Expert representation of design repository space: A comparison to and validation of algorithmic output

Katherine Fu, Department of Mechanical Engineering, Carnegie Mellon University, 5000 Forbes Ave., Pittsburgh, PA 15217, USA

Joel Chan and *Christian Schunn*, Department of Psychology, University of Pittsburgh, Pittsburgh, PA, USA

Jonathan Cagan, Department of Mechanical Engineering, Carnegie Mellon University, Pittsburgh, PA, USA

Kenneth Kotovsky, Department of Psychology, Carnegie Mellon University, Pittsburgh, PA, USA

Development of design-by-analogy tools is a promising design innovation research avenue. Previously, a method for computationally structuring patent databases as a basis for an automated design-by-analogy tool was introduced. To demonstrate its strengths and weaknesses, a computationally-generated structure is compared to four expert designers' mental models of the domain. Results indicate that, compared to experts, the computationally-generated structure is sensible in clustering of patents and organization of clusters. The computationally-generated structure represents a space in which experts can find common ground/consensus — making it promising to be intuitive/accessible to broad cohorts of designers. The computational method offers a resource-efficient way of usefully conceptualizing the space that is sensible to expert designers, while maintaining an element of unexpected representation of the space.

© 2013 Elsevier Ltd. All rights reserved.

Design by analogy, in which designers draw inspiration from cross-domain design solutions, is a promising methodology for design practice (e.g., Bhatta & Goel, 1996; Davies, Goel, & Nersessian, 2009; Goel, 1997; Herstatt & Kalogerakis, 2004; Vattam, Helms, & Goel, 2010). Not all analogies are equally useful (Casakin & Goldschmidt, 1999; Chan et al., 2011; Dunbar, 1997; Fu, Chan, et al., 2013; Gick & Holyoak, 1980; Weisberg, 2009), raising the question of how to organize the very large space of possible analogies to any given problem. A number of methods exist for organizing spaces of possible design analogies (e.g., Gentner et al., 1997; Gentner & Markman, 1997; Linsey, Markman, & Wood, 2008; Mcadams &

Corresponding author:
Katherine Fu
Katherine.Fu@gmail.
com

ELSEVIER

Wood, 2000; Murphy, 2011; Verhaegen, D'hondt, Vandevenne, Dewulf, & Duflou, 2011) that make use of varying computational and representational techniques. For example, the computational work presented in this paper considers a diverse set of possible structural types (e.g., trees, rings, chains) and uses Bayesian methods for choosing the best type and the best instance within each type.

Because design-by-analogy inherently has a human in the loop, it is useful to understand how meaningful expert designers find the algorithm-generated organizations of possible inspirational examples. We hypothesize that a technique previously found to capture psychologically meaningful structures in other contexts is especially likely to provide structures that align with the mental models that humans develop through the design process. Such alignment could aid in both the effectiveness of the tool at stimulating designer's thinking and in easing the transition for its adoption. The study presented here attempts to validate the proposed computational methodology and its output through judgment of its accuracy and adequacy by four design experts and also a direct comparison of its output to the experts' structuring of the same patent space, which presumably reflects their mental model of the space.

1 Background

Analogy and external stimuli in engineering design have been studied in a number of ways. Studies have been performed to understand how the introduction of analogies affects the ideation process and outcomes (e.g., (Christensen & Schunn, 2005; Dahl & Moreau, 2002; Goldschmidt & Smolkov, 2006; Linsey et al., 2008), with some studies specifically examining how the introduction of analogies with different levels of applicability to the design problem affects individual designers (Christensen & Schunn, 2007; Tseng, Moss, Cagan, & Kotovsky, 2008). For brevity, we do not dive deeply into the design by analogy literature, here. Most relevant to the current context, the literature finds that the distance of analogies to the current design problem can have a large influence on the helpfulness of the analogy, with some distance being useful (Chan et al., 2011; Dahl & Moreau, 2002; Dyer, Gregersen, & Christensen, 2011; Fu, chan, et al., 2013; Gentner & Markman, 1997; Wilson, Rosen, Nelson, & Yen, 2010), but analogies too far away show no benefit (Fu, chan, et al., 2013). Thus, finding a tool that organizes the space of possible analogs by distance to the current problem may be especially useful.

1.1 Computational design tools

The development of computational tools to assist designers during the design process is an area of considerable research in engineering design. To facilitate functional modeling, Stone and Wood created a functional basis that can serve as a universal language in the ideation process (Stone & Wood, 2000). This work has been extended and adapted a great deal, including, for example, the creation of a biological functional basis (Cheong, Shu, Stone, & Wood, 2008).

The U.S. patent database conveniently catalogs existing technology and engineering design and, thus, has been used in the development of numerous computational design aids. For example, TRIZ helps engineers overcome difficulties in functional reasoning by searching through patents based on heuristic rules (such as use of opposites) (Altshuller & Shapiro, 1956). One axiomatic conceptual design model relies on the combination of TRIZ and functional basis (Zhang, Cha, & Lu, 2007). Patent citation data has been used to identify the interrelatedness of technologies, and to demonstrate the benefits of tapping into the technology knowledge base created by competitors within a design field (Chakrabarti, Dror, & Nopphdol, 1993). The syntactic similarity between patent claims has been investigated as an aid in patent infringement research (Indukuri, Ambekar, & Sureka, 2007). Businesses have used patent repository tools and patent mining to predict potential market trends, identify prolific inventors, and for other purposes. Among the characteristics explored in patent mining have been the number of citations, number of claims, average number of words per claim, and number of classes that the patent spans (Kasravi & Risov, 2007). The potential for using the Design Repository at Oregon State University as a design aid was demonstrated by Bohm et al. by performing a function-based search using Chi Matrix and Morphological Matrix techniques to find components that were present in concepts generated by hand (Bohm, Vucovich, & Stone, 2005). More broadly, design repositories of CAD models of components and assemblies have been explored as resources for designers to streamline product design of complex engineering systems by reusing, revising or gaining insight into previous designs and models (Szykman, Sriram, Bochenek, Racz, & Senfaute, 2000). The computational work examined in the current study focuses on structuring design repositories, using more open-ended analogical transfer, and using the textual content of the patents, which may generate richer outcomes.

In order to assist in avoiding patent infringement, Mukherjea et al. created a BioMedical Patent Semantic Web that discovered semantic associations between important biological terms within biomedical patents, returning a ranked list of patent resources and a Semantic Web showing the relationships between the terms and between resources. Unlike our work, the Semantic Webs generated by Mukherjea et al. are fully connected graphs with no imposed structure, and the data includes only the abstract of the patents being examined. Further, a Bayesian inference approach was not use to generate the webs (Mukherjea, Bhuvan, & Kankar, 2005). Bayesian inference was used by Chakrabarti et al., however, to train a model using a small data set of documents, which was then used to categorize the remaining documents into "topics" and create a hierarchical structure (Chakrabarti, Dom, Agrawal, & Raghavan, 1998). Structures other than hierarchies were not explored and the structures created were not used as input for analogical design work.

PatViz, a tool developed by Koch et al. allows visual exploration of iterative and complex patent searches and queries using a variety of patent data,

including full text. Users may create one graph view with PatViz through a guided process, rather than an algorithm. The tool includes three visualizations of interest: Patent Graph, which creates a fully-connected web of patents; 3D IPD Treemap, which creates a 3D tree structure of the patents based on a predefined classification scheme; and Aggregation Tree, which is another tree view that uses predefined adjustable hierarchies (Koch, Bosch, Giereth, & Ertl, 2009). Unlike the work of Koch et al., our work does not rely on predefined or user-defined classification schemes; rather, it uses an exploration methodology to discover the best-fitting structures among multiple types of structures to describe the set of patents. The fit to the data being examined determines the form of the structure.

While the U.S. patent database represents a potentially productive repository of analogical or cross-domain design solutions, its size and complexity make it difficult for designers to use effectively. There have been many attempts to automate, aid, or streamline the search of the US patent database, including simple key word searches on the United States Patent and Trademark Office (USPTO) website or Google Patents, as well as theories like TRIZ and their resulting tools (Altshuller & Shapiro, 1956; Duran-Novoa, Leon-Rovira, Aguayo-Tellez, & Said, 2011; Hernandez, Schmidt, & Okudan, 2012a, 2012b; Houssin & Coulibaly, 2011; Krasnoslobodtsev & Langevin, 2005; Liang, Tan, & Ma, 2008; Nakagawa, 2012; Nix, Sherret, & Stone, 2011; Zhang et al., 2007). It is still difficult, however, to recognize and appreciate the characteristics that might be relevant to a particular design problem within the 'space' of patents. A number of computational "innovation support tools" have been developed (CREAX; Goldfire), as well as a number of research-driven design support tools, proposals for design support tools, and methodologies (Bhatta & Goel, 1996; Chiu & Shu, 2005; Goel, Bhatta, & Stroulia, 1997; Verhaegen et al., 2011; Vincent, Bogatyreva, Bogatyreva, Bowyer, & Pahl, 2006; Linsey, Wood & Markman, 2008). Many of these tools and methods, however, rely heavily on the users to generate the terms or analogies and comb through search results. Psychology literature suggests that the probability of retrieving similarities is low when surface dissimilarities exist, because retrieval of far-field analogies is cognitively difficult (Gick & Holyoak, 1980) and the range of remindings tend to be limited by surface similarity (Forbus, Gentner, & Law, 1994). If designers have a means to extract the interrelatedness and interconnectedness of patents in the space, especially with respect to a specific design problem, they may be able to more strategically choose which cross-domain designs to review or to traverse the space in a more intentional and meaningful exploratory way.

1.2 Computational basis

The computational methodology for discovering structural forms in patent spaces is based on the work of Kemp and Tenenbaum (2008). Kemp and Tenenbaum devised an algorithm that can discover the underlying structure that is

inherent to a particular data set, enabling the extraction of psychologically meaningful relationships or information. The algorithm discovers 1) the best-fitting form for a given set of data from a space of 8 possible forms, which include a partition, order, chain, ring, tree, hierarchy, grid, and cylinder; and 2) the best-fitting instantiation of that selected form, which is called a structure. For example, given human similarity judgments for a set of sixteen colors, the structural form discovered by the algorithm that best described the data was a ring, matching the color wheel with which we are all familiar from our elementary school days. Given data regarding votes of Supreme Court justices, the best structural form discovered was a chain, organizing the justices from liberal to conservative; again, a meaningful structuring. Given data regarding features that a set of animals have, a tree structure was discovered to be the best, similar in form type to the biological classification scheme proposed by Linnaeus. The form types explored are based on the psychology literature (Jaynes, 2003), and have been used in formal models in many contexts (Anderson, 1991; Bradley & Terry, 1952; Carroll, 1976; Collins & Quillian, 1969; Fiske, 1992; Guttman, 1944, 1954; Huelsenbeck & Ronquist, 2001; Inhelder & Piaget, 1969; Kohonen, 1997; Shepard, 1980; Sneath & Sokal, 1973; Wiggins, 1996). In addition, the algorithm is based on Bayesian inference, a statistical basis that has a long history of successfully describing and modeling human cognition (Griffiths, Kemp, & Tenenbaum, 2008). Jaynes describes human plausible reasoning as a calculation of the degree of plausibility of a particular hypothesis being true based on previous experience and common sense, and given the facts at hand, corresponding directly to the components that must be considered when calculating the posterior probability of a hypothesis being true given a set of data using Bayes Rule (Jaynes, 2003). This basis in human cognition is an important motivation for the choice to use the Kemp and Tenenbaum algorithm in the computation methodology being validated in this study, as it has the potential to be intuitive in the human mind and thus form the basis for an automated design inspiration tool.

2 The current study

The goal of the current study was to gain a better understanding of the output structures of the algorithm through comparison to human judgment. Structuring algorithms such as Kemp and Tenenbaum's (2008) have customarily employed comparisons to pre-existing "correct" structures as a method of validation. In a similar vein, we employed comparisons of the algorithm's output structures to similar structures generated by expert designers, and also solicited their qualitative reactions to particular aspects of the output structure. This method was judged as appropriate for validation for two reasons. First, to our knowledge, there exists no one "psychologically meaningful", "correct", or "analytical" solution to the question of how to best structure design repository data or patents. Second, the final application of this structuring work is intended to be the basis for an automated design by analogy stimulation tool, making it crucial that the output is intuitive or sensible to human minds.

Specifically, the study was performed to address two primary research questions:

1. In what ways is the algorithm's patent structure similar to or different from how an expert designer might structure a space of patents in preparation for inspiration through design by analogy with regard to a specific design problem?
2. If the structure is different, can experts nevertheless make sense of it and perhaps see new connections between patents?

The reasoning chain in Figure 1 represents possible outcomes that feed into three main potential inferences from the study. First, suppose that the algorithm structures patents in the exact same way that experts do. In that case, we might infer then that (1) **the algorithm does what an expert would be able to do, but more *efficiently***. This would not be an insignificant outcome, as limitations of attention and working memory capacity (Cowan, 2000), and time and effort effectively preclude the possibility of structuring large databases of potential analogical stimuli by hand. Furthermore, supporting new ways of representing the patent space are not the only way to help inspire innovation. Experts' mental representations are often based on deep structural similarity between knowledge structures and domains (Chi, Feltovich, & Glaser, 1981), and experts are often able to draw deep structural analogies from memory (Ahmed & Christensen, 2009; Ball & Christensen, 2009; Christensen & Schunn, 2007; Holyoak & Thagard, 1996; Novick, 1988). Accurate externalizations of experts' mental representations could potentially *inspire* innovation by supporting search for non-obvious analogous solutions to design problems that are hidden by surface dissimilarities (e.g., application domain, engineering discipline). Further, engineering design is inherently and increasingly interdisciplinary (Reich & Shai, 2012); even experts with one disciplinary focus may be inspired by representations that include insights from other relevant
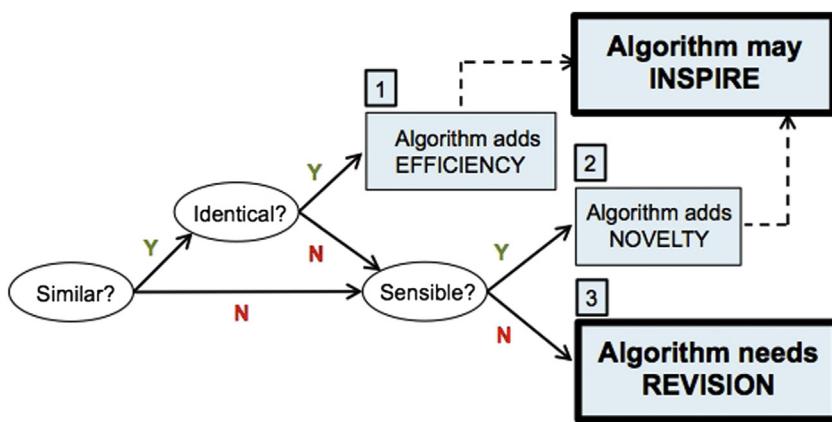

Figure 1 Reasoning chain of possible outcomes of study and main possible inferences.

engineering disciplines. The possibility of being able to aggregate expert representations across domains could be useful in catering to the increasingly interdisciplinary nature of engineering design. These aggregate representations could give access to solutions that might be hidden by surface dissimilarities, given that when one approaches a domain in which she is not expert, processing is easily dominated by surface features (Chi et al., 1981; Hmelo-Silver & Pfeffer, 2004; Novick, 1988).

Secondly, suppose that a) the algorithm produces a structure that is completely different from that of the experts, but the experts are nevertheless able to make sense of it, or b) the algorithm produces a structure that bears both similarities and differences to that of the experts, and the differences are not so large as to render the structure incomprehensible. From both these outcomes, we could infer that (2) **the algorithm may structure the analogical stimuli in a way that potentially allows for *novel* connections to be made between patents.** The algorithm's grouping of patents into nodes based on functional similarity could highlight features and functional principles that would otherwise potentially be overlooked (e.g., via analogical comparison; (Gentner et al., 1997; Kurtz, Miao, & Gentner, 2001)) yielding fresh insights for creative ideation that could *inspire* innovation. In the cognitive psychology literature, it has been shown that enabling changes in representation of objects and/or ideas, for example by leading problem solvers to attend to previously ignored features, is an effective way of dealing with "functional fixedness", where problem solvers have difficulty seeing a potential creative use of an object with which they are familiar (Kaplan & Simon, 1990; McCaffrey, 2012).

Finally, suppose that a) the algorithm produces a structure that is completely different from that of the experts, and it is incomprehensible to them, or b) the algorithm produces a structure that bears both similarities and differences to that of the experts, and the differences are so large as to render the structure incomprehensible. From both these outcomes, we could infer that (3) **the algorithm may need *revision* in order to be useful for supporting design-by-analogy**. These outcomes and inferences comprise the set of potential contributions of the study to the research and development of our automated design-by-analogy support tool.

3 Methods

To address our research questions, we asked a set of expert designers to 1) structure a set of patents, and 2) provide focused feedback on our algorithm's output structure. Structuring of the patents was conducted in the context of preparing to solve a particular design problem, given our interest in comparing the experts' judgments of patents' relevance to the design problem to the algorithm's ordering of patents' relevance as represented in the layout of the structure. The similarity of our algorithm's output structure to the experts' structures was evaluated both quantitatively and qualitatively. In the

following sections, we describe the details of our methods of data collection, and also the algorithm we used to produce the structure.

3.1 Participants

This study was conducted with four design experts in Pittsburgh, PA, USA. All participants were males with at least 10 years of experience in the field of product design, and an educational background in engineering or industrial design. Three participants were employed at local product development firms, and one was employed at a local biomedical apparatus design and manufacture firm. The participants were highly experienced and comfortable in reviewing and researching patents. Each participant was compensated with $100 for the 2 h of the study. Though there were only a small number of participants, the information is considered valuable because it is based on expert data, and includes deep analysis of interaction between 45 different pieces of information (patents) for each of the four experts. Further, significant insights in the study of knowledge representations have often been obtained with only a few participants whose behaviors are examined in depth (Chase & Simon, 1973; Chi et al., 1981; Chi & Koeske, 1983; Gobet & Simon, 2000; Huth, Nishimoto, Vu, & Gallant, 2012; Morais, Olsson, & Schooler, 2013), and a review of research on design revealed that published studies have often been conducted with qualitative analyses of the performance of a small number of experts (Mehalik & Schunn, 2007).

3.2 Materials

The experts were supplied with the following materials:

- Consent form
- Design problem description (see Section 3.3), on a 8.5 in by 5.5 in piece of paper, laminated, with mounting putty on the back
- Title, abstract text and one key figure from thirty-seven randomly selected mechanical patents and eight patents selected from previous studies presented in Chan et al. (2011) and Fu, chan, et al. (2013) (see Appendix A), each on a 8.5 in by 5.5 in piece of paper, laminated, with mounting putty on the back
- A 56 in by 107 in sheet of white craft paper, fixed to a wall
- Multiple colored magic markers, one black marker, and one pen
- Pad of sticky notes
- An 8.5 in by 11 in depiction of the algorithm's output structure for the same 45 patents
- Title, abstract text and one key figure from patents in 3 different clusters of the algorithm's output structure, each on a 8.5 in by 5.5 in piece of paper.

3.3 Design problem

The design problem supplied to the experts is one that has been used throughout the development and analysis of this work on the structuring of

patents for use in design by analogy, as presented in Chan et al. (2011) and Fu, chan, et al. (2013). The textual design problem description as supplied to the experts of this study is as follows:

Design a device to collect energy from human motion for use in developing and impoverished rural communities in places like India and many African countries. Our goal is to build a low-cost, easy to manufacture device targeted at individuals and small households to provide energy to be stored in a rechargeable battery with approximately 80% efficiency. The energy is intended to be used by small, low power draw electrical devices, such as a radio or lighting device, hopefully leading to an increase in the quality of life of the communities by increasing productivity, connection to the outside world, etc. The target energy production is 1 kW-hour per day, roughly enough to power eight 25 Watt compact florescent light bulbs for 5 hours each per day, or enough to power a CB radio for the entire day.

For reference, an average adult human can output about 200 watts with full body physical activity for short periods of time, with a significant reduction for sustained power output.

3.4 Patent selection

An initial random subset of 10,000 patents was selected from the patent database. From this initial set, patents were chosen for the final set if they were classified within the U.S. Patent classification system as "Body Treatment And Care, Heating And Cooling, Material Handling And Treatment, Mechanical Manufacturing, Mechanical Power, Static, and Related Arts". In addition, for continuity with previous studies and an understanding of human judgment of these patents, eight additional patents were included from two previous studies performed by the authors (Chan et al., 2011; Fu, chan, et al., 2013). Four of these patents were the "far" analogical stimuli from the study performed to understand how analogical distance, commonness, and modality affect ideation outcomes presented in Chan et al. (2011), and the other four were the "far" analogical stimuli from a study performed to test the use of structures of patents created using Kemp and Tenenbaum's discovery of structural form algorithm to choose "near" and "far" patents and measure their effect on ideation outcomes, presented in Fu, chan, et al. (2013). Both previous studies used the same design problem used here, described in Section 3.3. In total, 45 patents were provided to the experts as described in Section 3.2.

3.5 Algorithmic structure generation

The method used to produce the structure of patents shown in Figure 2 has been presented in depth in Fu (2012) and Fu, Cagan, Kotovsky & Wood (2013). The computational methodology for generating structures of patents
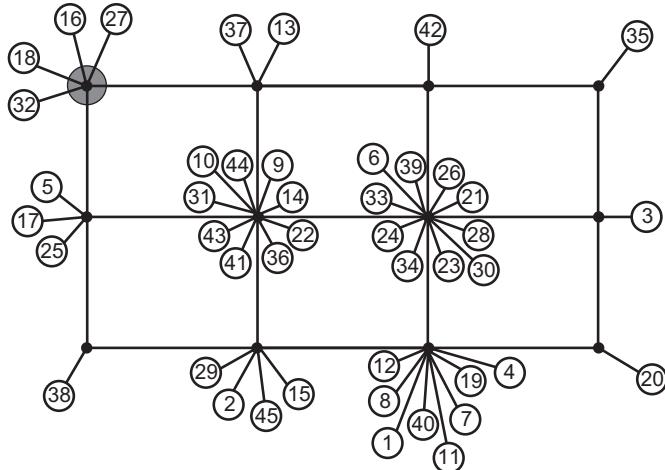
Figure 2 Algorithm output structure of 45 patents.

has three main parts. First, the set of 45 full-text patents are preprocessed using Latent Semantic Analysis (LSA), a computational text analysis tool that allows for the extraction of contextual similarity and synonymy between words and documents (Deerwester, Dumais, Furnas, & Landauer, 1990; Foltz, Kintsch, & Landauer, 1998; Landauer, Foltz, & Laham, 1998). The output from this first step is a similarity matrix, assigning a numeric value between 0 and 1, connoting semantic similarity, to all pairwise relationships between patents.

This similarity matrix is the input to the second portion of the methodology, which is an algorithm for the discovery of structural form in data, devised by Kemp and Tenenbaum (2008). The algorithm is a hierarchical Bayesian algorithm that discovers the structure and form type, chosen from a predefined set of form types, including the partition, order, chain, ring, hierarchy, tree, grid, and cylinder, that achieves the highest posterior probability value for the description of the data by a particular form type and structure. Within the algorithm, the form of the structure itself changes as the data being examined changes and has the potential of revealing deep structural relationships between patents.

Finally, LSA is used again to generate labels for the output structure, which describe the clusters of patents in the best structure, allowing a designer viewing the structure to understand the meaning of the clustering of the patents, and characterize the space for navigation and exploration. Labels were not included in the structures used in this study, as the researchers were interested in the alignment of expert-generated labels and computationally-generated labels. A further post-processing step is done to extract "regions" of similarity in the structure, which further facilitate efficient understanding and exploration of the space of patents. These regions were also excluded

from the computational output used in this study. The main goal of this work is a comparison of the judgment of experts to, and a validation of, the structure(s) generated by this methodology. The full text of the 45 patents used in the methodology described above, enumerated in Appendix A, were used to create the grid structure in Figure 2.

3.5.1 Cluster label generation

The structure shown to the experts (see Figure 2) did not include labels to indicate the similarity or meaning of the associations of patents within the clusters. Having the experts suggest labels, both for the clusters within the structure generated by the algorithm, as well as for the clusters within their own patent space or structure, was a way to collect data that could then be compared to labels that were generated computationally. The labels were generated by employing LSA, as described in Fu et al. (2013). Two different methods for label generation were employed, the results of which are presented next. The first method, the Highest Average Rank Labeling method, averaged the ranks of each word in the LSA space for each patent within a cluster, and used the top five highest average ranked words for that cluster. The second method, the Highest Cosine Similarity Labeling method, gathered the top twenty words with the highest cosine similarity to each patent within a cluster, pooled and sorted the words in descending cosine similarity value order, and used the top five words with the highest cosine similarity values in the pool of words for the patents within that cluster. The results yielded by Highest Average Rank Labeling method are shown in Figure 3. The results yielded by the Highest Cosine Similarity Labeling method are shown in Figure 4. The results of both methods are discussed for the three nodes in particular in Table 6 as compared to expert generated labels. Because four of the patents included in the space from the previous work presented in Chan et al. (2011) were not
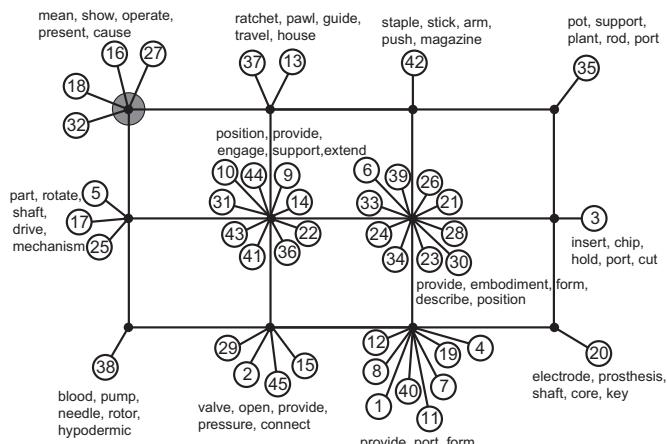

Figure 3 Algorithm output structure of 45 patents with cluster labels, highest average rank method.
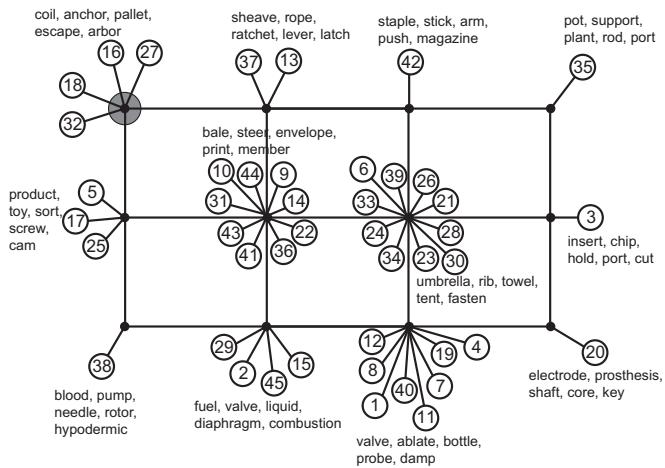
Figure 4 Algorithm output structure of 45 patents with cluster labels, highest cosine similarity method.

part-of-speech tagged, many unimportant words were included in the text, and thus the LSA space. These words were manually filtered out to increase the quality of the labels. This problem would not arise if all patents in the space have been pre-processed with part-of-speech tagging.

3.6 Experimental procedure

Each participant was allotted 2 h to complete the study, and the five sessions were run in three different locations. All locations were similar, in that they were large conference rooms, in which experts had ample light, table space on which to work, seating, and a large wall space for building the structure.

3.6.1 Experimenter's instructions and overview

The study began with one of the experimenters explaining the motivation of the study, the schedule of tasks, and the specific instructions for completing those tasks. The experts were told that the goal of the study was to understand how an expert, or professional product designer, might organize a collection of patents to facilitate searching for potentially relevant inspiration for design-by-analogy, and how this is similar or different from the output of an

Table 1 Correlations between participant and algorithm-generated structures of patents

	E1	*E2*	*E3*	*E4*	*Algorithm*
E1		0.22[a]	0.21[a]	0.02	0.24[a]
E2	0.22[a]		0.10[a]	0.01	0.11[a]
E3	0.21[a]	0.10[a]		0.00	0.15[a]
E4	0.02	0.01	0.00		0.17[a]
Algorithm	0.24[a]	0.11[a]	0.15[a]	0.17[a]	

Note: Values are Phi contingency coefficients.
[a] denotes statistical significance at the .01 level (2-tailed)

Table 2 Logistic regression of algorithm pairwise patent links on number of experts judging pairs as linked

	B	S.E.	Wald	df	p	Exp(B)	95% C.I. for Exp(B)	
							Lower	Upper
Constant	−2.59	.14	339.48	1	0.00	.08		
Expert links	0.93	.11	68.37	1	0.00	2.53	2.03	3.15

algorithm we have that tries to do this automatically for large collections of patents. The experts were then tasked with reading, processing, and placing the design problem and the 45 mechanical patents onto a large sheet of paper on the wall, such that the physical proximity between the patents connoted functional similarity, and physical proximity of patents to the design problem connoted relevance to the solving of that design problem. In addition, it was requested that the experts attempt to build the patent space in a way that facilitates searching through the collection of patents for potential inspiration for a design problem. It was emphasized that the space of patents should capture *functional* similarity, and not necessarily problem or technological domain similarity. To illustrate this point, experts were given the example of placing a patent for a car suspension system near a patent for a running shoe based on the fact that shock absorbing is a function that is key in both patents. This phase of the procedure took approximately 5 min.

3.6.2 *Participant patent space creation*

The experts were given the design problem description along with the first 5 patents (in numerical order, see Appendix A). They were told to first read the design problem and think about it for a few min. They were explicitly asked to try not to solve the design problem *per se*, but to think about the design problem as contextualizing information with which to approach the building of the patent space.

Table 3 Algorithm structure nodes evaluated by experts during interview

Node	# of nodes away from design problem	Patent #	Patent name
1	0	27	Induction loop vehicle detector
		16	Escapement mechanism for pendulum clocks
		18	Earthquake isolation floor
		32	Accelerometer
2	1	5	Interactive toy
		17	Single drive, multi-screw sorter with pusher means
		25	Hanger tilt mechanism for hanging transportation apparatus
3	3	15	Fuel injection system for linear engines
		29	Smoke generating apparatus
		2	Self-cleaning pressure relief and bypass valve, dispensing apparatus and method
		45	Animal waterer

Table 4 Mean expert reactions to each of the algorithm nodes

Participant	Node 1	Node 2	Node 3
E1	1.00	1.00	1.00
E2	1.00	0.25	1.00
E3	1.00	0.00	1.00
E4	1.00	1.00	0.75
Mean "Sensibility"	1.00	0.56	0.94
% Outright Yes	100%	50%	75%

1 = Outright Yes, 0.75 = Yes, but there's an odd one out, 0.25 = Ok, if you insist, but I really wouldn't put them together unless I had no choice, and 0 = Outright No

After this, experts were told to read each patent, look over the associated key patent figure, and place a sticky note on it, writing down key functions or sub-functions of the patent as described. This functionally focused note was meant to serve as a quick reference for them as they built and iterated on the patent space, as well as to remind them to focus on the functionality of the patents as they are placing them in the space. After these six items were read and processed, the experts were videotaped placing the patents on the large piece of paper on the wall. They were given markers and told that they could annotate the structure as they progressed, and should feel free to modify the placement of the patents or the content of the sticky notes at any point in the patent space building process.

After placing the design problem and first five patents in the space, the experts were given the remaining 40 patents in sets of 10 (in study index number order as shown in Appendix A), with a total of four sets of 10. Each set of 10 patents

Table 5 Summary of expert reactions to algorithm's distance orderings

Participant	Summary of response	Representative quotes
E1	Makes sense	*The ordering of it makes sense to me, the order of how far away they are.*
E2	Overall relative ordering makes sense, although he would like to split node 1 into 2 and make part of it (16 and 18) a little further out, but still closer than node 2.	*But I would actually almost do it like, these two, 32 and 27, then these two 16, 18. And then those [group 2], your other group, and then those [group 3]*
E3	Ordering of 1 and 3 makes sense (doesn't like 2, so doesn't want to talk about it)	*Um, so yeah that makes sense, although this one [group 2], I, like I said. So like if this one [group 1] were, if this were in the upper left hand corner, that's sort of your origin or your center. You know, I'd probably put this [group 3], you know, closer to that than I would put this, 'cause I just don't know what the heck this, this group [group 2] means.*
E4	Makes sense	*Ok, I understand why that is*

Table 6 Summary of expert proposed labels and LSA generated cluster labels

Participant	Node 1	Node 2	Node 3
E1	Motion dampening related to magnetic fields	Applying small forces to create larger movements	Fluid movement and control
E2	Energy transfer, energy capture	Complex mechanical systems	Hydraulic output control
E3	Motion conversion	N/A	Fluid transfer
E4	Oscillatory motion as input	Rotation as input	Dispensing fluid (in a discrete amount)
Highest Average Rank	Mean, show, operate, present, cause	Part, rotate, shaft, drive, mechanism	Valve, open, provide, pressure, connect
Highest Cosine	Coil, anchor, pallet, escape, arbor	Product, toy, sort, screw, cam	Fuel, valve, liquid, diaphragm, combustion

was read, processed, and placed on the wall before proceeding to the next set of 10. The placement of the patents on the wall was videotaped for all sets. Once finished, the experts were told to take a few min to ensure they were satisfied with their placement of patents in the space, and to move them around if necessary until the placement was found to be satisfactory. The patents were processed in the same order across experts to eliminate any effect order of exposure or processing of the patents might have on their creation of the patent space. Overall, this phase of the procedure took approximately 1 h and 15 min.

3.6.3 Participant explanation of structure

On video, the experts were then asked to walk the experimenters through the structure, explaining any groups that emerged, the logic behind the organization of the space and regions of patents within the space, and the significance of distance of patents from each other or from the design problem. They were asked to write down key words or labels on the paper itself describing their regions, or clusters of patents, as they talked through them. In addition, the experimenters asked the experts to draw lines indicating borders, clusters, and flows, etc. where appropriate to clarify their patent space and the meaning behind it. This phase of the procedure varied between experts, but took approximately 10 min, on average.

3.6.4 Participant interview regarding algorithm output structure

The experts were then asked to examine the structure that constituted the output of the algorithm (see Figure 2); the method used to generate this structure is described in Section 3.5. This structure examination was performed systematically for three different nodes, and audio and video recordings of these interviews were made. The design problem is in the upper left corner, and indicated by the shaded circle. The nodes examined were the upper left corner,

containing patents 32, 18, 16, and 27; the left middle, containing patents 5, 17, and 25; and the bottom middle left, containing patents 29, 2, 45, and 15. Experts were provided with half sheet descriptions of the patents in each of the three nodes, including the title, abstract, and one key figure, and were asked the following questions for each of the three nodes, examining one node at a time:

Node Grouping Evaluation Questions (asked regarding three separate nodes)

The algorithm puts patents in the same node that are closely functionally related. Does it make sense to you that these patents are in the same node? What key functional relationships do you think make them go together? Are any of these relationships ones you hadn't thought of before? Does the grouping of these patents yield any fresh insights or ideas that you think might be useful for the design problem?

Questions Regarding Reactions to Distance of Clusters from Design Problem

The algorithm places nodes with respect to the design problem based on potential relevance. So, the patents in Node 1 (distance 0) are supposed to be pretty similar/relevant to the design problem, followed by the patents in Node 2 (distance 1) and then Node 3 (distance 3). Does this node placement make sense? Why or why not?

General Reactions Question

Do you have any overall impressions about this structure with respect to its usefulness for supporting creative design ideation? If you were to use a structure like this for design by analogy ideation, would you find it useful to have labels describing the functional similarity of the patents within the clusters?

This phase of the procedure also varied between experts, but on average, the interview phase lasted 15–20 min.

4 Results

The data collected in this study was very rich, allowing for a number of ways to examine the outcomes. The main analyses presented here include a quantitative comparison of the structures generated by all experts to one another and to the algorithm's output structure, a comparison of expert-suggested cluster labels to LSA-generated cluster labels, and a summary of key patterns in the audio and video data. The patent spaces generated by each expert are included in Appendix B.

4.1 Pairwise associations of patents

The patent spaces or "structures" that were created by the experts were diverse in arrangement, meaning, and format. The clear commonality between the participant structures and the algorithm's output structure is the clustering or grouping of patents. Thus, to examine the degree to which experts (and the algorithm) agreed on patent clustering, we examined correlations between structures at the pairwise-link level. Specifically, for all 990 possible pairwise links between the 45 patents, we noted whether each expert (and the algorithm) posited a link or not. Table 1 shows that experts 1–3 had statistically significant similarities to each other at the pairwise-linking level; however, correlations are small, in part reflecting the correlation analysis's correction for chance agreement (which is high in binary-to-binary correlations, and especially high when the ratio of non-links to links is very large), and in part demonstrating substantive variation among experts' representations of the space. Some clarity on the reason for these variations is proposed in Section 4.5.2, in which the analyses of video and audio data are presented. Notably, the algorithm correlated significantly with all experts, including E4. This is important to highlight, as the lack of correlation between E4 and the other 3 experts suggests that this expert was thinking differently about the interrelationships between the patents and the design problem, and yet the algorithm was still able to incorporate some of that perspective into its structure.

From the perspective of developing a tool that maps a space of patents and enables automated search for potentially relevant and inspiring patents, correctly *not* linking two patents that experts generally agree should *not* be linked is equally as important as correctly linking two patents that experts agree should be linked. A high rate of spurious links would render the tool practically useless, as any query for potentially inspiring patents within the structure would return a deluge of results that included a large number of irrelevant patents. Furthermore, as there seems to exist nontrivial variation in terms of expert representations of patent spaces (as indicated by the statistically significant but small inter-correlations), a useful tool should probabilistically track the degree of overlap between expert representations, such that patent links that at least one expert proposed should be proposed by the tool with a low but nonzero probability, and patent pairs that more experts agree are linked should be linked by the tool with a higher probability.

To formally test whether our algorithm not only correctly proposed expert-agreed-upon links, but also correctly rejected expert-agreed-upon *non*-links, *and* whether the algorithm's link assignments probabilistically tracked the degree of agreement among experts for links, we fitted a logistic regression model predicting the probability of the algorithm linking any two patents, with *expert links* as the predictor, where *expert links* indexed the degree of agreement among experts on a given link, ranging from a minimum of 0 if no expert

judged the pair as linked, and a maximum of 4 if every expert judged the pair as linked. The results of this analysis are shown in Table 2. The overall model was statistically significant, $\chi^2(1) = 70.75$, $p < .000$, and a Hosmer and Lemeshow test of deviations from the logistic model turned out nonsignificant, $\chi^2(1) = 0.31$, $p = .58$, indicating a model with adequate fit.

As shown in Figure 5, for patent pairs for which *no* experts proposed a link, the algorithm proposed links for less than 10% of those pairs, the proportion of algorithm-proposed links increased monotonically as a function of the number of expert-links, and the algorithm found 100% of the links that all four experts agreed upon. This function is quantified in the results of the logistic model, which estimated that, with each additional expert judging a patent pair as linked, the probability of the algorithm linking the pair more than doubles ($\exp(B) = 2.53$; see Table 2).

Taken together, the results of the pairwise patent link analysis point to the idea that the algorithm clustering of patents can be judged as "common ground" or consensus among experts within this study, both in terms of correctly grouping patents together that experts agree should be grouped, *and* correctly *not* grouping patents together that experts would agree should not be grouped. Importantly, the algorithm also appears to be able to incorporate an "outlier" perspective in its grouping of the patents, as evidenced by the algorithm's link correlations with E4, despite a lack of significant correlations between that expert and the other experts.

4.2 Expert reactions to algorithm output structure

The interview of the experts' reactions to the algorithm output structure was comprised of two parts: 1) evaluating the sensibility of the clustering for each node, and 2) evaluating the sensibility of the ordering of node distances
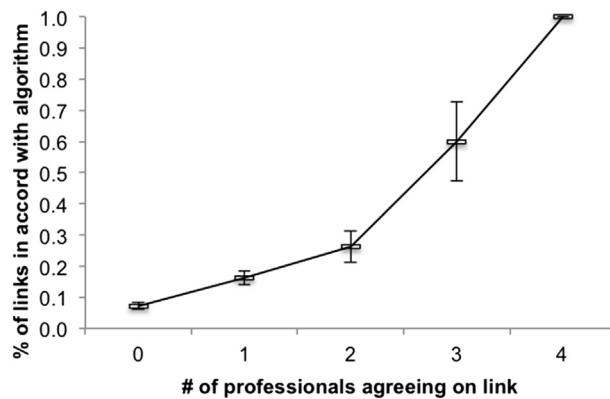

Figure 5 Proportion of patent pairs linked in algorithm output structure by number of experts judging pairs as linked.

from the design problem. The condensed findings for these two interview phases are presented next.

4.2.1 Expert reactions to sensibility of clustering

The expert reactions to the algorithm's structure were obtained as they examined three separate nodes in detail, at varying distances from the design problem. Table 3 shows the contents of each node, along with their distances from the design problem. Table 4 shows the summary of expert reactions to the clustering of patents in the three nodes that were examined. The expert judgments were quantified by the first two authors on a scale from 0 to 1, where 1 = "Outright Yes", 0.75 = "Yes, but there's an odd one out", 0.25 = "Ok, if you insist, but I really wouldn't put them together unless I had no choice", and 0 = "Outright No". In addition, the proportion of experts that reported that the clustering "made sense" or not was calculated, also presented in Table 4.

These results indicate that, overall, the experts were generally satisfied with the clustering created by the algorithm output structure of patents. Node 2 was the only cluster of which the participants were not in majority agreement with the algorithm clustering. This may be due to different conceptualizations of the design problem space, or attending to different features in the patents themselves when considering their clustering. Further discussion of these possible explanations is presented in Section 4.5. Overall, however, the clustering of patents within the nodes examined was affirmed by the experts, and we emphasize that only one expert outright rejected the clustering in Node 2 as invalid.

4.2.2 Expert reactions to sensibility of node distances

The second portion of gathering expert reactions to the algorithm output structure was gauging their feelings about the placement of the three clusters being examined in the structure with respect to one another and to the design problem location, indicated by the red circle in Figure 2. As summarized in Table 5, the experts generally found the node distances to make sense. Participant 3, as seen in the previous section, did not agree with the clustering of Node 2, and thus did not want to discuss its location in the structure. Participant 2 would have liked to see Node 1 further subdivided into two smaller clusters, grouping patents 32 and 27 in one, and 16 and 18 in the other. Overall, the locations of the nodes examined within the structure, both relative to one another and to the design problem were affirmed as well.

4.3 Exploratory quantitative comparison of distance orderings

To further explore the degree to which the algorithm's distance orderings were valid, we undertook an exploratory quantitative comparison of the distance orderings between the algorithm and the experts' structures. Unfortunately, 3 of the 4 experts did not appear to have an explicit overall similarity space

that organized their structures above the cluster level, and also did not explicitly represent distance of clusters to the design problem. E1 said he was not sure how any of the specific patents were related to the design problem, and did not talk about the relative relevance of each of his clusters to the design problem; E3 did include a "not interesting" category of patents (i.e., patents unlikely to be relevant to the design problem), but the overall *spatial* ordering of his structure was governed more in terms of a functional flow structure than by relevance to the design problem; E4 explicitly opposed any notion of ordering of relevance, contending that all clusters were potentially relevant. Only E2's structure bore evidence of spatial ordering of relevance to the design problem, both in the structure itself, and in the verbal content of his explanation of the structure. It is unclear to us why only E2 was willing to spatially order the patents by relevance — perhaps further studies with more experts might uncover whether this is due to timing (e.g., a general reluctance of expert designers to prune out stimuli as irrelevant in early stages of ideation), scale (e.g., greater willingness to order by relevance with more stimuli), or some other reason.

Nevertheless, we thought it would be informative to examine the correlation between the distances of patents from the design problem in E2's structure as compared to the algorithm's structure, while bearing in mind that inferences made from this exploratory analyses should be made with care.

To measure distances, the following procedure was used. First, an image of E2's structures was printed out on letter sized (8.5 × 11 in) paper. Next, dot markers were drawn by one of the authors at the centroid of each patent cluster. Finally, a standard metal ruler was used to measure the distance of each dot from the middle of the design problem. Measurements were in cm, in increments of 0.1 cm. Each patent in the cluster received the distance value of the centroid dot from the design problem. For example, if patents 1 and 4 were in cluster 1, and the distance of the centroid of cluster 1 from the design problem was 6 cm, then patents 1 and 4 were regarded as being 6 cm from the design problem. In the algorithm's structure, each patent's distance from the design problem was the number of nodes (shortest path) from the patent's node location to the design problem.

There is a pragmatic and a theoretical motivation for assigning distances at the cluster level rather than the patent level. Pragmatically, we had reason to believe that the experts did not pay much attention to distances at the patent level, due to constraints (either perceived or real) of the structuring procedure. For example, our instructions asked the experts to (if possible) avoid covering the titles/numbers of patents, which forces a minimum distance on the patents inside a cluster. Theoretically, this cluster-level way of measuring distance is closer to how the algorithm does so (distance from design problem defined at the node level).

In this exploratory analysis, a statistically significant and fairly sizable correlation was found between the distances of E2 and the algorithm (Pearson $r = 0.55$, $p < 0.01$) (Figure 6).

This finding provides a useful complement to the interview data, given the relatively large effect size and explicit examination of the agreement in distance orderings for all 45 patents (as opposed to the 11 in the qualitative reactions in the previous section), and lends strength to our conclusion that the algorithm achieves some success in organizing patents in the space in a way that is meaningful and sensible to experts.

4.4 Cluster label analysis

The experts were also asked to propose functional labels for each node. Table 6 shows the labels they proposed, in addition to the labels generated by the two cluster label generation methods presented in Fu et al. (2013), Highest Average Rank and Highest Cosine Similarity. In addition, the full cluster label set for both labeling methods are presented on the algorithm output structure in Figures 2 and 3. On the surface, there seems to be a fair bit of overlap of the text within the labels suggested by the experts for Node 3, to a slightly lesser extent with Node 1, and much less with Node 2. The experts tend to agree that Node 3 is a cluster of patents related to controlling or moving fluids. The Highest Average Rank method for cluster label generation yielded labels that indicate that fluid is the topic area common to the patents within Node 3, but also the functionalities or components that are employed in the patents, like "connect" and "valve." The Highest Cosine Similarity method of generating cluster labels yielded words that are on topic with fluid applications, but much more specific in terms of actual processes or components, such as "combustion" and "diaphragm."

Experts suggested labels for Node 1 that mostly are related to motion, with one expert suggesting more energy-related labels. Similarly, the Highest
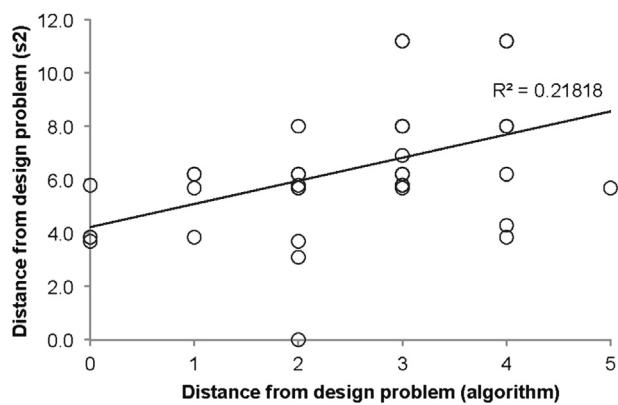

Figure 6 Scatterplot of patent distances in E2's vs. algorithm's structures.

Average Rank labels for Node 1 seem to be general functional terms, not specific to motion or energy capture; and the Highest Cosine Similarity labels, like those for Node 3, are topical to the specific content of the patents within the node — though they reference mechanisms that would be useful when trying to capture energy from motion, like "escape" (as in escapement) and "coil" (as in induction).

The expert-suggested labels for Node 2 are diverse in the specific words chosen, though they all relate to mechanical systems (note, node 2 labels from participant 2 were erroneously not collected). Correspondingly, the Highest Average Rank labels make clear this common thread, with function and component labels that tie the expert labels together, like "drive" and "mechanism" and "shaft." The Highest Cosine Similarity labels again contain words that point more toward the specific patents in the cluster, like "toy."

In general, as exhibited by the detailed node analysis and as can be observed in Figures 2 and 3, the Highest Average Rank method tends to yield labels that are more general, indicating the common functionality or components explaining the association of the patents within the cluster. In many cases, however, and particularly with nodes that have only one patent in them, the two methods of cluster label generation yield the same results. The Highest Cosine Similarity method, on the other hand, tends to yield labels that are very specific to the content of the individual patents within the cluster, although sometimes containing words that may only pertain to one or two patents in the cluster. It should be noted that the patent text was not part-of-speech tagged for patents 16, 18, 27, and 32, as these were hand picked patents from the previous design by analogy study presented in Chan et al. (2011), which may be a reason for the less descriptive label output from the Highest Average Rank labeling method for Node 1.

## 4.5 Explaining inter-expert and expert-algorithm differences

4.5.1 Similarities in patent function labels
Why did expert structures not completely agree with one another? Did they agree what the categories of clusters should be but made somewhat different decisions about which patents below inside each cluster? An analysis of the overlap among experts between function labels given to each patent can provide some insights. Because of synonyms and near synonym relationships among words, expert—expert similarity between labels should be determined along a semantic overlap dimension rather than exact match dimension. The cosines between words in the LSA vector word representation used by the computational algorithm can serve as an approximation of semantic overlap between words (Landaurer & Dumais, 1997).

Figure 7 plots the mean cosines across patents for three comparisons: (1) between experts' function labels for each patent (1st column), (2) between
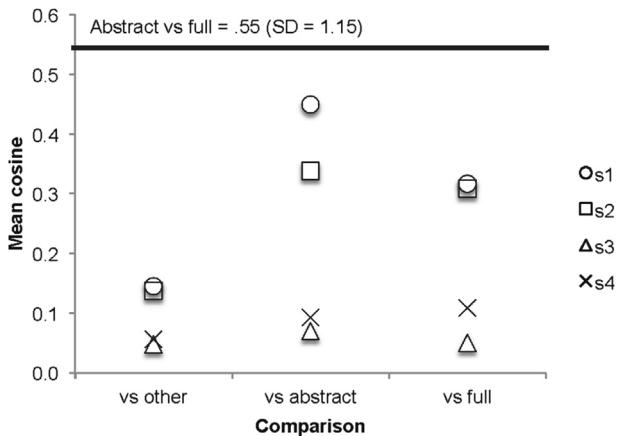
Figure 7 Mean semantic cosine similarity of expert functional patent labels compared to other experts, patent abstracts, and patent full text.

experts' function labels and abstracts for each patent (2nd column), and (3) between experts' labels and full-text for each patent. The mean cosine between abstract and full-text across patents is used as a baseline measure against which to compare the mean cosines. This choice of baseline reflects our assumption that, in situations where a patent is transformed into a different, potentially sparser representation (e.g., moving between different experts' representations of a patent, summarizing patent in an abstract), the transformation from patent full-text to abstract would be a sensible "gold standard" with respect to degree of similarity between transformations, to the extent that patent writers attempt to create a reasonably comprehensive and accurate summary of the patent in the abstract.

The first comparison column (vs. other) gives the mean cosine of each expert against the other experts, averaged across experts for each patent, and then averaged across patents. If the experts in general were attending to similar aspects of each patent, one would expect their mean cosine with each other for each patent to at least approximate the mean cosine between the abstract and full-text for that patent. The figure clearly shows that this is manifestly not the case: on the contrary, the cosine values clearly indicate that the average similarity between experts is dramatically lower than the average abstract–full-text similarity.

The second comparison column (vs. abstract) gives the mean cosine of each expert's functional representation of each patent with the patent's abstract, averaged across patents. The figure clearly shows a range of similarities with the abstract, suggesting divergent approaches to representing each patent, perhaps with some experts attending to functional aspects of each patent in a more holistic, high-level manner (E1 and E2), and others attending to

much more focused, detailed aspects of each patent (E3 and E4). The third comparison column yields a similar insight.

Overall, the authors interpret these data to suggest that the divergence between experts in the overall structures stems at least in part from their attending to different aspects of each patent. In support of this interpretation, the inter-expert cosine for each patent was significantly positively correlated with both average expert vs. abstract ($r = 0.49$, $p < 0.01$) and expert vs. full-text cosines ($r = 0.35$, $p < 0.05$), suggesting that the degree to which, on average, the experts were attending to the patents in a holistic manner (i.e., in accord with the full-text or abstract-level summary of the patent, rather than focusing on specific aspects of the patent apart from its core functionality) was a key determinant of the average similarity between experts. Considering the average complexity of patents, this behavior of the experts is not unexpected.

4.5.2 Protocol video and audio interview analysis

It is clear not only from qualitative analysis of the structures, but also from the correlations shown in Table 1, that the structures created by the experts are diverse. The video and audio explanation of the structures both during the creation and during the "walk through" shed light on the reasons for this variability. Some representative quotes from the interviews with the experts at different stages in the experiment are listed as follows. The common thread among these expert quotes is that they all offer some indication that the framing of the space prior to working with the patents, or contextualizing the space with the design problem itself, leads to the experts building the space, categorizing the patents, and attending to the content within the patents in different and highly influenced ways. The quotes have been edited to remove speech disfluencies.

> "Well I guess in general I thought about people, people moving around, and people moving things around and those are things that, those are activities that, that happen, particularly, in the third world there's a lot of mechanic-, use of the body for labor. So, so when I thought about that, I thought about well how do you attach things to the body, and I thought about that. Let's see… And then also how do you attach things, how do you attach a cart, or connect a cart, and so you might be carrying water or something. And you might as well carry it on a cart, and it has wheels on it, and maybe those wheels spin, and there's a small amount of generation of electricity from that, that rotation of the wheels. And then, you're pushing a cart or you're pulling a cart. So these are all related to the collecting power from a cart" (E3, while explaining structure)

> "It depends what path you were going down. If you said I'm doing something mechanical and you're describing the escapement mechanism, then that's fine. But otherwise I'm not sure. Again, I would go more with bicycle and

gears or bicycle and rotary motion." (E2, while evaluating node distances in algorithm output structure)

In addition, there was evidence that the experts experienced significant cognitive load when performing the task, indicating that the automated generation of structures would be beneficial in terms of removing the expense of energy and time from the designer during patent exploration and organization. Some representative quotes are below.

"Some of it, I think, some of it is wherever your starting filter is, how you're going to sort these things, and what my groupings are. You've already, once you start to choose those, you've headed down a path, so if you forced a, had some kind of forced starting grouping methodology or some kind of grouping filter" (E1, while evaluating clustering in algorithm output structure)

"If I could see all the groups, maybe I'd put it in a different, different kind of sort." (E4, while evaluating node distances in algorithm output structure)

5 General discussion

The research questions underlying this work were the following:

1. How is the algorithm's patent structure similar to or different from how an expert designer might structure a space of patents in preparation to be explored for inspiration through design by analogy with regard to a specific design problem?
2. If the structure is different, can experts nevertheless make sense of it and perhaps see new connections between patents?

Our central finding was that the algorithm output structure of patents was a good representation of the consensus of experts (to the extent that such consensus exists) regarding the associations and clustering of patents in the space. Specifically, the pairwise links generated by the algorithm correlate significantly with all experts' links, even with E4, whose links did not correlate with the other experts, thereby capturing a diversity of expertise, which is important in the increasingly diverse and interdisciplinary nature of engineering design (Reich & Shai, 2012). Additionally, our logistic regression analysis demonstrated that the algorithm was able to probabilistically track the degree of expert consensus on patent pairwise links, pairing patents experts agreed should be paired, *not* pairing patents experts agreed should *not* be paired, and assigning links with nonzero probabilities as a function of degree of expert agreement in areas of potential disagreement.

From the verbal accounts given by the experts, both during their working with the patents to create their own patent space, and during the interview

regarding the algorithm-generated structure, some additional key findings emerged. First, the general consensus with regard to the algorithm output structure clustering is that it makes sense, and in some cases, was both sensible and surprising — indicating that the output from the algorithm is intuitive to expert designers, and can potentially even help them conceptualize the space in new and useful ways, indicated by some interview data. Second, the node distances with respect to the design problem and the interrelationships between the clusters examined were found to make sense to the experts. The exploratory quantitative comparison of distance orderings in Section 4.3 converges with this finding, showing a significant correlation between the node distances created by E2 and the node distances generated by the algorithm. This strong overlap is of particular interest, as it indicates that the organization of the overall space itself can be meaningful, not just the clustering of patents — motivating the use of structures for exploring patents as opposed to simple categorization or grouping methods.

Thus, the answer to the research questions posed is that the algorithm did produce similar *and* different patent spaces when compared to expert thinking, but in beneficial ways for both outcomes. The algorithm-generated patent space was *similar* to expert thinking in that it was strongly correlated with their common ground, and found by the experts to be logical and sensible in clustering, node distances, and node labels. The algorithm-generated patent space was *different* from expert thinking in that the clustering of patents by functional similarity was not highly correlated with any individual expert, and found to be surprising at times (based on interview data), while still being sensible to the experts. Most importantly, though, the experts did find the results meaningful.

Returning to the potential outcomes outlined in Figure 1, we interpret the results of this work to be a combination of outcome 1, **the algorithm does what an expert would be able to do** (in this particular case, while further studies using larger expert pools are needed), **but more *efficiently***, and outcome 2, **the algorithm may structure the analogical stimuli in a way that potentially allows for novel connections to be made between patents.** The algorithm-generated patent space was similar to the consensus of the experts' understanding and representation of the patent space, making it much more efficient than experts in creating a meaningful patent space. However, the algorithm-generated patent space was also meaningfully different from each individual expert's representation of the patent space, and found at times to be surprising to individual experts while still being sensible. This suggests that designers who traverse the algorithm's structures may be able to discover novel connections between patents.

From a practical standpoint, the algorithm output structure offers a more objective conceptualization of the space, while saving the designer time

and energy. The experts reported that the task of organizing the patents into the space was one in which it was difficult not to get locked into a representation early on, and keeping track of all of the information and aspects was challenging. This indicates that the automated patent structuring achieved by the methodology presented in Fu et al. (2013) could provide significant advantages to designers by presenting them with useful, sensible, and even sometimes surprising representations of patent spaces while saving them significant mental effort and time. It took the algorithm less than 45 min to optimally generate 8 different structures, running on a Windows 64-bit dual core desktop computer, while it took the experts approximately 1.25 h to generate one representation of the patent space. As the number of patents grows, the time for the human would increase likely in a non-linear way due to fundamental limitations with human memory (Cowan, 2000). Perhaps more importantly, it is unlikely that a designer would take the time to structure large numbers of patents by hand, even though this work has shown it potentially valuable. The algorithmic structuring of patents could be viewed as operating like another expert, providing a view that overlaps with that of other experts, while being somewhat different as well.

6 Conclusion

From an inspirational standpoint, the algorithm output structure offers a conceptualization of the space that is sensible to four expert designers, while still maintaining an element of surprise and unexpected representation of the space, making it a promising basis for a computational design by analogy inspiration tool. By allowing for more efficient and insightful access to external analogical stimuli, designers have the potential to create more innovative design solutions. With a way to extract the interrelatedness and interconnectedness of patents in the space, designers might be able to strategically choose which cross-domain designs to expose themselves to, or even traverse the space in a more intentional and meaningful exploratory way. Additionally, by making novel connections between patents, designers might be able to discover novel solutions via altered representations of the problem or design space. Explicit tests of such effects should be explored in future work.

Appendix A

Patents used in algorithm generated structure

Study index number	US patent number	Patent title
1	6,082,923	Converging Sphere Joint Assembly[b]
2	5,984,148	Self-Cleaning Pressure Relief and Bypass Valve, Dispensing Apparatus and Method
3	5,375,948	Cutting Insert for Cutting and Grooving Tools
4	6,367,521	Gravity Feed Fluid Dispensing Valve
5	6,497,607	Interactive Toy
6	4,853,977	Patient Garment
7	5,993,410	Adjustable Probe
8	4,223,996	Apparatus for Mixing Solid and Liquid Constituents of Mortar or The Like
9	4,589,668	Wheeled Cart with Removable Skis
10	3,962,735	Movable Bulkhead with Guiding and Overcanting Prevention Means
11	4,124,051	Shock Absorbing Wheel Hub
12	4,984,583	Air Bubbling Mats for Therapeutically Agitating Bath Water
13	7,175,212	Latch Having Releasable Cinching Mechanism
14	4,259,034	Bale Handling Apparatus
15	6,634,325	Fuel Injection System for Linear Engines
16	4,139,981	Escapement mechanism for pendulum clocks[a]
17	5,909,815	Single Drive, Multi-Screw Sorter with Pusher Means
18	4,402,483	Earthquake isolation floor[a]
19	4,103,708	Ventilated Poppet Damper
20	3,964,473	Bone Prosthesis
21	4,705,064	Safety Seal for an Operating Lever
22	6,142,689	Envelope Leveler for Printer Feeder
23	5,273,173	Screw Top
24	5,438,724	Method for Using Plastic Fasteners for Shoe-Lasting Applications
25	5,234,096	Hanger Tilt Mechanism For Hanging Transportation Apparatus[b]
26	4,867,134	Fluid-Heating Solar Collector
27	4,568,937	Induction loop vehicle detector[a]
28	6,109,282	Self-Erecting Loop Structure
29	4,303,397	Smoke Generating Apparatus
30	5,899,571	Beach Towel, Tote Bag and Beach Umbrella System
31	6,234,452	Hand Operable Motorcycle Stand
32	4,335,611	Accelerometer[a]
33	5,085,240	Shelter Structure[b]
34	6,634,044	Compact Stretcher
35	4,270,310	Support Device for an Upstanding Plant Support Rod in a Plant Pot
36	5,423,097	Emergency Drop Fowler and Gatch
37	5,277,276	Compensating rope sheave tie down[b]
38	3,938,909	Single Needle Alternating Flow Blood Pump System
39	5,647,066	Safety Helmet Visor
40	6,119,041	Apparatus and Method for Linear Lesion Ablation
41	4,484,762	Ski Binding and Boot
42	4,762,262	Side-Fed Stapler
43	6,164,698	Steering Device for Automobiles
44	6,062,856	Dental Implant Hole Guide Extension
45	4,739,727	Animal Waterer

[a] Far Analogical Stimuli from Chan et al. (2011).
[b] Far Analogical Stimuli from Fu et al. (2012).

Appendix B

EXPERT GENERATED PATENT SPACES
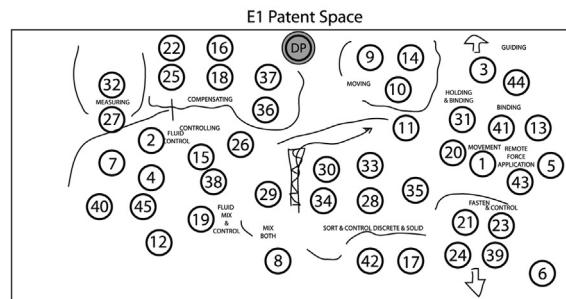
E1 Patent Space
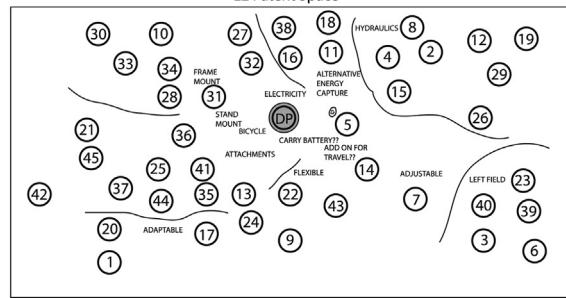
E2 Patent Space

E3 Patent Space
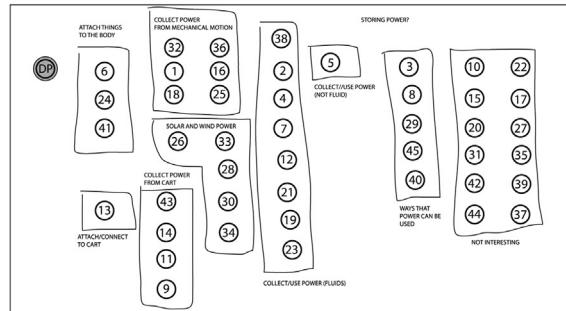
E4 Patent Space
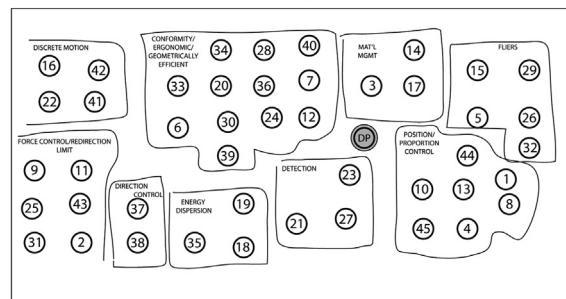
References

Ahmed, S., & Christensen, B. T. (2009). An in situ study of analogical reasoning in novice and experienced design engineers. *Journal of Mechanical Design, 131*(11), 111004.

Altshuller, G. S., & Shapiro, R. B. (1956). О Психологии изобретательского творчества. [On the psychology of inventive creation]. *Вопросы Психологии (The Psychological Issues), 6*, 37−39, (in Russian).

Anderson, J. R. (1991). The adaptive nature of human categorization. *Psychological Review, 98*(3), 409−429.

Ball, L. J., & Christensen, B. T. (2009). Analogical reasoning and mental simulation in design: two strategies linked to uncertainty resolution. *Design Studies, 30*(2), 169−186.

Bhatta, S., & Goel, A. (1996). From design experiences to generic mechanisms: model-based learning in analogical design. *AIEDAM, 10*(2), 131−136.

Bohm, M. R., Vucovich, J. P., & Stone, R. B. (2005). *Capturing creativity: Using a design repository to drive concept innovation*. Paper presented at the Proceedings of DETC2005, DETC05/CIE-85105, Long Beach, California.

Bradley, R. A., & Terry, M. E. (1952). Rank analysis of incomplete block designs. 1. The method of paired comparisons. *Biometrika, 39*, 324−345.

Carroll, J. D. (1976). Spatial, nonspatial and hybrid models for scaling. *Psychometrika, 41*, 439−463.

Casakin, H., & Goldschmidt, G. (1999). Expertise and the use of visual analogy: implications for design education. *Design Studies, 20*(2), 13−175.

Chakrabarti, A. K., Dror, I., & Nopphdol, E. (1993). Interorganizational transfer of knowledge: an analysis of patent citations of a defense firm. *IEEE Transactions on Engineering Management, 40*(1), 91−94.

Chakrabarti, S., Dom, B., Agrawal, R., & Raghavan, P. (1998). Scalable feature selection, classification and signature generation for organizing large text databases into hierarchical topic taxonomies. *The VLDB Journal, 7*(3), 163−178.

Chan, J., Fu, K., Schunn, C., Cagan, J., Wood, K., & Kotovsky, K. (2011). On the benefits and pitfalls of analogies for innovative design: ideation performance based on analogical distance, commonness, and modality of examples. *Journal of Mechanical Design, 133*(8). http://dx.doi.org/10.1115/1.4004396.

Chase, W. G., & Simon, H. A. (1973). The mind's eye in chess. In W. G. Chase (Ed.), *Visual information processing* (pp. 215−281). New York, NY: Academic Press.

Cheong, H., Shu, L. H., Stone, R., & Wood, K. L. (2008). *Translating terms of the functional basis into biologically meaningful keywords*. Paper presented at the Proceedings of the ASME International Design Engineering Technical Conference, New York, NY, USA.

Chi, M. T. H., Feltovich, P. J., & Glaser, R. (1981). Categorization and representation of physics problems by experts and novices. *Cognitive Science, 5*(2), 121−152.

Chi, M. T. H., & Koeske, R. D. (1983). Network representation of a child's dinosaur knowledge. *Developmental Psychology, 19*(1), 29−39.

Chiu, I., & Shu, L. H. (2005). *Bridging cross-domain terminology for biomimetic design*. Paper presented at the ASME IDETC/CIE, Long Beach, CA, USA.

Christensen, B. T., & Schunn, C. D. (2005). Spontaneous access and analogical incubation effects. *Creativity Research Journal, 17*(2−3), 207−220.

Christensen, B. T., & Schunn, C. D. (2007). The Relationship of analogical distance to analogical function and preinventive structure: the case of engineering design. *Memory & Cognition, 35*(1), 29−38.

Collins, A. M., & Quillian, M. R. (1969). Retrieval time from semantic memory. *Journal of Verbal Learning and Verbal Behavior, 8*(2), 240−247.

Cowan, N. (2000). The magical number 4 in short-term memory: a reconsideration of mental storage capacity. *Behavioral and Brain Sciences, 24*, 87−185.

CREAX. *CREAX: Creativity for innovation.* Retrieved 07.09.12, from http://www.creax.com.

Dahl, D. W., & Moreau, P. (2002). The Influence and value of analogical thinking during new product ideation. *Journal of Marketing Research, 39*(1), 47−60.

Davies, J., Goel, A., & Nersessian, N. (2009). A computational model of visual analogies in design. *Journal of Cognitive Systems Research, Special Issue on Analogies − Integrating Cognitive Abilities Cognitive Systems Research, 3*, 204−215.

Deerwester, S., Dumais, S. T., Furnas, G. W., & Landauer, T. K. (1990). Indexing by latent semantic analysis. *Journal of the American Society for Information Science, 41*(6), 391−407.

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

Duran-Novoa, R., Leon-Rovira, N., Aguayo-Tellez, H., & Said, D. (2011). Inventive problem solving based on dialectical negation, using evolutionary algorithms and TRIZ heuristics. *Computers in Industry, 62*, 437−445.

Dyer, J. H., Gregersen, H. B., & Christensen, C. M. (2011). *The innovator's DNA: Mastering the five skills of disruptive innovators.* Boston, MA: Harvard Business Review Press.

Fiske, A. P. (1992). The four elementary forms of sociality: framework for a unified theory of social relations. *Psychological Review, 99*(4), 689−723.

Foltz, P. W., Kintsch, W., & Landauer, T. K. (1998). The measurement of textual coherence with latent semantic analysis. *Discourse Processes, 25*(2−3), 285−307.

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

Fu, K. (2012). *Discovering and exploring structure in design databases and its role in stimulating design by analogy.* (Ph.D. Dissertation), Carnegie Mellon University, Pittsburgh, PA, USA.

Fu, K., Cagan, J., Kotovsky, K., & Wood, K. (2013). Discovering structure in design databases through function and surface based mapping. *Journal of mechanical Design, 135*(3), 031006.

Fu, K., Chan, J., Cagan, J., Kotovsky, K., Schunn, C., & Wood, K. (2013). The meaning of "near" and "far": the impact of structuring design databases and the effect of distance of analogy on design output. *ASME Journal of Mechanical Design, 135*(2), 021007.

Gentner, D., Brem, S., Ferguson, R. W., Wolff, P., Markman, A. B., & Forbus, K. D. (1997). Analogy and creativity in the works of Johannes Kepler. In T. B. Ward, S. M. Smith, & J. Vaid (Eds.), *Creative thought: An investigation of conceptual structures and processes* (pp. 403−459). Washington D.C.: American Psychological Association.

Gentner, D., & Markman, A. B. (1997). Structure mapping in analogy and similarity. *American Psychologist, 52*(1), 45−56.

Gick, M. L., & Holyoak, K. J. (1980). Analogical problem solving. *Cognitive Psychology, 12*(3), 306−355.

Gobet, F., & Simon, H. A. (2000). Five seconds or sixty? Presentation time in expert memory. *Cognitive Science, 24*(4), 651−682.

Goel, A., Bhatta, S., & Stroulia, E. (1997). Kritik: an early case-based design system. In M. Maher, & P. Pu (Eds.), *Issues and applications of case-based reasoning in design* (pp. 87−132). Mahwah, NJ: Erlbaum.

Goel, A. K. (1997). Design, analogy, and creativity. *IEEE Expert, 12*(3), 62−70.

Goldfire, I. M. *Invention machine goldfire: Unleashing the power of research.* Retrieved 19.02.12, from http://inventionmachine.com/products-and-services/innovation-software/goldfire-Research/.

Goldschmidt, G., & Smolkov, M. (2006). Variances in the impact of visual stimuli on design problem solving performance. *Design Studies, 27*, 549−569.

Griffiths, T. L., Kemp, C., & Tenenbaum, J. B. (2008). Bayesian models of cognition. In R. Sun (Ed.), *Cambridge handbook of computational psychology* (pp. 59−100). New York, NY: Cambridge University Press.

Guttman, L. (1944). A basis for scaling qualitative data. *American Sociological Review, 9*(9), 139−150.

Guttman, L. (1954). A new approach to factor analysis: The radex. In P. F. Lazarsfeld (Ed.), *Mathematical thinking in the social sciences* (pp. 258−348). New York, NY: Free Press.

Hernandez, N. V., Schmidt, L. C., & Okudan, G. E. (2012a). *Experimental assessment of TRIZ effectiveness in idea generation.* Paper presented at the ASEE Annual Conference, San Antonio, TX, USA.

Hernandez, N. V., Schmidt, L. C., & Okudan, G. E. (2012b). *Systematic ideation effectiveness study of TRIZ.* Paper presented at the ASME IDETC/CIE, Chicago, IL, USA.

Herstatt, C., & Kalogerakis, K. (2004). *How to use analogies for breakthrough innovations.* Working Papers/Technologie-und Innovationsmanagement, 24(11 April 2009). Retrieved from. http://biblioteca.universia.net/html_bura/ficha/params/id/39800023.html.

Hmelo-Silver, C. E., & Pfeffer, M. G. (2004). Comparing expert and novice understanding of a complex system from the perspective of structures, behaviors, and functions. *Cognitive Science, 28*(1), 127−138.

Holyoak, K. J., & Thagard, P. (1996). The analogical scientist. In K. J. Holyoak, & P. Thagard (Eds.), *Mental leaps: Analogy in creative thought* (pp. 185−209). Cambridge, MA: MIT Press.

Houssin, R., & Coulibaly, A. (2011). An approach to solve contradiction problems for safety integration in innovative design process. *Computers in Industry, 62*, 398−406.

Huelsenbeck, J. P., & Ronquist, F. (2001). MRBAYES: Bayesian inference of phylogenetic trees. *Bioinformatics, 17*(8), 754−755.

Huth, A., Nishimoto, S., Vu, A., & Gallant, J. (2012). A continuous semantic space describes the representation of thousands of object and action categories across the human brain. *Neuron, 76*(6), 1210−1224.

Indukuri, K. V., Ambekar, A. A., & Sureka, A. (2007). *Similarity analysis of patent claims using natural language processing techniques.* Paper presented at the International Conference on Computational Intelligence and Multimedia Applications.

Inhelder, B., & Piaget, J. (1969). *The early growth of logic in the child.* New York, NY: W W Norton & Company.

Jaynes, E. T. (2003). *Chapter 1: Plausible reasoning probability theory: The logic of science.* Cambridge: Cambridge University Press.

Kaplan, C., & Simon, H. A. (1990). In search of insight. *Cognitive Psychology, 22*(3), 374−419.

Kasravi, C., & Risov, M. (2007). *Patent mining − Discovery of business value from patent repositories.* Paper presented at the Proceedings of the 40th Hawaii International Conference on System Sciences.

Kemp, C., & Tenenbaum, J. B. (2008). The discovery of structural form. *Proceedings of the National Academy of Sciences of the United States of America, 105*(31), 10687−10692.

Koch, S., Bosch, H., Giereth, M., & Ertl, T. (2009). *Iterative integration of visual insights during patent search and analysis*. Paper presented at the IEEE Symposium on Visual Analytics Science and Technology, Atlantic City, NJ, USA.

Kohonen, T. (1997). *Self-organizing maps*. New York, NY: Springer.

Krasnoslobodtsev, V., & Langevin, R. (2005). *TRIZ application in development of climbing robots*. Paper presented at the First TRIZ Symposium, Japan.

Kurtz, K., Miao, C., & Gentner, D. (2001). Learning by analogical bootstrapping. *Journal of the Learning Sciences, 10*(4), 417−446.

Landauer, T. K., & Dumais, S. T. (1997). A solution to Plato's problem: the latent semantic analysis theory of acquisition, induction, and representation of knowledge. *Psychological Review, 1*, 211−240.

Landauer, T. K., Foltz, P. W., & Laham, D. (1998). An introduction to latent semantic analysis. *Discourse Processes, 25*(2−3), 259−284.

Liang, Y., Tan, R., & Ma, J. (2008). *Patent analysis with text mining for TRIZ*. Paper presented at the IEEE ICMIT, Bangkok, Thailand.

Linsey, J., Markman, A. B., & Wood, K. L. (2008). *WordTrees: A method for design-by-analogy*. Paper presented at the ASEE Annual Conference.

Linsey, J. S., Wood, K. L., & Markman, A. B. (2008). Modality and representation in analogy. *Artificial Intelligence for Engineering Design, Analysis & Manufacturing, 22*, 85−100. http://dx.doi.org/10.1017/S0890060408000061.

Mcadams, D. A., & Wood, K. L. (2000). *Quantitative measures for design by analogy*. Paper presented at the ASME Design Engineering Technology Conference.

McCaffrey, T. (2012). Innovation relies on the obscure: a key to overcoming the classic problem of functional fixedness. *Psychological Science, 23*(3), 215−218.

Mehalik, M., & Schunn, C. (2007). What constitutes good design? A review of empirical studies of design processes. *International Journal of Engineering Education, 22*(3), 519−532.

Morais, A. S., Olsson, H., & Schooler, L. J. (2013). Mapping the structure of semantic memory. *Proceedings of Cognitive Science, 37*(1), 125−145.

Mukherjea, S., Bhuvan, B., & Kankar, P. (2005). Information retrieval and knowledge discovery utilizing a biomedical patent semantic web. *IEEE Transactions on Knowledge and Data Engineering, 17*(8), 1099−1110.

Murphy, J. T. (2011). *Patent-based analogy search tool for innovative concept generation*. (Ph.D. Dissertation), The University of Texas, Austin, TX.

Nakagawa, T. (2012). Creative problem-solving methodologies TRIZ/USIT: overview of my 14 years in research, education, and promotion. *The Bulletin of the Cultural and Natural Sciences in Osaka Gakuin University, 64*.

Nix, A. A., Sherret, B., & Stone, R. B. (2011). *A function based approach to TRIZ*. Paper presented at the ASME IDETC/CIE, Washington, D.C., USA.

Novick, L. R. (1988). Analogical transfer, problem similarity, and expertise. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 14*(3), 510−520.

Reich, Y., & Shai, O. (2012). The interdisciplinary engineering knowledge genome. *Research in Engineering Design, 23*(3), 251−264.

Shepard, R. N. (1980). Multidimensional scaling, tree-fitting, and clustering. *Science, 210*(4468), 390−398.

Sneath, P. H., & Sokal, R. R. (1973). *Numerical taxonomy: The principles and practice of numerical classification*. San Francisco, CA: Freeman.

Stone, R., & Wood, K. L. (2000). Development of a functional basis for design. *Journal of Mechanical Design, 122*, 359−370.

Szykman, S., Sriram, R. D., Bochenek, C., Racz, J. W., & Senfaute, J. (2000). Design repositories: engineering design's new knowledge base. *IEEE Intelligent Systems, 1094-7176/00*, 48−55.

Tseng, I., Moss, J., Cagan, J., & Kotovsky, K. (2008). The role of timing and analogical similarity in the stimulation of idea generation in design. *Design Studies, 29*, 203−221.

Vattam, S., Helms, M., & Goel, A. (2010). A content account of creative analogies in biologically inspired design. *AIEDAM, Special Issue on Biologically Inspired Design, 24*, 467−481. http://dx.doi.org/10.1017/S089006041000034X.

Verhaegen, P., D'hondt, J., Vandevenne, D., Dewulf, S., & Duflou, J. R. (2011). Identifying candidates for design-by-analogy. *Computers in Industry, 62*, 446−459.

Vincent, J. F. V., Bogatyreva, O. A., Bogatyreva, N. R., Bowyer, A., & Pahl, A. K. (2006). Biomimetics: its practice and theory. *Journal of the Royal Society Interface, 3*(9), 471−482.

Weisberg, R. W. (2009). On 'out-of-the-box' thinking in creativity. In K. W. A. Markman (Ed.), *Tools for innovation* (pp. 23−47). New York, NY: Oxford University Press.

Wiggins, J. S. (1996). An informal history of the interpersonal circumplex tradition. *Journal of Personality Assessment, 66*(2), 217−233.

Wilson, J. O., Rosen, D., Nelson, B. A., & Yen, J. (2010). The effects of biological examples in idea generation. *Design Studies, 31*(2), 169−186.

Zhang, R., Cha, J., & Lu, Y. (2007). *A conceptual design model using axiomatic design, functional basis and TRIZ*. Paper presented at the Proceedings of the 2007 IEEE IEEM.