draw parallels and connections between disparate topics or domains. Metaphors and similes promote analogical thinking and can allow solutions to be applied from one domain to another. The random input method involves randomly selecting stimulus, e.g. a picture from a catalogue, to relate back to the design problem through a series of word associations. The process of relating the problem to random stimulus, which may be non-obvious and unexpected, may provide new perspectives and thus stimulate creative design (de Bono 1970).

The above examples show that language is frequently used to facilitate design. However, fewer researchers have attempted to examine and model the effect of language on design. Nagai and Noguchi (2003) developed a model of the thinking path required for creative thinking. In an experiment where designers used keywords while designing, Nagai and Noguchi observed that difficult and remote keywords caused designers to extend their thinking pathways. They concluded that extending thinking pathways may help to realise creative concepts. In our experiments, we observed that oppositely related keywords result in designer behaviours that may increase concept creativity. We then propose a model to explain the effects of oppositely related stimuli on increasing concept creativity.

2.2. *Opposition in design*

We study opposite versus similar-language stimuli because opposition and opposing relationships are common in language and reasoning. Antonyms, or opposite words, are universally found in language, and most people demonstrate good intuition in recognising antonym/synonym pairs (Fellbaum 1993, Murphy 2003). The antonym/synonym relationship is also one of only two valid verb relationships, with the other being the hypernym/hyponym relationship. The hypernym/hyponym relationship can be thought of as a super-ordinate/sub-ordinate relationship, where words are hierarchically related in either a more general or more specific manner.

Design methods using opposition include TRIZ, design-by-analogy/contrast, argumentation and argumentative negotiation (Rittel and Webber 1984, Altshuller and Shulyak 1996, Fantoni *et al.* 2006, Jin *et al.* 2007). In TRIZ, the Russian abbreviation for the theory of inventive problem solving, the problem at hand is first phrased in contradictions to identify parameters to be improved and those degraded as a consequence. Fantoni *et al.* (2006) proposed a method of design-by-analogy/contrast involving the use of synonyms and antonyms as design stimuli. Argumentation and argumentative negotiation involve the verbalisation of contradictory demands and then a move towards agreement to produce novel solutions in collaborative engineering (Rittel and Webber 1984, Jin *et al.* 2007).

Hubka and Eder (1996) even speculate that opposites and dissimilarities may contribute to creativity in engineering design through the resolution of 'cognitive dissonance', e.g. resolving ideas from intuitive versus intellectual modes of thinking. Festinger, the originator of the theory of cognitive dissonance, theorised tensions occur when an individual becomes simultaneously aware of two inconsistent thoughts. To resolve the tensions, the individual must implement change (Myers 1999), thus creating new, alternative solutions. In design, resolving tensions between the problem statement and other factors, e.g. the degraded parameter versus the improved parameter in TRIZ, and stimuli oppositely related to the design problem, may lead to creative design.

2.3. *Measuring creativity in design*

Many approaches have been developed to assess the creativity of a concept or idea. In general, most agree that creativity is multi-dimensional (Torrance 1974, Amabile 1983, Shah *et al.* 2000, etc.) and that using a single measure of creativity may result in identifying strange or even incorrect ideas as being creative. For example, in Wilson's (1951) method of statistical infrequency, infrequent ideas are considered novel, and therefore creative. Two generally agreed upon measures of creativity within science and engineering are *novelty* and *usefulness*, defined below:

Novelty: A creative idea must contain some degree of newness, originality or surprise (Wilson 1951, Torrance 1974, Hubka and Eder 1996, Howard *et al.* 2008, Shai *et al.* 2009, Brown 2008).

Usefulness: A creative idea must contain some degree of appropriateness and value (Besemer and Treffinger 1981, Amabile 1989, de Bono 1992, Akin and Akin 1998, Howard *et al.* 2008, Shai *et al.* 2009, etc.). Amabile (1989) defines appropriateness in the sciences as being correct. Usefulness is especially emphasized in the engineering literature, often in the context of functionality (Pahl and Beitz 1996, Dieter 2000, Shah *et al.* 2000, Ullman 2003).

Some researchers assert that creative ideas only need to be novel and useful, e.g. Amabile (1989) and Akin and Akin (1998). However, novelty and usefulness alone may not sufficiently measure creativity, especially at the abstract conceptual level. Many suggest that the wholeness, clarity, elaboration, or *cohesiveness* of an idea must also be considered:

Cohesiveness: A creative idea must contain some degree of wholeness, elaboration, detail, style and clarity (Torrance 1974, Besemer and Treffinger 1981, Hubka and Eder 1996, Adams 2001, Kudrowitz and Wallace 2010).

Other measures of creativity may also be used to assess the effectiveness of a creativity method or individual. These measures can include fluency, or the quantity of concepts, generated by an individual or the creativity method (Torrance 1974, Shah *et al.* 2000, Yang 2009) and variety, or the number of different categories of concepts generated (Torrance 1974, Shah *et al.* 2000). This paper focuses on the use of the direct measures of novelty, usefulness and cohesiveness for the evaluation of each individual concept.

2.4. *Previous work*

Previous exploratory experiments showed that designers provided with opposite-stimulus words tend to generate concepts that were more novel. These previous experiments are summarised below.

The first exploratory experiment was a pen-and-paper study where participants indicated their concepts on worksheets. Forty-two participants were provided with opposite and similar words simultaneously for a series of four problems and then instructed to use the words as stimuli for design. For three of the problems, participants who chose at least one opposite-stimulus word generated more novel concepts (Chiu and Shu 2008a). Novelty was determined using statistical infrequency, where less frequently occurring concepts were deemed more novel (Wilson 1951, Shah *et al.* 2000). However, because all participants were provided with opposite- and similar-stimulus words *simultaneously*, results may have been confounded, i.e. were participants affected by only the opposite words, or by the pairs of opposite/similar words? From this experiment, we were also unable to gain further insight into *how* participants used language stimuli due to the nature of data collected from pen-and-paper experiments.

The second exploratory experiment was a small-scale between-subjects experiment where participants verbalised all thoughts as they designed. Six participants received either opposite words or similar words while generating concepts for one problem. Those receiving opposite stimuli generated more novel concepts, however the difference was not significant. Novelty in this experiment was determined by two independent human raters. Two raters were recruited because two is the minimum number of human raters required for judgement tasks (Landis and Koch 1977). Spearman's correlation between the two raters was $r = 0.51, p = 0.054 \sim 0.05$, showing a large, borderline significant agreement between the two raters (Chiu and Shu 2008b). No training was provided to the raters in this exploratory experiment.

Using the session transcripts from the second experiment, we were able to conduct an explicit content analysis to examine participants' language use and behaviours with respect to different stimulus types. Specifically, we examined the stimulus part-of-speech (POS), to determine whether a given stimulus word was used as a noun, verb or adjective. While stimulus words were given as root verbs, e.g. 'remove' and not 'removing' or 'removed', participants were not told that the stimuli were verbs. We found that opposite-stimulus participants used stimulus words significantly more often as verbs than similar-stimulus participants. Furthermore, in this second exploratory experiment, we found that opposite-stimulus participants tended to introduce more new words and phrases (NWPs) in their concept generation process, but not significantly so. NWPs are identified by comparing words and phrases given in a problem statement with the words and phrases generated by participants in relationship to the stimulus words. NWPs may represent new concept elements that have been expressed, or lexicalised, within concept generation. Increased NWP introduction appears linked to the use of stimuli as verbs rather than nouns. Increased NWPs appear advantageous for concept generation because NWPs may form the basis of creative concepts. An expanded verbal protocol experiment, i.e. in terms of participants and problems, would allow for more thorough analysis.

Overall, previous experiments showed that opposite stimuli appear to increase concept novelty, one of the measures that contribute to total concept creativity. These preliminary results serve as motivation for current experiments (described in Sections 3 and 4) that overcome limitations of previous work described above. On the basis of our preliminary results, we hypothesise the following:

(1) Opposite-stimulus words increase concept creativity in terms of all creativity measures;
(2) Opposite-stimulus words used as verbs increase the introduction of NWPs

The following sections will describe two experiments where we focus on further quantifying the above two hypotheses.

3. Experiment 1: a between-subjects pen-and-paper experiment

Experiment 1 was a fully between-subjects pen-and-paper experiment where participants generated concepts for a series of four problems: (1) Bushing problem (2) Egg problem (3) Grinding problem and (4) Sunflower problem. Participants generated concepts under one of the following experimental conditions:

(1) Opposite stimulus;
(2) Similar stimulus.

Four independent raters were recruited to judge the concepts based on the creativity components of novelty, usefulness and cohesiveness. For two of the experiment problems, the Sunflower and Egg problems, raters judged opposite-stimulus concepts as being significantly more novel, useful and cohesive than similar-stimulus concepts. For the other two problems, the Bushing and Grinding problems, there were no significant differences in any of the creativity component scores. Details are given below.

3.1. *Method*

3.1.1. *Participants and procedure*

Participants consisted of nine graduate engineering students from the Department of Mechanical and Industrial Engineering at the University of Toronto. At the time of the experiment, participants were enrolled in a graduate design course. Participants consisted of two females and seven males with an average age of 27.2 years (sd = 5.38). Eight out of nine participants indicated they had industry design experience ranging from a few months to five years.

Participants were randomly assigned to one of two experimental conditions where they generated concepts using either opposite stimuli or similar stimuli. All participants completed four problems, but in a different, random order as determined by a random number generator. Participants were allotted a total time of 10 min per problem to review given stimulus words; to select stimulus words; and to generate and describe their concepts using selected stimulus words on provided worksheets. The total experiment duration was 40 min. The worksheets were collected for analysis. The exact instructions are given below:

> This is an experiment investigating the use of stimuli in concept generation. The following are four unrelated design problems. For each design problem, a set of related word stimuli is supplied. For each problem:
> (1) Review the problem and associated word stimuli.
> (2) Perform a functional decomposition, e.g. what needs to be done?
> (3) Select the word(s) you want as stimuli and indicate your selection(s). Each stimulus set is only relevant to the associated design problem.
> (4) Use selected word(s) to develop concepts to solve the design problem. If you determine that you cannot complete a concept using your selected word(s), you may select another word.
> (5) Please consider each problem in the order given.
> You have 10 minutes for each problem. Record all concepts. You may write, sketch, calculate, etc., on your worksheets.

Pen-and-paper experiments are a fairly time-efficient method for collecting design data. Others have used pen-and-paper experiments to study various aspects of conceptual design such as effects of random stimulus words, analogical similarity, generation of form alternatives and sketching abilities (Thomas and Carroll 1984, Tseng *et al.* 2008, Yang 2009, Corremans 2011).

3.1.2. *Experiment problems and stimuli*

Participants were provided with four problems and related-stimulus sets on worksheets. Stimulus words were obtained by first performing a functional decomposition of the problem. Functional decomposition identifies the functions, i.e. physical actions or behaviours, required to transform an initial state into the desired final state (Pahl and Beitz 1996). A high-level functional decomposition yielded functional keywords that were then used to generate similar- or opposite-stimulus words using a thesaurus (Merriam-Webster.com 2008) and WordNet 3.0 (Princeton University 2008). Problems and stimuli are summarised in Table 1.

For the Sunflower problem, the required high-level functions to transform whole seeds into oil were 'extracting' the seeds from the shell, and then 'separating' the seeds from the shell fragments for the production of oil. From the original functional keyword 'extract', we generated using a thesaurus and WordNet similar words 'empty' and 'withdraw', and opposite words 'insert' and 'fill'. From 'separate', similar words generated were 'disconnect' and 'divide', and opposite words were 'join' and 'combine'. The original functional keywords were not provided to participants.

3.1.3. *Concept evaluation*

Concept evaluation for our previous studies (described in Section 2.4) was limited to the evaluation of concept novelty, using statistical infrequency or human raters. Since the creativity literature suggests that creativity is more than merely novelty, and that creativity is ultimately a human judgment, e.g. from a customer's point of view (Cross 2006, Brown 2008), we recruited human

Table 1. Summary of problems and stimulus sets for Experiment 1.

		Stimulus words	
Problem	Problem description and decomposition	Similar	Opposite
Sunflower-seed shelling	Sunflower seed oil is a nutritious and valuable commodity in sub-Saharan West Africa. Mechanical presses to make oil from the shelled seeds exist locally, but machines to remove the shells do not. At present, there exists no alternative to the laborious and time-consuming process of shelling the sunflower seeds individually by hand before loading them into the press. Develop a concept for shelling sunflower seeds that can be used locally with minimal resources (DTM 2006)	Empty Withdraw Disconnect Divide	Fill Insert Join Combine
Soft-material grinding	Grinding of metals is quite common to obtain a fine surface finish and tight tolerances. But when grinding soft materials such as rubber or plastic, the grinding wheels quickly become clogged. Repeated dressings (sharpening and shaping of the grinding wheel) do not help. Develop concepts that will enable surface finishing (with or without grinding wheels) to be used on soft materials (Kosse 2004)	Smooth Subtract Clean Remove	Roughen Add Clog Replace
Egg orientation	Develop concepts to automatically orient raw chicken eggs with the pointed ends all facing one direction (Kosse 2004)	Select Detect Pivot Move	Reject Miss Fix Restrain
Bushing-and-pin assembly	Parts that are automatically mated, e.g. a bushing and a pin, must be positioned so that their axes coincide. Using chamfers on mating parts does not solve the alignment problem. Develop a concept to centre mating parts that does not require high positioning accuracy (Kosse 2004)	Straighten Match Inject Install	Skew Mix Eject Extract

Low novelty				Average novelty			High novelty			
0	1	2	3	4	5	6	7	8	9	10
Not novel										**Very novel**

Figure 1. Rating scale for novelty. Reprinted with permission. Copyright ASME 2010.

Table 2. Example anchor concepts for Sunflower problem.

Low/high	Novelty	Usefulness	Cohesiveness
Low	Place whole seeds in machine. Divide the shell by machine with chisel/knife edge and open like hinge securing shell. Empty the seeds from the shell and discard the shells from the hinge device	Machine grips two sides of seed. Seed is cut into half (divide). The seed is emptied using vibration (empty). Shells are discarded. Separate seeds processed as before	Divide the shell into two. Empty shell and remove seed from shell
High	Combine two wheels with sticky surfaces to crack the seeds. Shells stick to the surface and get carried away	Whole seeds go into a hopper. Use edible solvent to dissolve shells. Clean seeds. Dry seeds. Grind to get oil	Refer to high usefulness concept. High usefulness concept is also a high cohesiveness concept

raters to judge concept creativity for this study. Human raters assessed concept creativity using all three components described previously: *Novelty, usefulness* and *cohesiveness*. Each component was scored on an 11-point scale between 0 and 10. The rating scale for novelty is illustrated in Figure 1, with the same numeric scale applied to usefulness and cohesiveness.

For this experiment, four independent raters were recruited to score each concept. Raters consisted of one female and three males, all with an engineering background and interest in design. Raters were not provided with the identity of the designers, nor the stimulus condition under which the concepts were generated. Concepts were presented to raters in random order.

Raters were provided with low anchor and high anchor concepts, i.e. 'not novel/useful/cohesive' and 'very novel/useful/cohesive' concepts obtained from previous experiments, and instructed to evaluate all concepts based on those anchors. This method of anchoring and obtaining human judgements is known as direct scaling and is commonly used in psychophysics, a branch of psychology that deals with relating physical stimuli with mental phenomena (Engen 1972). The direct scaling method has been applied to rating pleasantness of smells, perception of heaviness, and even to rating of emotions and beauty. Examples of anchor concepts for the Sunflower problem are provided in Table 2.

3.2. *Results*

3.2.1. *Inter-rater agreement*

Kendall's W concordance coefficient was calculated to measure inter-rater agreement. Kendall's W measures the agreement of more than two raters scoring N entities and ranges from 0, signifying no agreement or random ratings, to 1, signifying consensus between the multiple raters. As Kendall's W is related to Spearman's correlational coefficient, Kendall's W can be interpreted similarly to a correlation coefficient, where 0.1, 0.3 and 0.5 are small, medium and large agreements, respectively (Siegal 1956). The inter-rater agreements are given in Table 3.

Although the coefficients above suggest medium to large degrees of agreement, the sample sizes are too small to draw confident conclusions ($p \gg 0.05$). Upon aggregating ratings for all concepts from all four problems, the coefficients indicate medium to large degrees of agreement that are significant, or borderline significant, as seen in Table 4.

Table 3. Inter-rater agreement for Experiment 1 concepts.

Problem	Novelty	Usefulness	Cohesiveness
Bushing and pin ($N = 9$)	$W = 0.3711$ $\chi^2(8) = 8.91$ $p = 0.35$	$W = 0.5482$ $\chi^2(8) = 13.16$ $p = 0.11$	$W = 0.6181$ $\chi^2(8) = 14.83$ $p = 0.06$
Egg ($N = 15$)	$W = 0.5531$ $\chi^2(14) = 23.23$ $p = 0.06$	$W = 0.6502$ $\chi^2(14) = 27.31$ $p = 0.02$	$W = 0.5278$ $\chi^2(14) = 22.17$ $p = 0.08$
Grinding ($N = 16$)	$W = 0.6004$ $\chi^2(15) = 27.02$ $p = 0.03$	$W = 0.5024$ $\chi^2(15) = 22.61$ $p = 0.09$	$W = 0.4634$ $\chi^2(15) = 20.85$ $p = 0.14$
Sunflower ($N = 13$)	$W = 0.4839$ $\chi^2(12) = 17.42$ $p = 0.13$	$W = 0.3261$ $\chi^2(12) = 11.74$ $p = 0.47$	$W = 0.497$ $\chi^2(12) = 17.89$ $p = 0.12$

Table 4. Kendall's W for aggregated results of Experiment 1 concepts.

Problem	Novelty	Usefulness	Cohesiveness
Aggregated results ($N = 53$)	$W = 0.507$ $\chi^2(52) = 79.091$ $p = 0.009$	$W = 0.4426$ $\chi^2(52) = 69.052$ $p = 0.06$	$W = 0.4908$ $\chi^2(52) = 76.5592$ $p = 0.015$

Therefore, based on the aggregated results, there is significant agreement among the raters (p approximately 0.05 or less).

3.2.2. *Creativity results*

In this section, we will first present examples of participant concepts and resulting rater scores to illustrate the experimental and concept rating methodology. Then, the overall results of the quantitative comparisons are provided.

Three concept examples from the Sunflower problem are given in Table 5 along with associated average rater scores. Concepts are unedited and as collected directly from the worksheets. Stimulus words used by participants are italicised for reference.

Table 5. Example concepts and creativity scores for the Sunflower problem.

Example Concept	Participant experimental condition	Avg. rater score ($N = 4$)		
		Nov.	Use.	Coh.
Concept 2: To *divide*: Orient the seed so that it lies flat, then use a knife to split along the flat side of the shell. Or cut into half, then use vibration or gravity to *empty*, blow it out to extract inner part	Similar	2	2.5	2.25
Concept 13: (1) Air hose *insert* into seed. (2) Air *fills* the seed, cracking the shell using pressure. (3) The newly cracked seed is collected and *joined* with the rest of the seeds from before	Opposite	8	4.5	5.25
Concept 16: Second concept – Grind shell and seed together. Then find a fluid, (hopefully water will do) that has the proper density to float the lighter shells to the surface and leave the heavier seeds on the bottom	Opposite	5.25	7.5	7.5
Average concept creativity for all Sunflower concepts ($N = 13$)		4.35	4.88	5.0

Concept 2 was generated under the similar-stimulus condition, while concepts 13 and 16 were generated under the opposite-stimulus condition. Note that concept 16 does not incorporate stimulus words. However, the participant had indicated that this was his second concept for this problem, and stimulus words were used in his first concept for this problem; demonstrating he had reviewed the stimulus words. Concept 2 was rated as the least creative concept, and concepts 13 and 16 were rated as being above-average concepts. While raters judged concept 13 to be the most novel, concept 16 was rated as more useful and cohesive. This may reflect that raters considered batch processing of multiple seeds simultaneously (concept 16) to be more practical than processing of individual seeds (concept 13 and concept 2).

For each problem, rater scores for concepts were averaged to facilitate analysis, and independent T-tests were conducted to compare concept novelty, usefulness and cohesiveness. Rater scores were averaged because our main interest is in designer behaviour, not rater behaviour. For two of the experiment problems, the Sunflower and Egg problems, raters judged opposite-stimulus concepts as being more novel, useful and cohesive than similar-stimulus concepts. The T-tests showed that these differences were either significant, or borderline significant, in all but the Egg problem novelty scores. This can be seen in Figure 2 and Table 6.

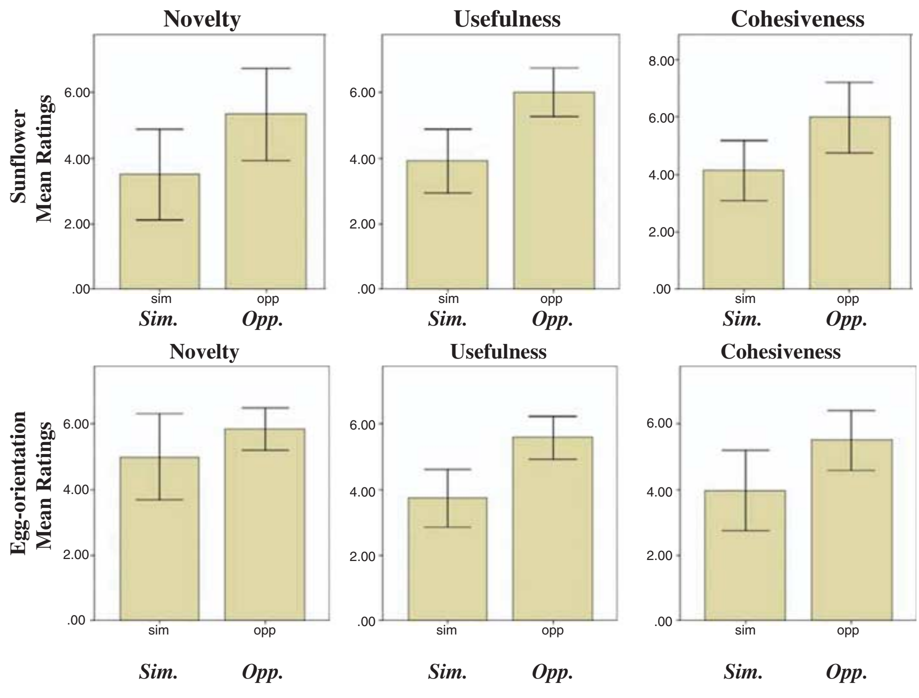

Figure 2. Graph of novelty, usefulness and cohesiveness results for the Sunflower and Egg problems.

Table 6. T-test results for Sunflower and Egg concept creativity.

	Sunflower				Egg			
Creativity Component	Sim. Mean Rating ($N = 7$)	Opp. Mean Rating ($N = 6$)	$t(11)$	p(one-tail)	Sim. Mean Rating ($N = 9$)	Opp. Mean Rating ($N = 6$)	$t(13)$	p(one-tail)
Novelty	3.51	5.33	−1.80	0.048	5.00	5.83	−0.96	0.16
Usefulness	3.93	6.00	−3.32	0.004	3.75	5.58	−2.98	0.006
Cohesiveness	4.15	5.97	−2.25	0.02	3.97	5.50	−1.80	0.048

For the other two problems, the Bushing and Grinding problems, there were no significant differences in the creativity component scores, and thus different stimuli had no significant effects.

3.2.3. *Language results*

While pen-and-paper experiments are not the most conducive to examining language use, we were able to examine aggregated frequency of stimulus use by counting and categorising stimulus words used by the participants. Many instances of stimulus words found on worksheets were merely listed and not used in a phrase or sentence. When these ‘unknown’ uses of stimuli are subtracted from the frequency totals, opposite-stimulus concepts incorporate significantly more stimulus words, $t(51) = 2.791, p = 0.0035 < 0.05$, than similar-stimulus concepts. This can be seen in Figure 3, and suggests that using more opposite-stimulus words may result in more creative concepts.

3.3. *Experiment 1 discussion*

There was strong inter-rater agreement with regard to aggregated concept creativity measures. This is illustrated by the below-average usefulness scores for concepts involving individual seed processing (concepts 2 and 13), versus a higher score for concepts involving batch processing of multiple seeds (concept 16). This may reflect that batch processing is seemingly more practical. However, it should be noted that strong agreement does not necessarily mean that the raters were correct (Siegal 1956), nor does it necessarily account for other factors which may have influenced concept creativity. Despite any drawbacks associated with human ratings of creativity, creativity is a human judgment, and what is considered creative is not necessarily ‘correct’ or ‘incorrect’.

For two of the problems, the Sunflower and Egg problems, opposite-stimulus concepts were significantly more creative than similar-stimulus concepts. However, for the Bushing and Grinding problems, opposite- and similar-stimulus concepts were found to be equally creative. Results may reflect the difference in problem type and problem novelty. The Sunflower and Egg problems are general-domain problems that may be more novel to the participants, while the Bushing and Grinding problems are more technically oriented and should be familiar to most engineers, i.e. participants *and* raters. The results shown above suggest that opposite-stimulus words can

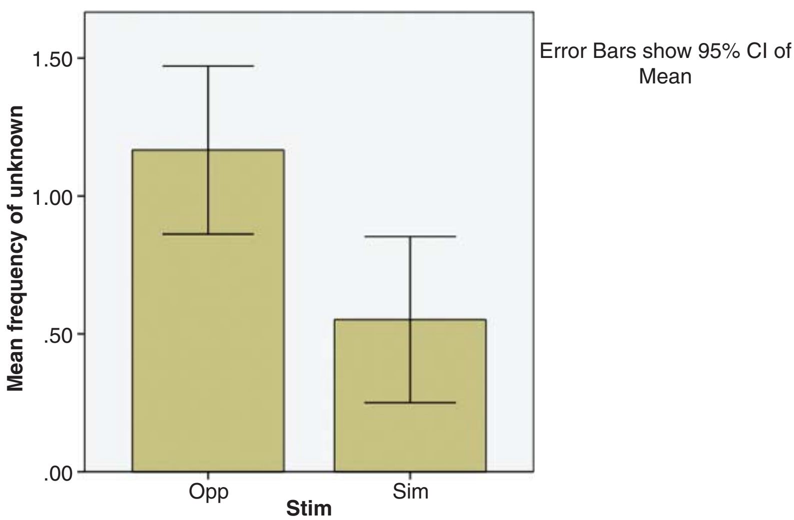

Figure 3. Aggregated frequency of stimulus word-use with no unknown uses of stimulus words.

stimulate more creative concepts for general-domain problems or problems that may be more novel to the designers. In cases where problems are more technically oriented, experience may neutralise the effects of stimuli.

Both rater and participant experience affect creativity; experience is an important factor in individual creativity (Amabile 1989, Akin 1990). It is difficult to determine if participant concepts are based on designs encountered elsewhere, and difficult to separate historically creative concepts, concepts that never existed before in the world, from personally creative concepts, concepts that never existed before in the participant's mind, but existing in the world (Amabile 1983). In addition, a more experienced rater may tend to judge a participant's personally creative concepts as less creative even though they are creative and novel in the context of the experiment. A less experienced rater may judge the same concept as more creative. This may have been the case with the technically oriented problems, where both participants and raters should be familiar with a variety of alignment/insertion and material removal solutions. Raters may also judge concepts to be less creative if the participant explicitly refers to existing designs, even if concepts appear creative. Overall, it would be difficult to control for participant and rater experience.

Regarding stimulus-word use, results show that opposite-stimulus concepts incorporated more stimulus words within the concept. This supports that opposite-stimulus participants were better able to use stimuli to introduce and develop new ideas in concept generation, which suggests that opposite-stimulus words are more useful for stimulating creative concepts. A more comprehensive verbal protocol, or talk-out-loud, experiment may provide further insight into how opposite-stimulus words increase concept creativity. The next section describes such an experiment.

4. Experiment 2: An expanded verbal protocol experiment

Experiment 2 is a verbal protocol experiment involving 14 participants generating concepts for three problems under one of the following experimental conditions:

(1) No stimulus (control);
(2) Similar stimulus;
(3) Opposite stimulus.

Language use in relation to participant behaviour, as well as concept creativity, is examined in detail in Experiment 2.

4.1. *Method*

4.1.1. *Participants and procedure*

All 14 participants are fluent English speakers recruited from the Department of Mechanical and Industrial Engineering at the University of Toronto. Participants consisted of 13 males and one female, ranging from fourth-year undergraduate students to second-year PhD students.

In individual sessions, participants first completed three training problems to habituate them to verbalising. Then, participants were instructed to verbalise all thoughts as they completed a series of three design problems. Fifteen minutes were allotted for each problem for a total experiment duration of 45 min. Ten participants were provided with stimulus words, either opposite or similar words, while four were not provided with any stimuli. Of the 10 stimulus participants, five switched stimulus type between problems. Table 7 details the experimental design.

Similar to the previous pen-and-paper experiment, participants in verbal protocol experiments were provided with worksheets presenting the problem statements and stimulus words. In contrast to pen-and-paper experiments, participants in verbal protocol experiments were also instructed to

Table 7. Experimental design for Experiment 2.

	Stimulus participants										Control participants				Condition subtotals		
Prob.	TH	JS	VT	SW	JL	DRO	DR	UG	MM	DH	DL	JM	AF	AP	Opp.	Sim.	None
Bushing	S	S	S	S	S	S	O	O	O	O	N	N	N	N	4	6	4
Snow	O	O	O	S	S	S	S	O	S	O	N	N	N	N	5	5	4
Coal	S	S	S	S	S	S	O	O	O	O	N	N	N	N	6	4	4

Notes: Reprinted with permission. Copyright ASME 2010. S: Similar stimuli; O: Opposite stimuli; N: No stimuli.

verbalise all thoughts as they worked on the design tasks. The sessions were recorded and fully transcribed for analysis.

Verbal protocols are common for studying cognitive processes such as human machine interaction (Bainbridge *et al.* 1968), medical decision-making (Lutfey *et al.* 2008) and are considered a relatively objective and appropriate method for studying phenomena in design (Cross *et al.* 1996, Hubka and Eder 1996, McNeill *et al.* 1998, Chakrabarti *et al.* 2004, Cross 2006, Visser 2006). In fact, verbalisation may be the most popular method for studying design cognition (Coley *et al.* 2007). However, there are some debates associated with verbal protocol studies. Nisbett and Wilson (1977) have questioned the accuracy of the data obtained from verbalisations as they have found that verbal reports do not necessarily match recordings of the reported event. Ericsson and Simon (1993), on the other hand, contend that as long as verbalisations are immediate and do not require recall from memory, verbalisations accurately describe events being reported. Additionally, since verbal protocol experiments are a time and resource intensive method, the number of participants involved is usually small. A survey of design studies using this method reveals that the typical number of participants is low, e.g. 4 participants in a design cognition modelling study (Benami and Jin 2002), 8 participants in a personal creativity and design activities study (Kim *et al.* 2011), 10 participants in a design education study (Atman and Bursic 1996), and 20 participants in a design stimulation study (Jin and Benami 2010). Our sample size of 14 is reasonable considering the range of sample sizes typical of verbal protocol experiments.

4.1.2. *Problems and stimuli*

Participants were provided with three problems and related stimulus sets on worksheets. Specifically, the problems were (1) Bushing-and-pin assembly (2) Snow insulation of houses and (3) Coal storage. Again, stimulus sets for the opposite and similar stimulus conditions were verbs in the root form generated by using a combination of a thesaurus (Merriam-Webster.com 2008) and WordNet 3.0 (Princeton University 2008), starting from the original functional keywords. Some keywords do not have antonyms in the resources consulted, e.g. 'to insulate' from the snow insulation problem, so opposite stimuli were generated based on opposition to the problem itself, e.g. 'to pack', as the problem specifically stated that 'packing' of snow is undesirable. As generating opposite and similar verbs is not possible for all keywords, and oppositely related words are sparse for verbs to start with, this strategy was used for the other problems as well. Problems and stimulus sets are given in Table 8.

4.1.3. *Concept identification and evaluation*

To reduce bias, design sessions were transcribed by an independent transcriptionist and an independent concept reviewer was recruited to identify and code concepts from the free-form transcripts and worksheets. The following is a transcript excerpt representing approximately 30 s of one experiment session. Lines are numbered for referencing during analysis.

Table 8. Summary of problems and stimulus sets for Experiment 2.

Prob.	Problem description	Opposite stimulus words	Similar stimulus words
Bushing-and-pin assembly	Parts that are automatically mated, e.g. a bushing and a pin, must be positioned so that their axes coincide. Using chamfers on mating parts does not solve the alignment problem. Develop a concept to centre mating parts that does not require high positioning accuracy (Kosse 2004)	Original keywords: Opposite of align and insert Stimulus words: Change, disorder, disarrange, scramble, randomise, misalign, tumble, skew, move, expel, pull, eject, evict	Original keywords: align and insert Stimulus words: Inject, transplant, sandwich, connect, skew, mount, misalign, attach, join, reorient, adjust, modify, match
Snow	In Canada, snow is readily available in winter and has good insulating qualities due to the amount of air in it. However, if the snow is packed to the point, it becomes ice, it is less insulating due to the loss of air. Come up with a concept to enable snow to be used as an additional layer of insulation for houses in the winter	Original keywords: pack and compact Stimulus words: Arrange, bundle, change, compress, constrict, contract, force, impact, move, push, squeeze, tighten, wad	Original keywords: insulate and surround Stimulus words: Blanket, control, cover, defend, enclose, immerse, pack, preserve, prevent, restrain, restrict, submerge, touch
Coal	Clean coal and clean coal combustion technologies make it possible to generate cleaner electricity. That, combined with the increasing cost of oil and natural gas, power plant operators may consider converting or reconverting their power plants from oil or natural gas back to coal. However, there may not be enough land area near the plant that can be used for on-the-ground coal storage. Propose alternative solutions to a conventional coal pile (adapted from Dieter 2000)	Original keyword: opposite of store Stimulus words: Abandon, discard, discharge, dispense, disperse, dispose, distribute, export, remove, scatter, spread, waste	Original keyword: store Stimulus words: Accumulate, collect, displace, distribute, feed, give, heap, keep, place, supply, transfer, withhold

(1) *I think the obvious …*
(2) *the first thing that comes to mind is that you'd like blanket the house … uh*
(3) *essentially blanket the house in a layer, in a thin layer of snow …um*
(4) *If the snow is packed to the point that it becomes ice …*
(5) *I guess, you'd obviously try to figure out what amount of packing you'd have to do*
(6) *to restrict the snow from becoming ice due to over packing.*

Finished transcripts were corrected for minor spelling errors, e.g. 'chamfer' for 'camphor', 'pedal' for 'petal', but were otherwise not annotated nor changed.

Concepts were identified by reviewing transcripts and worksheets. Concepts identified by the independent reviewer were compared with the concepts identified by the investigators who also reviewed all worksheets and transcripts. In cases where there was disagreement between the independent reviewer and investigators, both identified concepts were added to the set of concepts. For example, in a transcript segment from the Coal problem, both the independent reviewer and the investigators identified an 'underground storage concept' in which the coal would be stored underground. However, only the independent reviewer also identified a 'storage pile' as a concept. The investigators did not regard the 'storage pile' as a concept because the problem statement specifically required concepts other than a 'conventional coal pile'. Although the investigators disagreed with the independent reviewer with regard to the identification of the 'storage pile' as a concept, the 'storage pile' concept was added to the set of concepts to be evaluated by the raters. This concept identification process helped to ensure that all possible concepts were included for evaluation.

The independent reviewer also summarised participants' instances of similar concepts into a single concept type. For example, in one transcript segment for the Coal problem, multiple references to the 'tower concept' or 'condominium concept' are in fact only one concept that involved storing coal in a tall structure. A total of 195 concepts were identified between all participants and problems, 59, 59 and 77 concepts for the Bushing, Snow and Coal problems, respectively.

Concept creativity was evaluated using the anchoring and direct-scaling method developed and described in Experiment 1. Three raters were recruited and consisted of one female and two males, all with knowledge and interest in engineering design. Raters were not provided with the identity of the designers, nor the stimulus condition under which the concepts were generated. Concepts were presented to the raters in random order.

4.1.4. *Analysis*

First, inter-rater agreeability was calculated using Kendall's W. Then, rater scores for each concept were averaged, and all concept scores for the same participant were aggregated to facilitate analysis. This produces an aggregated *novelty, usefulness and cohesiveness* score for each of the 14 participants. Aggregated scores were analysed using a mixed-model analysis of variance (ANOVA). As in Experiment 1, rater scores were averaged to facilitate analysis because our main interest is not in rater behaviour, but in participant behaviour. Because five participants switched stimulus type between problems during the experiment (identified as TH, JS, VT, MM and DR in Table 7), pseudo-replicates were created to model these participants as independently contributing to each stimulus condition, effectively increasing the sample size to 19 from 14. This is a common technique to deal with scenarios where not all participants contribute independently to only one experimental condition over multiple trials. Generally, the use of pseudo-replicates results in a conservative estimate of differences (L. Duquette, personal communication, 2009).

4.2. *Results*

4.2.1. *Inter-rater agreement*

Kendall's W was calculated for each of the problems to examine inter-rater agreement and shown in Table 9.

All values for W show a medium to large agreement between the raters, and are statistically significant, $p < 0.05$. Therefore, there is significant agreement between the raters.

4.2.2. *Creativity results*

Overall, raters judged opposite-stimulus concepts to be more creative than similar-stimulus concepts. This difference was significant, or borderline significant, for all three creativity metrics. See

Table 9. Kendall's W for Experiment 2 concepts.

Problem	Novelty	Usefulness	Cohesiveness
Bushing ($N = 59$)	$W = 0.62$ $\chi^2(58) = 143.61$ $p < 0.0001$	$W = 0.53$ $\chi^2(58) = 122.76$ $p < 0.0001$	$W = 0.57$ $\chi^2(58) = 131.28$ $p < 0.0001$
Snow ($N = 59$)	$W = 0.55$ $\chi^2(58) = 127.25$ $p < 0.0001$	$W = 0.60$ $\chi^2(58) = 138.49$ $p < 0.0001$	$W = 0.65$ $\chi^2(58) = 150.17$ $p < 0.0001$
Coal ($N = 77$)	$W = 0.51$ $\chi^2(76) = 153.99$ $p < 0.0001$	$W = 0.41$ $\chi^2(76) = 126.13$ $p = 0.0003$	$W = 0.47$ $\chi^2(76) = 143.68$ $p < 0.0001$

Table 10. Planned contrast results for concept novelty, usefulness and cohesiveness.

	Comparison of estimated means		
	Expt. Cond. 1	Expt. Cond. 2	Contrast t- and p-values
Nov.	Opp: 4.09	None: 4.26	$t(28.456) = 0.32, p = 0.378, p > 0.05$
	Opp: 4.09	Sim: 2.88	$t(27.65) = -3.02, p = 0.0025, p < 0.05$
Use.	Opp: 3.75	None: 3.47	$t(31.51) = -0.73, p = 0.24, p > 0.05$
	Opp: 3.75	Sim: 3.17	$t(31.51) = -1.61, p = 0.059, p \sim 0.05$
Coh.	Opp: 4.09	None: 3.74	$t(30.16) = -0.92, p = 0.18$[a]$, p > 0.05$
	Opp: 4.09	Sim: 3.52	$t(29.42) = -1.62, p = 0.058$[a]$, p \sim 0.05$

Notes: Reprinted with permission. Copyright ASME 2010. Shaded rows show significant or borderline significant differences.
[a]Adjusted for effects of problem order.

Table 10 and Figure 4 for planned contrasts and interaction graphs of all creativity results. Note the lines in the interaction graph do not signify a relationship, but assist in visualising variable effects and interactions.

For novelty, a significant main effect was found for Stimulus Type $F(2, 27.58) = 7.09$, $p = 0.003$, $p < 0.05$. Planned contrasts comparing individual experiment conditions, e.g. opposite-stimulus concepts versus no-stimulus concepts, show a significant difference between opposite-stimulus and similar-stimulus concepts, $t(27.65) = -3.02$, $p = 0.0025$, $p < 0.05$. For usefulness and cohesiveness, planned contrasts show that opposite-stimulus concepts are borderline significantly more useful and cohesive than similar-stimulus concepts, $t(31.51) = -1.61$, $p = 0.059$, $p \sim 0.05$ and $t(29.42) = -1.62$, $p = 0.058$, $p \sim 0.05$, respectively. Problem order, the order in which problems were completed, was found to have an effect on cohesiveness and was corrected in the planned contrasts.

For novelty, usefulness and cohesiveness, planned contrasts show no significant difference between opposite-stimulus and no-stimulus (control) concepts.

Overall, the ANOVAs and planned contrasts support the original hypothesis that opposite-stimulus concepts are more novel, useful and cohesive than similar-stimulus concepts. However, we also observe two results that contradict much of the literature with regard to design stimulation:

(1) Opposite-stimulus and no-stimulus concepts were found to be equally creative;
(2) No-stimulus (control) concepts were found to be more creative than similar-stimulus concepts.

These results will be further discussed in a later section.

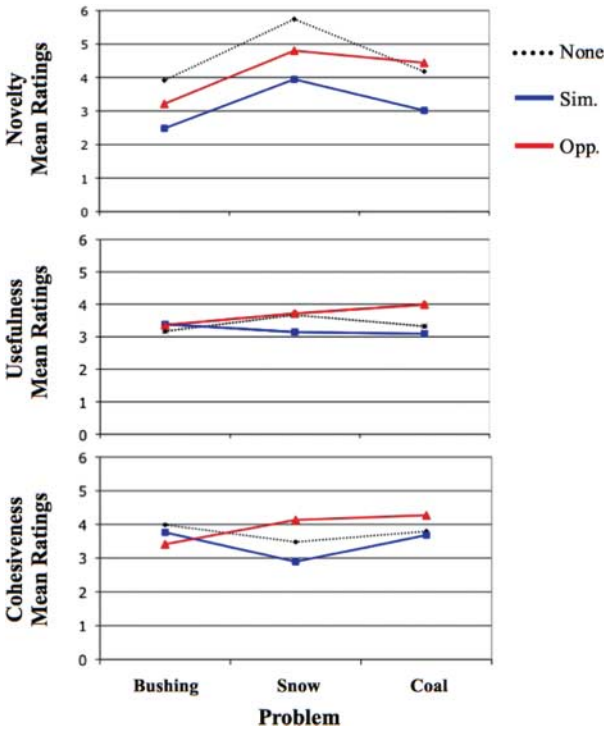

Figure 4. Graphs of concept novelty, usefulness and cohesiveness ratings. Note the lines in the interaction graph do not signify a relationship, but assist in visualising variable effects and interactions. Reprinted with permission. Copyright ASME 2010.

4.2.3. *Language results*

An explicit content analysis was performed using experiment transcripts to determine how stimulus participants used stimulus words in the concept generation process. A parameter of interest is the POS in which the stimulus words were used, and therefore, stimulus words used by participants were categorised as a noun, verb or adjective. Another parameter of interest corresponds to the NWPs introduced by stimulus words. NWPs are of interest because they may indicate that new concept elements were being introduced into the concept generation process. In turn, new concept elements may form the basis of creative concepts. The investigators were able to identify NWPs by comparing words and phrases in the problems and stimulus sets with the words and phrases generated by the participants. In an example from the Snow problem, given the stimulus word 'constrict', sentence A below contains an NWP (underlined), while sentence B does not contain an NWP.

(A) *Constrict* the <u>motion of heat</u> [from leaving the house]
(B) …*constrict* snow.

In sentence A, the phrase 'motion of heat', associated with the stimulus word 'constrict', was not given as part of the design problem nor stimulus set. However in sentence B, the word 'snow', also associated with the stimulus word 'constrict', was given as part of the design problem.

Overall, results suggest (1) verbs introduce more NWPs and (2) opposite-stimulus verbs introduce the most NWPs for two of the problems, the Bushing and Snow problems. This can be seen in the interaction graphs in Figures 5–7.

For each problem, a two-way ANOVA was used to compare the effect of Stimulus POS (verb, noun or adjective) and Stimulus Type (opposite or similar stimuli) on NWPs. First, the ANOVA is

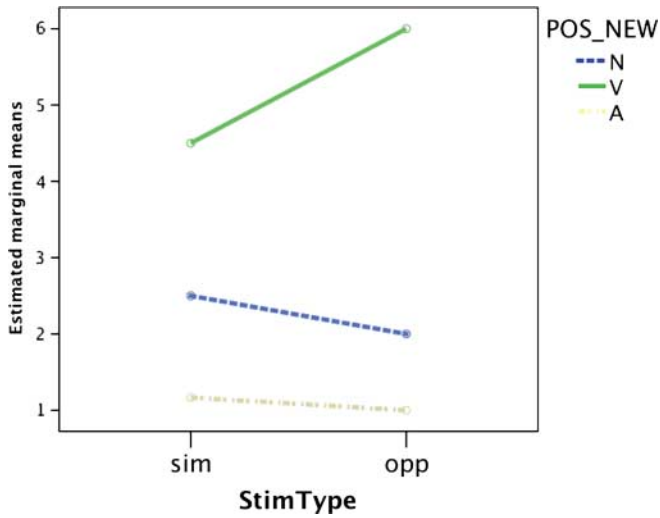

Figure 5. Bushing language results comparing effects of POS and Stimulus Type on estimated marginal means of NWP.

used to determine if there are Stimulus POS effects on NWP introduction. We found that Stimulus POS has a significant effect on the introduction of NWPs, $F(2, 16) = 7.28, p = 0.006, F(2, 16) = 40.42$, $p = 0.000$ and $F(2, 16) = 10.24$, $p = 0.001$, for the Bushing, Snow and Coal problems respectively. For all three problems, planned contrasts show that stimuli used as verbs introduce significantly more NWPs, or borderline significantly more NWPs, as seen in Tables 11–13.

Second, the ANOVA is used to determine if there are Stimulus Type and Stimulus Type*POS interaction effects on NWPs. For the Bushing problem, Figure 5 suggests that opposite stimuli introduce the most NWPs. However, there is no significant effect of Stimulus Type on NWPs nor Stimulus Type*POS interaction, meaning that one Stimulus Type POS, e.g. opposite-stimulus verbs, does not significantly introduce more or fewer NWPs.

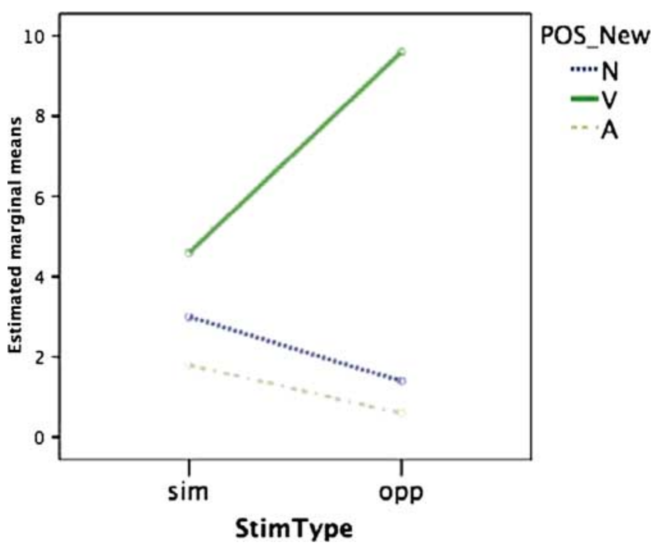

Figure 6. Snow language results comparing effects of POS and Stimulus Type on estimated marginal means of NWP.

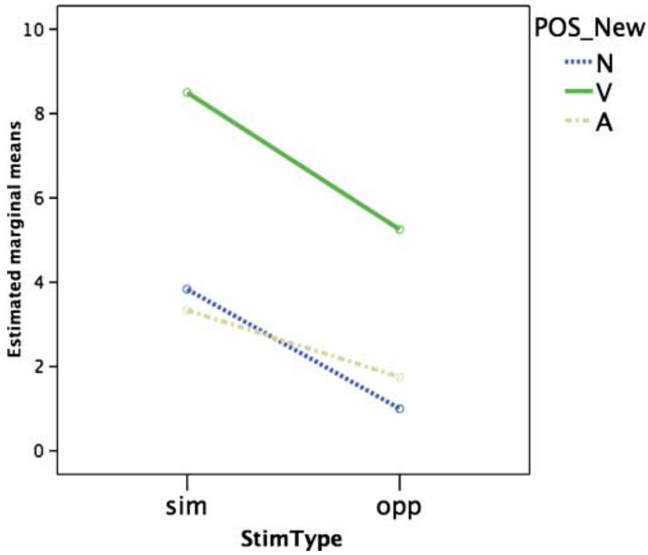

Figure 7. Coal language results comparing effects of POS and Stimulus Type on estimated marginal means of NWP.

Table 11. Bushing problem contrast results.

Mean NWP introduction		F- and p-values
Verb: 5.10	Noun: 2.30	$F(1, 8) = 4.61, p = 0.064–0.05$
Adj: 1.10	Noun: 2.30	$F(1, 8) = 2.80, p = 0.13$

Table 12. Snow problem contrast results.

Mean NWP introduction		F- and p-values
Verb: 7.10	Noun: 2.20	$F(1, 8) = 41.76, p = 0.000 < 0.05$
Adj: 1.20	Noun: 2.20	$F(1, 8) = 2.90, p = 0.13$

Table 13. Coal problem contrast results.

Mean NWP introduction		F- and p-values
Verb: 7.20	Noun: 2.70	$F(1, 8) = 47.58, p = 0.000 < 0.05$
Adj: 2.70	Noun: 2.70	$F(1, 8) = 2.90, p = 0.13$

For the Snow problem, Figure 6 suggests that opposite stimuli introduce the most NWPs. There is a significant effect of Stimulus Type on the introduction of NWPs, $F(1, 8) = 5.26, p = 0.051 \sim 0.05$. Additionally, there is a significant interaction of Stimulus Type*POS, $F(2, 16) = 13.88$, $p = 0.000$. Contrasts for Stimulus Type*POS interaction comparing opposite-stimulus verbs to other POS show significant differences, $F(1, 8) = 18.94$, $p = 0.002$, indicating that different stimulus POS have different effects depending on the Stimulus Type. In this case, it indicates that opposite-stimulus verbs introduce significantly more NWPs.

For the Coal problem, unlike the Bushing and Snow problems, Figure 7 suggests that similar stimuli introduce more NWPs rather than opposite stimuli. The ANOVA shows a significant effect of Stimulus Type on NWP introduction, $F(1, 8) = 16.76, p = 0.003$, while there is no significant interaction of Stimulus Type*POS. These results are discussed in detail in the next section.

4.3. *Experiment 2 discussion*

Experiment 2 shows opposite-stimulus concepts were significantly, or borderline significantly, more novel, useful and cohesive than similar-stimulus concepts. These results support our original hypothesis that opposite stimuli result in more creative concepts. These results are consistent with previous results obtained from Experiment 1. However, surprisingly, Experiment 2 also shows that stimulus concepts in general were *not* more creative than no-stimulus (control) concepts. This is contrary to the literature and results of other design stimulation studies.

It is unclear why no-stimulus concepts were judged more creative than some stimulus concepts, specifically similar-stimulus concepts. Further work is required to explain the discrepancy between results found in this experiment, and results reported by others. However, we have noted that many other design stimulation experiments were pen-and-paper, e.g. Tseng *et al.* (2008). The discrepancy observed in Experiment 2 may have resulted from a limitation of experimental methodology; it is possible that the requirement to use stimuli in addition to verbalising and designing increased the cognitive workload of stimulus participants past optimal performance. It is well-known that increased mental workload can decrease human performance (Wickens and Hollands 2000, Drews *et al.* 2009), where human performance is creativity in this case. Table 14 enumerates the tasks performed by no-stimulus participants versus stimulus participants and shows that stimulus participants perform an additional task compared to no-stimulus participants.

A cognitive workload assessment, e.g. NASA Task Load Index, can be used to determine if stimulus use in verbal protocol design experiments will increase cognitive workload (and thus decrease performance in terms of creativity). For example, the NASA Task Load Index asks participants to rate perceived workload in several dimensions during or after completion of a task (Hart and Staveland 1988). Other methods can also be used to determine the cognitive workload associated with design tasks. For example, Tang and Zeng (2009) investigated the use of body movements to quantify a designer's mental stress during the design process. As it may not be possible to equalise the workload between the stimulus and control conditions, i.e. comparing designs generated using stimulus versus no stimulus, this may be a serious limitation of verbal protocol experiments as applied to design research. Review of other verbal protocol experiments reveal at least one other reported case in which an experimental manipulation (which should improve design overall) did not produce better design concepts despite an increase of other metrics (Atman and Bursic 1996). In Atman and Bursic's study, reading a short design text before verbalising and designing improved metrics that should indicate improved design, e.g. time spent on designing, but did not result in better quality designs. More details with respect to potential limitations of verbal protocol design experiments are described by Chiu and Shu (2010).

Language analyses were conducted to gain insight into how designers used stimulus words in concept generation. First, stimulus words used in concept generation were examined to determine

Table 14. Task comparison between control and stimulus participants.

No-stimulus tasks	Stimulus tasks
1. Design	1. Design
2. Verbalise	2. Verbalise
–	3. Use stimuli

Note: Reprinted with permission. Copyright ASME 2010.

Table 15. Breakdown of average unknown usage of stimulus words.

Problem	Average unknown usage of opposite stimulus words	Average unknown usage of similar stimulus words	Comparison t-values	p-values
Bushing	26.0	16.3	$t(8) = 1.10$	0.15
Snow	18.6	17.8	$t(8) = 0.12$	0.45
Coal	31.0	17.0	$t(8) = 2.07$	0.04

the POS in which they were used, i.e. verb, noun, adjective or unknown, and then any associated NWPs were examined. Language analyses showed that verbs in general introduced more NWPs than nouns or adjectives. Furthermore, results indicate that opposite verbs may introduce the highest number of NWPs, and correspond to increased creativity measures as determined by the independent raters.

Combined results of the Bushing and Snow problems suggest that opposite-stimulus verbs may be the mechanism causing participants to introduce more NWPs, thus resulting in more creative concepts. However, the results of the Coal problem were contrary to those of the other two problems, in that similar-stimulus verbs introduced more NWPs than opposite-stimulus verbs. Despite the contradicting NWP introduction results for the Coal problem, opposite-stimulus concepts were still judged as more creative than similar-stimulus concepts. An explanation for this inconsistent result may be found by examining the high rate of unknown stimulus use in the Coal problem. Re-examining stimulus-use frequency for the Coal problem shows that opposite-stimulus participants had significantly more instances of unknown stimulus use, e.g. stimulus words listed without context and hence with unknown POS, on average than in any other experimental condition or problem, $t(8) = 2.07$, $p = 0.04$. The average unknown stimulus POS per problem is shown in Table 15.

Stimulus words used as an unknown POS do not introduce NWPs, but frequent instances of recorded unknown stimuli may indicate that the participant frequently looked at specific stimulus words and likely thought about those words as they designed. It is possible that participants were fatigued during the Coal problem, the last problem in the experiment, and did not verbalise all thoughts related to the task.

5. Overall discussion

Our experiments showed that oppositely related stimuli can increase concept creativity. Results suggest that use of opposite stimuli is most effective for general domain problems, or problems that are novel to the designer, e.g. the Sunflower, Egg and Snow problems. Experiment 2 also suggests that stimuli in general may be detrimental to designer performance, but this may be a limitation of the experimental method. Despite any potential methodological limitations, Experiment 2 did allow a comparative study of how language stimuli may support more creative concept generation.

Experiment 2 results show that opposite-stimulus words are associated with more NWPs, which further suggests that opposite-stimulus verbs may force introduction of NWPs so that they are correct and consistent in the context of the problem. While participants used stimuli as verbs and nouns often (and adjectives infrequently), verbs may be better at introducing NWPs because verbs are more flexible than nouns (Gentner and France 1988). Verbs are also known to unconsciously and automatically evoke concepts corresponding to the semantic filler roles typically associated with the event they denote (Lyons 1977, McRae *et al.* 2005). Common semantic filler roles include

Table 16. Semantic properties of nouns and verbs, adapted from Croft (1991).

Property	Noun	Verb
Semantic class	Object	Action
Pragmatic function	Reference	Predication
Stativity	State	Process
Persistence	Persistent	Transitory
Valency/relationality	0	1+

patients (often direct objects), agents (often subjects), instruments and locations. For example, the verb 'hammer' will commonly evoke 'carpenter' as an agent role, and 'nail' as an object role.

In the Snow problem, opposite-stimulus participants were provided with the stimulus word 'constrict'. Those using it as a verb automatically realised that the verb 'constrict' had to '*constrict* something'. '*Constrict* snow' (to the point where it turns to ice) was inconsistent with the problem because participants were explicitly instructed not to compress snow to the point of ice. However, introducing an NWP in '*constrict* the motion of the heat', is both new, and consistent. The verb 'constrict' implies that something must be constricted, and it is flexible enough to allow for different arguments while still 'making sense'.

Examining semantic properties of nouns and verbs further explain the relationship between stimulus verbs and NWPs. Table 16 summarises some key properties and differences between nouns and verbs.

Table 16 shows verb properties that may benefit conceptual design more than noun properties. First, nouns are used to reference objects, which are usually fixed and static, while verbs represent transitory actions and processes (Croft 1991). Fixedness is contrary to the purpose of conceptual design, which is to expand the solution space, and not to fix, or limit the space. The abstractness of actions and processes allows verbs to avoid naming an actual design solution, which renders verbs more advantageous for expanding the solution space. Second, nouns have a valency of zero, while verbs have at least a valency of one. The valency, or relationality, of a word refers to the implied entities associated with the use of the word. When a noun is used, there is no other implied entity, e.g. 'book' does not imply the existence of any other entity. However, when a verb is used, at least one other entity is implied, e.g. 'hit' implies the existence of a hitter and the object hit (Lyons 1977, Croft 1991, McRae *et al.* 2005). For verbs, the implied entities are the semantic filler roles.

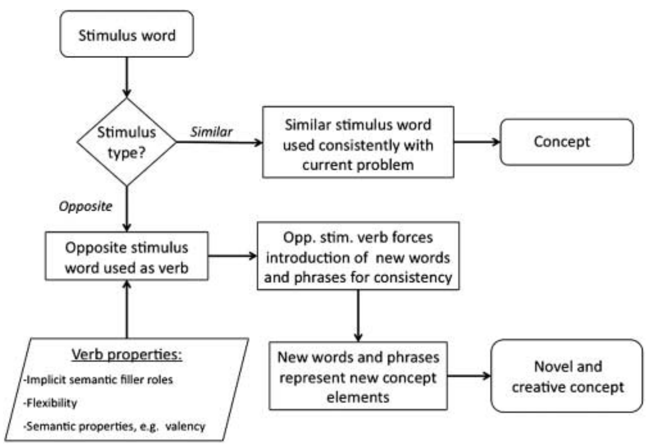

Figure 8. Explanatory model of opposite-verb NWP introduction and its effect on concept creativity.

Examining the semantic properties of nouns versus verbs aids the explanation of the empirical results obtained in Experiment 2. Both linguistic theory and empirical results suggest that verbs used as design stimuli may increase concept creativity. Furthermore, empirical results show that oppositely related verbs may be the most effective at stimulating NWPs, which may lead to more creative concepts. With similarly related verbs, NWP introduction is unnecessary because existing problem-statement words and phrases capture the current problem state, e.g. problem objects like 'snow' or 'bushings'. Similar stimuli are already consistent with the problem and do not need to be resolved by the introduction of NWPs. Figure 8 models the effect of opposite-stimulus verbs on the generation of creative concepts.

6. Summary and concluding remarks

We investigate the effects of language on design because connections between language and cognition may be used to facilitate creative and successful design. Specifically, we study oppositely related words because they may stimulate creative concept generation by being unexpected and non-obvious, while being available for systematic retrieval in lexical resources. Through experiments, we observed the effects of oppositely and similarly related stimulus words on concept creativity and designer behaviours with respect to the different stimulus types in concept generation.

Empirical results support our original hypotheses and show that opposite stimuli may increase:

(1) Concept novelty as well as other creativity metrics;
(2) Introduction of NWPs that may form the basis of novel and creative concepts.

Empirical results combined with linguistic theory allow us to speculate on the mechanism in which opposite stimuli interacts with the conceptual design process to produce more creative concepts. While similar-stimulus words can be used 'as-is' to reason about the problem and concepts consistently, opposite-stimulus words must be used with NWPs to maintain consistency within the problem. These new words or phrases may be key to the formation of more creative concepts.

Results also reinforced that problem novelty and designer experience may be a factor (Experiment 1), and that methodological issues with verbal protocols may interfere with results (Experiment 2). However, overall, our results show that opposite stimuli appear a practical means of stimulating creative design, that is simple to implement, e.g. using flashcards or worksheets. Unlike some creative design methods, e.g. TRIZ, which require training and related materials, e.g. contradiction tables or software, opposite stimuli only require the generation of words that are oppositely related to the problem. These words, which are familiar to most as antonyms, can be obtained from conventional thesauri or WordNet. The ease of obtaining oppositely related words, i.e. in antonyms, increases the ease by which this method can be integrated into engineering practice. While the designer's previous experience with similar design problems may offset the creative advantages offered by opposite-stimulus words, opposite-stimulus words are suitable for instances of conceptual design where the engineer may not have already gained familiarity with the new domain.

It is unlikely that the effects of language on design cognition and creativity can be fully understood in the near future. However, insights from our experiments combined with linguistic theory allow us to propose an explanatory model of interaction between language stimuli and design cognition. This knowledge can be used to facilitate more creative and successful design.

Acknowledgements

We gratefully acknowledge the Natural Sciences and Engineering Research Council of Canada (NSERC) for funding support. We would also like to acknowledge the participants and raters who provided their time.

References

Adams, J.L., 2001. *Conceptual blockbusting, a guide to better ideas 4/e*. New York, NY: Perseus Press.
Akin, O., 1990. Necessary conditions for design expertise and creativity. *Design Studies*, 11 (2), 107–113.
Akin, O. and Akin C., 1998. On the process of creativity in puzzles, inventions and designs. *Automation in Construction*, 7 (2–3), 123–138.
Altshuller, G.S. and Shulyak, L., 1996. *And suddenly the inventor appeared: TRIZ, the theory of inventive problem solving, 2/e*. Worcester, MA: Technical Innovation Center.
Amabile, T.M., 1983. *The social psychology of creativity*. New York: Springer-Verlag.
Amabile, T.M., 1989. *Growing up creative, nurturing a lifetime of creativity.* New York, NY: Crown Publishers, Inc.
Atman, C.J. and Bursic K., 1996. Teaching engineering design: can reading a textbook make a difference? *Research in Engineering Design*, 8 (4), 240–250.
Bainbridge, L., *et al.*, 1968. A study of real-time human decision making using a plant simulator. *Operational Research Quarterly*, Special Conference Issue, 19, 91–106. Reprinted in E. Edwards and F.P. Lees, eds. (1974). *The human operator in process control*. London: Taylor and Francis.
Benami, O. and Jin, Y., 2002. Creative stimulation in conceptual design. *Proceedings of the ASME DETC/CIE*, Montreal, Canada, DETC2002/DTM-34023.
Besemer, S.P. and Treffinger, D.J., 1981. Analysis of creative products: review and synthesis. *Journal of Creative Behavior*, 15 (3), 158–178.
de Bono, E., 1970. *Lateral thinking: creativity step by step.* New York, NY: Harper & Row.
de Bono, E., 1992. *Serious creativity*. New York, NY: HarperCollins.
Brown, D.C., 2008. Guiding computational design creativity research. *Proceedings of NSF International Workshop on Studying Design Creativity '08*, University of Provence, France.
Bruner, J.S., 1964. The course of cognitive growth. *American Psychologist*, 19 (1), 1–15.
Carroll, D.W., 1999. *Psychology of language, 2/e.* Pacific Grove, CA: Brooks/Cole Publishing Company.
Chakrabarti, A., Morgenstern, S., and Knaab, H., 2004. Identification and application of requirements and their impact on the design process: a protocol study. *Research in Engineering Design*, 15 (1), 22–39.
Chiu, I. and Shu, L.H., 2007. Biomimetic design through natural language analysis to facilitate cross-domain information retrieval. *Artificial Intelligence for Engineering Design, Analysis and Manufacturing*, 21 (1), 45–59.
Chiu, I. and Shu, L.H., 2008a. Use of opposite-relation lexical stimuli in concept generation. *Annals of the CIRP*, 57 (1), 149–152.
Chiu, I. and Shu, L.H., 2008b. Effects of dichotomous lexical stimuli in concept generation. *Proceedings of ASME international design engineering technical conferences*, New York City, NY, USA, 3–6 August 2008. DETC2008-49372 (DTM).
Chiu, I. and Shu, L.H., 2010. Potential limitations of verbal protocols in design experiments. *Proceedings of ASME international design engineering technical conferences*, Montreal, Quebec, Canada, 15–18 August 2010. IDETC2010-28675 (DTM).
Cheong, H., *et al.*, 2011. Biologically meaningful keywords for functional terms of the functional basis. *Journal of Mechanical Design*, 133 (2), 021007: 1–11.
Chomsky, N., 1968. *Language and mind*. New York, NY: Harcourt Brace Jovanovich.
Coley, F., Houseman, O., and Roy, R., 2007. An introduction to capturing and understanding the cognitive behaviour of design engineers. *Journal of Engineering Design*, 18 (4), 311–325.
Corremans, J.A.M., 2011. Measuring the effectiveness of a design method to generate form alternatives: an experiment performed with freshman students product development. *Journal of Engineering Design*, 22 (4), 259–274.
Croft, W., 1991. *Syntactic categories and grammatical relations: the cognitive organization of information*. Chicago, IL: University of Chicago Press.
Cross, N., 2006. *Designerly ways of knowing.* London: Springer-Verlag.
Cross, N., Christiaans, H., and Drost, K., 1996. Introduction: the Delft protocols workshop. *In*: N. Cross, H. Christiaans, and K. Drost's, eds. *Analysing design activity*. West Sussex: John Wiley & Sons, 1–16.
Dieter, G.E., 2000. *Engineering design: a materials and processing approach*. 3rd ed. New York, NY: McGraw-Hill.
Dong, A., Hill, A.W., and Agogino, A.M., 2003. A document analysis method for characterizing design team performance. *Journal of Mechanical Design*, 126 (3), 378–385.
Dong, A., 2006. Concept formation as knowledge accumulation: a computational linguistics study. *Artificial Intelligence for Engineering Design Analysis & Manufacturing*, 20 (1), 35–53.
Drews, F.A., *et al.*, 2009. Text messaging during simulated driving. *Human Factors: The Journal of the Human Factors and Ergonomics Society* [online]. Available from: http://hfs.sagepub.com [Accessed 17 January 2010].
DTM, 2006. *Design that matters design challenge portfolio: shelling machines* [online]. Available from: http://www.designthatmatters.org/ [Accessed 17 October 2006].

Engen, T., 1972. Psychophysics II. Scaling methods. *In*: J.W. Kling, L.A. Rigg, Woodworth & Scholsberg, eds. *Experimental psychology*. 3rd ed. New York, NY: Holt Rinehart & Winston, 47–86.

Ericsson, K.A. and Simon, H.A., 1993. *Protocol analysis: verbal reports as data*. Cambridge, MA: MIT Press.

Fantoni, G., Taviani, C., and Santoro, R., 2006. Design by functional synonyms and antonyms: a structured creative technique based on functional analysis. *Proceedings of the Institution of Mechanical Engineers, Part B, Journal of Engineering Manufacture*, 221 (4/2007), 673–683.

Fellbaum, C., 1993. *English verbs as a semantic net, in five papers on WordNet*, 40–61 [online]. Available from: ftp://ftp.cogsci.princeton.edu/pub/wordnet/5papers.ps [Accessed 7 July 2008].

Gentner, D. and France, I., 1988. The verb mutability effect: studies of the combinatorial semantics of nouns and verbs. *In*: S. Small, G. Cottrell, and M. Tanenhaus, eds. *Lexical ambiguity resolution*. Los Altos, CA: Morgan Kaufmann, 343–382.

Gero, J.S., Sushil, J.L., and Kundu, S., 1994. Evolutionary learning of novel grammars for design. *Artificial Intelligence for Engineering Design Analysis & Manufacturing*, 8 (2), 83–94.

Gordon, W.J.J., 1961. *Synectics, the development of creative capacity*. New York, NY: Harper & Row.

Hacco, E. and Shu, L., 2002. Biomimetic concept generation applied to design for remanufacture. *Proceedings of ASME design engineering technical conferences*, Montreal, DETC2002/DFM-341.

Hart, S.G. and Staveland, L.E., 1988. Development of NASA-TLX (task load index): results of empirical and theoretical research. *In*: P.A. Hancock and N. Meshkati, eds. *Human mental workload*. Elsevier Science: NorthHolland, 139–183.

Hatchuel, A. and Weil, B., 2009. C-K design theory: an advanced formulation. *Research in Engineering Design*, 19 (4), 181–192.

Holt, K., 1993. Computer-aided creativity in engineering design. *Journal of Engineering Design*, 4 (4), 371–376.

Howard, T.J., Culley, S.J., and Dekoninck, E., 2008. Describing the creative design process by the integration of engineering design and cognitive psychology literature. *Design Studies*, 29 (2), 160–180.

Hubka, V. and Eder, W.E., 1996. *Design science – introduction to the needs, scope and organization of engineering design knowledge*. London: Springer-Verlag.

Jackendoff, R., 1983. *Semantics and cognition*. Cambridge, MA: MIT Press.

Jin, Y. and Benami, O., 2010. Creative patterns in conceptual design, artificial intelligence for engineering design. *Analysis and Manufacturing*, 24 (2), 191–209.

Jin, Y., Geslin, M., and Lu, S.C.-Y., 2007. Impact of argumentative negotiation on collaborative engineering. *Annals of the CIRP*, 56 (1), 181–184.

Kalmar, I. and Davidson, D., 1997. *Anthropological linguistics and semiotics*. 2nd ed. Toronto, Canada: Quirk Press.

Kan, J.T.W. and Gero, J.S., 2007. Using the FBS ontology to capture semantic design information in design protocol studies. *In*: J. McDonnell and P. Lloyd, eds. *DTRS7*. London: University of the Arts, 155–165.

Keller, R., Eckert, C.M., and Clarkson, P.J., 2009. Using an engineering change methodology to support conceptual design. *Journal of Engineering Design*, 20 (6), 571–587.

Kim, Y.S., Jin, S.T., and Lee, S.W., 2011. Relations between design activities and personal creativity modes. *Journal of Engineering Design*, 22 (4), 235–257.

Kosse, V., 2004. *Solving problems with TRIZ: an exercise handbook*. 2nd ed. Southfield, MI.: Ideation Int'l Inc.

Kudrowitz, B.M. and Wallace, D.R., 2010. Assessing the quality of ideas from prolific, early-stage product ideation. *Proceedings of the ASME IDETC/CIE 2010*, Montreal, Quebec, Canada, August 15–18, 2010 (DETC2010-28991).

Landis, J.R. and Koch, G.G., 1977. The measurement of observer agreement for categorical data. *Biometrics*, 33 (1), 159–174.

Levinson, S., 1996. Language and space. *Annual Review of Anthropology*, 25, 353–382.

Li, P. and Gleitman, L., 2002. Turning the tables: language and spatial reasoning. *Cognition*, 83 (3), 265–294.

Lutfey, K.E., *et al.*, 2008. How are patient characteristics relevant for physicians' clinical decision making in diabetes? An analysis of qualitative results from a cross-national factorial experiment. *Social Science & Medicine*, 67 (9), 1391–1399.

Lyons, J., 1977. *Semantics*. vol. 2. Cambridge, NY: Cambridge University Press.

McNeill, T., Gero, J.S., and Warren, J., 1998. Understanding conceptual electronic design using protocol analysis. *Research in Engineering Design*, 10 (3), 129–140.

McRae, K., *et al.*, 2005. A basis for generating expectancies for verbs from nouns. *Memory & Cognition*, 33 (7), 1174–1184.

Merriam-Webster.com, 2008. *Merriam-Webster online dictionary* [online]. Available from: http://merriam-webster.com [Accessed 23 September 2008].

Miller, G., *et al.*, 1993. *Introduction to WordNet: an on-line lexical database, in Five papers on WordNet, 1–25* [online]. Available from: ftp://ftp.cogsci.princeton.edu/pub/wordnet/5papers.ps [Accessed 7 July 2008].

Murphy, M.L., 2003. *Semantic relations and the lexicon: antonymy, synonymy and other paradigms*. Cambridge: Cambridge University Press.

Myers, D.G., 1999. *Social psychology, 6/e*. Boston, MA: McGraw-Hill Publishers.

Nagai, Y. and Noguchi, H., 2003. An experimental study on the design thinking process started from difficult keywords: modeling the thinking process of creative design. *Journal of Engineering Design*, 14 (4), 429–437.

Nagai, Y. and Taura, T., 2006. Formal description of concept-synthesizing process for creative design. *In*: J.S. Gero, ed. *Design computing and cognition '06*. Dordrecht, The Netherlands: Springer, 443–460.

Nisbett, R. and Wilson, T., 1977. Telling more than we can know: verbal reports on mental processes. *Psychological Review*, 84 (3), 231–259.

Nuseibeh, B. and Easterbrook, S., 2000. Requirements engineering: a roadmap. *In*: A.C.W. Finkelstein, ed. *The future of software engineering*. Limerick, Ireland: IEEE Computer Society Press, 35–46.

Pahl, G. and Beitz, W., 1996. Engineering design, a systematic approach. *In*: K. Wallace, L. Blessing, and F. Bauert, trans., K. Wallace, eds. 2/e., London: Springer-Verlag London Ltd.
Pinker, S., 2007. *The stuff of thought, language as a window into human nature*. New York, NY: Viking.
Ratner, N. and Gleason, J., 1993. An introduction to psycholinguistics. *In*: J.B. Gleason and N.B. Ratner, eds. *Psycholinguistics*. Orlando, FL: Harcourt Brace College Publishers, 1–40.
Rittel, H.W.J. and Webber, M.M., 1984, Planning problems are wicked problems. *In*: N. Cross', ed. *Developments in design methodology*. Bath: John Wiley & Sons, 135–144.
Saeed, J.I., 2003, *Semantics*. 2nd ed. Oxford: Blackwell Publishing.
Segers, N., 2004. *Computational representations of words & representations of words & associations in architectural design, development of a system support creative design*. Thesis (PhD). Bouwstenen 78, Technische Universiteit Eindhoven.
Shah, J., Kulkarni, S., and Vargas-Hernandez, N., 2000. Evaluation of idea generation methods for conceptual design: effectiveness metrics & design of experiments. *Journal of Mechanical Design*, 122 (4), 377–384.
Shai, O., Reich, Y., and Rubin, D., 2009. Creative conceptual design: extending the scope by infused design. *Computer Aided Design*, 41 (4), 117–135.
Siegal, S., 1956. *Non parametric statistics for the behavioral sciences*. New York, NY.: McGraw-Hill, Inc.
Simon, H., 1969. *The sciences of the artificial*. Cambridge, MA: MIT Press.
Stone, R.B. and Wood, K.L., 2000. Development of a functional basis for design. *Journal of Mechanical Design*, 122 (4), 359–369.
Tang, Y., and Zeng., Y., 2009. Quantifying designer's mental stress in the conceptual design process using kinesics study. *Proceedings of International Conference on Engineering Design, ICED '09*, 24–27 August 2009, Stanford University, Stanford, CA.
Thomas, J.C. and Carroll, J.M., 1984. The psychological study of design. *In*: N. Cross's, ed. *Developments in design methodology*. Chichester: John Wiley & Sons, 221–235.
Torrance, E.P., 1974. *Torrance tests of creative thinking*. Bensenville, IL: Scholastic Testing Service, Inc.
Tseng, I., *et al.*, 2008. Overcoming blocks in conceptual design: the effects of open goals and analogical similarity on idea generation. *Proceedings of ASME IDETC/CIE*, Brooklyn, NY, August 3–6, DEC2008-49276.
Ullman, D., 2003. *The mechanical design process*. 3rd ed. New York, NY: McGraw-Hill.
Visser, W., 2006. *The cognitive artifacts of designing*. Mahwah, NJ: Lawrence Erlbaum Associates, Publishers.
Wickens, C.D. and Hollands, J.G., 2000. *Engineering psychology and human performance*. 3rd ed. Upper Saddle River, NJ: Prentice Hall.
Wilson, R.C., 1951. An operational definition of originality. *American Psychologist*, 6 (6), 297..
Princeton University, 2008. *About WordNet*. WordNet., Princeton University. Available from: http://WordNet.Princeton.edu [Accessed 7 July 2008].
Yang, M.C., 2009. Observations on concept generation and sketching in engineering design. *Research in Engineering Design*, 20 (1), 1–11.