

## Patent stimuli search and its influence on ideation outcomes

Binyang Song<sup>1</sup>, V. Srinivasan<sup>2</sup> and Jianxi Luo<sup>1,3</sup>

1 *Engineering Product Development Pillar, Singapore University of Technology and Design, Singapore*

2 *Instrument Design and Development Centre, Indian Institute of Technology, Delhi, India*

3 *SUTD-MIT International Design Centre, Singapore University of Technology and Design, Singapore*

### Abstract

Prior studies on design ideation have demonstrated the efficacy of using patents as stimuli for concept generation. However, the following questions remain: (a) From which part of the large patent database can designers identify stimuli? (b) What are their implications on ideation outcomes? This research aims to answer these questions through a design experiment of searching and identifying patent stimuli to generate new concepts of spherical rolling robots. We position the identified patent stimuli in the home, near and far fields defined in the network of patent technology classes, according to the network's community structure and the knowledge proximity of the stimuli to the spherical rolling robot design. Significant findings are: designers are most likely to find patent stimuli in the home field, whereas most patent stimuli are identified in the near field; near-field patents stimulate the most concepts, which exhibit a higher average novelty; combined home- and far-field stimuli are most beneficial for high concept quality. These findings offer insights on designers' preferences in search for patent stimuli and the influence of stimulation distance on ideation outcomes. The findings will also help guide the development of a computational tool for the search of patents for design inspiration.

**Key words:** design ideation, concept generation, novelty, patent, network analysis

Received 8 April 2017  
Revised 2 November 2017  
Accepted 2 November 2017

**Corresponding author**  
B. Song  
binyang\_song@mymail.sutd.edu.sg

Published by Cambridge University Press  
© The Author(s) 2017  
Distributed as Open Access under a CC-BY 4.0 license  
(<http://creativecommons.org/licenses/by/4.0/>)

*Des. Sci.*, vol. 3, e25  
journals.cambridge.org/dsj  
DOI: 10.1017/dsj.2017.27

the **Design Society**  
a worldwide community

 **CAMBRIDGE**  
UNIVERSITY PRESS

### 1. Introduction

Design creativity is the ability of an agent to address a design opportunity by developing outcomes that are both novel and useful (Sarkar & Chakrabarti 2011). Concept generation is an early phase in the design process where solution principles are conceived to address design opportunities (Jensen *et al.* 2009; Taura & Yukari 2012). Concept generation is a significant phase of the design process because a successful product is likely to be an outcome of an exploration of a variety of solution principles (Pahl & Beitz 2013). Owing to the ease of making changes that are less expensive in this phase, the scope for design creativity is greater in this early phase than the downstream phases (French 1985). Several guidelines, methods and tools have been proposed to foster creativity during the concept generation phase.

Providing stimuli to designers in order to identify analogies from them for generating concepts is one of the most potent and useful methods (Chakrabarti *et al.* 2005; Chan *et al.* 2011). A stimulus is beneficial for concept generation by helping develop creative solutions, enhance novelty, inhibit fixation, etc. (Qian &

Gero 1996; Goel 1997; Linsey *et al.* 2010; Chan *et al.* 2011). Simultaneously, certain stimuli can also inhibit concept generation by causing bias and fixation (Jansson & Smith 1991). Therefore, stimuli need to be carefully chosen before using them. Several kinds of aids to foster the use of stimuli and analogies have been proposed and found to be effective at improving quantity, novelty and creativity of solutions (Chakrabarti *et al.* 2005; Linsey *et al.* 2010). Prior studies have found that it is easier to analogize with stimuli from near than far domains to the target domain, because stimuli from near domains have more structural similarities to the target design problem than stimuli from far domains (Christensen & Schunn 2007). However, stimuli from far domains, owing to their surface dissimilarities, are the best sources for novelty and creative breakthroughs (Gentner & Markman 1997; Ward 1998). Several researchers investigated the effects of stimulation from near and far analogical distances on the outcomes of ideation, for instance Wilson *et al.* (2010), Fu *et al.* (2013*b*) and Chan & Schunn (2015). However, to date, the characterization of near and far stimuli have been inconsistent in existing studies, and so, their findings cannot be generalized across these studies.

Meanwhile, patents have been increasingly explored as sources of stimuli for engineering design (Fantoni *et al.* 2013; Fu *et al.* 2013*b*, 2015; Murphy *et al.* 2014; Srinivasan *et al.* 2017*a,b*). Patents contain technical descriptions of products and processes, which are both novel and functional, from various domains. The patent database is an enormous reservoir of design precedents. The growing patent data, as inventors continually file patent applications over time, present both opportunities and challenges for using them as design stimuli. Therefore, efficient methods and tools are required for designers to retrieve the most relevant from among millions of patents in the vast patent databases. While it is acknowledged that patents are useful for inspiration, questions persist pertaining to: from where in the complex database can designers find useful patent stimuli, and which of these patent stimuli can most effectively inspire designers to generate novel and valuable concepts. Moreover, due to a lack of uniform characterization of near and far stimuli, there is no single method to characterize stimuli as near or far, and consequently, not much work has been done to identify from where in the patent database can near and far patent stimuli be identified.

As a solution, network analysis techniques have been increasingly exploited to uncover the knowledge structure in the patent database to facilitate engineering design. For example, Fu *et al.* (2013*a*) analyzed the similarities of occurrences of functional verbs between patents to construct the Bayesian networks of patents. Such a network provides information of functional similarity between individual patents, which in turn has potential for a patent recommendation system for design stimulation (Fu *et al.* 2015). At a higher level, the patent classification and citation information have also been analyzed to measure knowledge proximity or distance between different classes of patents and construct technology network maps to approximate the total technology space (Kay *et al.* 2014; Leydesdorff, Kushnir & Rafols 2014; Alstott *et al.* 2017*a*; Yan & Luo 2017). In these network maps, nodes are technology classes that represent various technology categories and contain patents related to corresponding technology categories. These nodes are connected according to the knowledge proximity between them. A structural analysis of the networks can allow one to define and identify the technology classes near or far from a given design problem in the technology space. In this study, we will utilize such a technology space network to locate the patents that

designers found useful in an ideation exercise, according to the proximity between technology classes in the network. Herein, we consider a patent as useful if it is used as a stimulus for concrete concepts generated by designers.

The broad objectives of this research are: (a) to identify locations within the network of technology classes from where designers identify useful patent stimuli and (b) to study the implications of using such patent stimuli for ideation on the outcomes of ideation. Toward these broad objectives, the research in this paper examines the effects of using patents – sourced from technology classes which are located at the home field, near field and far field to a design problem – as stimuli for ideation on the outcomes, based on the data from an open concept generation exercise. The three fields in the technology space are defined based on community detection within the network of technology classes. The home field entails the technology classes that are directly relevant to the design problem, the near field comprises the technology classes that are in the same cohesive network communities as those in the home field, and the far field includes the technology classes in all the other communities in the technology network.

In the following sections, we review prior literature relevant to the theories and methods grounding our research (Section 2), introduce our data and research method (Section 3), present and discuss our findings (Sections 4 and 5).

## 2. Literature review

This study is theoretically motivated and grounded by the literature on design by analogy. Within the field of Design Science, the area of analogical design has been extensively researched. However, to fit the scope of this research, only those prior studies that use patents for stimulation in ideation or analyze the effect of analogical distance on the performance of ideation are reviewed here.

### 2.1. Patent stimuli and design by analogy

Many researchers have studied the use of patents as stimuli for design and developed tools for the search and analysis of patents. For example, several tools have been developed to search for patents to facilitate the use of TRIZ principles (Altshuller & Shapiro 1956) in solving design problems (Mukherjea, Bamba & Kankar 2005; Cascini & Russo 2006; Souili *et al.* 2015) developed the Biomedical Patent Semantic Web for retrieving patents based on the semantic associations between biological terms within the abstracts of biomedical patents. Particularly, a recent strand of research has focused on analyzing and using patents to aid in design by analogy.

Fu *et al.* (2013a) developed a computational tool for automatically identifying patent stimuli at different analogical distances. They extracted verb and noun content from the technical descriptions of patents, used semantic analysis to quantify the functional and surface similarities between patents, and created function- and surface-based Bayesian networks of patents, respectively. In the networks, a design problem can be located as the starting point, and the ‘analogical distance’ between the problem and patents is defined as the length of path between them. Murphy *et al.* (2014) proposed a functional vector approach to systematically search and identify functional analogies from the patent database. The following steps constitute the methodology: (a) process patents to identify a vocabulary of functions, (b) define a set of functions in patents comprising

primary, secondary and correspondent functions, (c) index patents using the functional set to create a vector representation of the patent database, (d) develop methods for generating query and estimate relevance of patents to a query, and (e) retrieve and display patents relevant to the query. Fu *et al.* (2015) empirically tested the functional vector approach of Murphy *et al.* (2014), to aid in the search for functional analogies from patent databases to stimulate design concepts, and found the experimental group generated solutions of higher novelty than the control group. Srinivasan *et al.* (2017a) tested the efficacy of using patents as design stimuli through a concept generation experiment, and found that the average quality and novelty of the concepts generated with patent stimuli individually or in combination with other resources is higher than those generated without any stimuli.

## 2.2. Stimulation distance

Design by analogy leverages existing solutions from source fields to solve design problems in target fields (Gick & Holyoak 1980; Weisberg 2006; Linsey 2007). The distance between the source and target fields is referred to as the stimulation or analogical distance. The Conceptual Leap hypothesis states that stimuli from far sources, owing to their surface dissimilarities, provide the best stimulation for creative breakthroughs (Gentner & Markman 1997; Ward 1998). Some anecdotal evidence exists in support of this hypothesis. However, empirical findings related to the validation of this hypothesis have not been consistent.

Chan *et al.* (2011) observed that far-field analogies help develop concepts of higher novelty, higher variability in quality and greater solution transfer but stimulate fewer concepts than near-field analogies. Chan & Schunn (2015) reasoned that the most creative solutions are more likely to be developed from near distance than far distance stimuli, owing to better perception and connection to the problem at hand. Srinivasan *et al.* (2017b) observed that as analogical distance of patent stimuli from the design problem increases, novelty of concepts generated using these stimuli increases but quality of concepts decreases. However, Wilson *et al.* (2010) observed no distinctions between stimuli from far sources and near sources. Fu *et al.* (2013b) found that stimuli from near sources or 'middle ground' help generate solutions of higher 'maximum novelty' than far sources; no significant differences were seen in 'average novelty' between near and far sources. Fu *et al.* also observed that both the 'mean quality' and the 'maximum quality' of solutions generated using stimuli from near sources are higher than those generated using stimuli from far sources. Consequently, they argued stimuli from 'middle ground' to be more beneficial for developing creative solutions. With these findings, Fu *et al.* (2013b) posited that comparisons of effects of analogical distance across different studies are hard owing to different metrics being used to measure distance in these studies. They also argued about the terms 'near' and 'far' as being relative and not being able to completely characterize these across different studies due to lack of a common metric to measure distance.

## 2.3. Network of technologies by distance or proximity

These prior studies have implied the potential value for designers to make use of the knowledge of the relative distance or proximity between technologies in

the search for design stimuli from either near or far sources. For example, to use patents as design stimuli, the Bayesian network of patents of Fu *et al.* (2013a) quantifies and visualizes the analogical distance between patents and a design problem, and thus designers can potentially use the network to identify patent stimuli from near or far distance from the design problem. However, the network of patents is only applicable for a small set of patents, whereas the total patent database contains millions of patents that may provide varied inspirations from different distances to a design problem.

According to the patent classification systems, such as the International Patent Classification (IPC) system, each patent is classified in one or multiple technology classes, which are categories of patents and represent different technology fields. This presents a structure for locating patents in the enormous database. A few recent studies have proposed methods to measure the knowledge proximity between the patent technology classes and used such proximity information to construct the network map of technology classes (Kay *et al.* 2014; Leydesdorff *et al.* 2014; Yan & Luo 2017). The network of all technology classes in the patent database can be used to approximate the total technology space (Alstott *et al.* 2017a). Such a network of technology classes, given the proximity information, may serve as a framework to define the near or far field of design stimulation. In turn, such a network map will allow the designers to be better informed of the proximity (or distance) between the source field of potential patent stimuli and the target field where a design problem or opportunity is located, or be better oriented to identify patents specifically from either the near or far field from the design problem.

In particular, the key requirement to create such a network is the measure of knowledge proximity between the patent technology classes, i.e., link weight in the network. In the literature, a variety of measures of knowledge proximity have been reported. One group of measures are computed using the data of patent references. For example, Jaccard index can be adopted to calculate the number of shared references of a pair of classes normalized by the total number of all unique references of patents in either class (Jaccard 1901; Small 1973) as an indicator of knowledge proximity. Alternatively, the cosine similarity index can be calculated between two vectors indicating patent references made from the patents in a pair of classes to all classes respectively (Jaffe 1986; Kay *et al.* 2014; Leydesdorff *et al.* 2014), i.e., class-to-class reference vectors. For a higher granularity, Yan & Luo (2017) extended the cosine similarity measure to class-to-patent vectors, concerning references to specific patents instead of aggregated classes. Another group of measures use the 'co-classification' information, i.e., how often two classes are co-assigned to individual patents, to compute knowledge proximity. For instance, the cosine similarity index can be calculated between two vectors of the occurrences of a pair of classes with all other classes in patents (Breschi, Lissoni & Malerba 2003; Ejerimo 2005; Kogler, Rigby & Tucker 2013). The normalized co-classification index measures the deviation of the actual observed co-occurrences of class pairs in patents from random expectations (Teece *et al.* 1994; Dibiaggio, Nasiriyar & Nesta 2014; Yan & Luo 2017) have reviewed and compared various knowledge proximity measures used in patent mapping. Note that, this strand of research on measuring the knowledge proximity between different patent technology classes was not previously engaged in the engineering design literature.

## 2.4. Summary

In brief, patents have been used as stimuli to foster ideation; however, while there exists evidence that the use of such stimuli is beneficial, the observations on the effect of analogical distance on for example, the attributes of design outcomes have not been consistent. Moreover, several metrics have been used to measure the proximity between stimuli and design problems and distinguish near- and far-field stimuli (Fu *et al.* 2013a). However, most of the prior studies are based on the textual analysis of small sets of patents selected from the patent database. No efforts, to our knowledge, have been pursued at identifying near and far fields to a design problem in the total technology space, and at searching for patent stimuli in the total patent database. The network of all technology classes may serve as a macro and consistent framework to define home, near and far fields to a target design problem or more open-ended design interest.

In the present study, we make use of the patent technology class network to classify the patents in the total patent database into home, near and far fields to a design problem. On this basis, we seek to answer the following questions:

- (1) Where are the sources of useful patent stimuli in the technology class network: home, near or far fields?
- (2) What are the implications of using patent stimuli from these different fields on the outcomes of ideation?

## 3. Method and data

This study analyzes the data, including the patent stimuli and generated concepts, from an ideation exercise. In this section, we will introduce the exercise and the methods used to analyze the patent stimuli and concepts.

### 3.1. Ideation exercise and data

Data from an ideation exercise of 30.007 Engineering Design and Project Engineering, a course offered at the Engineering Product Development (EPD) Pillar (<https://epd.sutd.edu.sg/>) of Singapore University of Technology & Design (SUTD) (<https://www.sutd.edu.sg/>) is used for this research. This course is mandatory for the second-year undergraduate students in the EPD Pillar and provides a holistic understanding and competency in engineering design. All the students participating in this ideation exercise had undertaken several design courses and structured design projects prior to this course. The ideation exercise was an early part of a design project, which ran throughout the course. The objective in this project was to conceive, design and develop an innovative spherical rolling robot (SRR) concept of self-defined system requirements, and fabricate a functional prototype. This objective was deliberately kept open to provide students the flexibility and room for creativity and innovation.

Before ideation, all the student designers were provided with Sphero™, a SRR toy manufactured by the company Sphero Inc. (<http://www.sphero.com/sphero>), to play, analyze and understand the structure and functioning of a SRR. Sphero is propelled by a self-contained cart and installed with an on-board micro-controller unit. Users may manipulate its motion remotely via a smartphone or tablet. Sphero represents a generic design of SRRs and is also a successful commercial product in the market. The designers were also offered access to 15 prototypes of SRRs



developed earlier at SUTD. The purpose of such sharing before ideation is to allow the students to rapidly learn and build up the basic design knowledge of SRRs.

The research team prepared two sets of patents for student designers to read and get inspired. The most cited US patent from each of the 121 3-digit technology classes defined in the IPC system was provided. The number of forward citations received by a patent is highly correlated to its realized value or importance (Trajtenberg 1990; Hall, Jaffe & Trajtenberg 2000). These 121 patents constituted the first set (Most Cited set). In addition, a randomly identified patent from each of the 121 3-digit IPC technology classes was also provided. These 121 random patents were identified using a random number generator and constituted the second set (Random set). The participants were provided with the title, abstract and images of the patents. If the participants found these contents relevant and inspirational for their problem, they could read the technical descriptions of the patents. Note that it was not mandatory for participants to use the provided patents as stimuli. In addition to the 242 given patents, all the participants were allowed to search and use other patents and resources (such as internet and books) for inspiration. The two sets of patents provide a basic coverage of patents from all the 121 technology classes in the total technology space, and complement the intuitive unguided search of the participants by bringing all the technology classes to the attention of the searchers.

The participants were instructed to generate functional and novel concepts, but no limit was fixed on the number of concepts they must generate. The participants were given a week to generate concepts and asked to sketch or render concepts with annotations and briefly explain how they work. At the end of the exercise, they needed to submit a report for each concept generated. Figure 1 shows an example of the submitted reports. Specifically, the participant must report which patents were used as stimuli and their justification, other resources accessed, and how stimuli were transformed into the new SRR concept (see Figure 1a), in addition to a textual description and a sketch of the generated SRR concept (Figure 1b). In the end of the concept generation exercise, 138 SRR concepts were generated using 231 patent stimuli. Among these patent stimuli, 39 patents were from the Most Cited set, 33 from the Random set, and the rest were searched and identified by the student designers on their own.

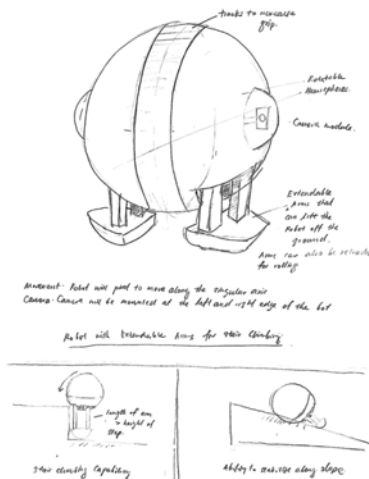
In addition, a consent form seeking the approval of participation was also collected from all the participants. A pre- and a post-ideation survey were conducted to collect information relating to age, gender, academic background, nationality, and other demographic data of the participants, to understand their experience of using patents *a priori* and *posteriori* to this exercise and the effects of their use.

## 3.2. Evaluation of ideation outcome

From the concept generation reports from individual participants, the stimuli used to generate each concept were identified, and novelty and quality of generated concepts were assessed based on their sketches, renderings and annotations. In the literature, researchers have proposed various metrics to assess the performance of ideation, in terms of the attributes of ideation outcomes, such as quantity, quality, novelty, variety, fluency, usefulness, feasibility, and similarity (McAdams & Wood 2002; Shah *et al.* 2003; Sarkar & Chakrabarti 2011). In this research, novelty and quality were used as metrics to assess performance of ideation.

Student Name: [REDACTED]	Student ID: [REDACTED]	Concept Number: 201701
<b>Design stimuli</b>		
1. Did you use patents for generating this concept? Yes If YES, (i) State patent number(s): US726422 B2		
Which above did you read technical description in addition to abstracts and figures? US726422 B2		
2. Why did you choose these patents (stated in 1-ii) to read technical description in detail? US726422 B2 has an interesting mechanism that allows it to extend a certain tool out of the spherical body. I read the technical description to understand the mechanisms behind this extension capability, which may provide the solution to stair climbing problem for spherical robots.		
3. Any other resources that provided inspirations for generating this concept: <a href="https://www.youtube.com/watch?v=UFPd7A210&amp;list=PLuUOP5FCooNDMqUEFF6mBR_YGWMymD00">https://www.youtube.com/watch?v=UFPd7A210&amp;list=PLuUOP5FCooNDMqUEFF6mBR_YGWMymD00</a> (A video demonstration stair climbing robots)		
4. Detail on how did you transform the above-mentioned design stimuli into a new concept? Substitution: Instead of extending a mechanical arm as proposed in past patents, we propose extending a segment of the spherical casing that is capable of lifting the robot off the ground.		
5. Description of the new concept (sketches and annotations)		
Stair climbing is an advantageous capability for surveillance bots as it provides lateral maneuverability in built up environments. Most stair climbing robots have elongated bodies and climbing mechanisms positioned at the front and rear most ends of their bodies. Otherwise, they are equipped with long caterpillar tracks to roll up the stairs, bypassing the need to even "climb". Spherical robots however do not have the requisite length.		
The idea of "bypassing the need to climb" is applicable to the spherical bot. Specifically, we propose a mechanism to extend a segment of the spherical case to lift the robot off the ground, to a height that is higher than the step of the stair. This movement is looped until the robot has completely climbed the stairs.		

(a) Reported information in the concept generation report



(b) Concept sketch with annotations in the concept generation report

**Figure 1.** A concept generation report from a student designer. (a) Reported information in the concept generation report. (b) Concept sketch with annotations in the concept generation report.

Novelty of a design outcome is a measure of unusualness or unexpectedness of the outcome in comparison to other outcomes that perform the same overall function. An expert in robotics and SRRs rated the novelty of the concepts on a 4-point scale (0–3), corresponding to no, low, medium and high novelty. This expert has extensive knowledge of prior arts in SRRs, based on which novelty of the generated concepts was evaluated. For example, the concept shown in Figure 1(b) can climb stairs by extending its arms, which had been seldom seen in prior designs. Therefore, this concept obtained a novelty score of 3.



Quality of an outcome is the degree of the fulfillment of requirements for which the outcome is developed. In the assessment of quality, three abstraction levels, namely functional, working principle, and structural levels, were considered. Quality of a concept was assessed using the formula:

$$Q = 0.5 \times f + 0.3 \times w + 0.2 \times s \quad (1)$$

where  $Q$  is the overall quality of a concept,  $f$  is a measure of the degree of fulfillment of the identified requirements by the functions in the concept,  $w$  is the degree of fulfillment of the identified functions by the working principles in the concept, and  $s$  is the degree of fulfillment of the working principles by the components and their relations in the concept. A weighting scale of 0.5, 0.3 and 0.2<sup>1</sup> was used corresponding to the function, working principle, and structural levels, respectively, because higher abstraction levels are the basis for building the lower abstraction levels.  $f$ ,  $w$  and  $s$  were rated by one of the authors using a 3-point scale (0–2)<sup>2</sup>, corresponding to no, partial and complete fulfillment. Therefore, the overall quality of a concept also varied between 0 and 2. For example, the sketch of the design for climbing stairs (see Figure 1b) describes a full set of functions required to fulfill the stated objectives (rolling on ground and climbing stairs) including rotate two hemispheres for propelling and steering, increase grip, monitor environment with camera, and extend arms for lifting the robot. So, it received 2 points for fulfilling requirements. This concept lacks mechanism details of how to propel the robot and extend the arms, and so, it received 1 point for partially fulfilling the functions identified earlier. Due to the absence of working principles, the design also lacks information of structural features required to fulfill the missing working principles, such as the transmission system for propelling, and so, it received 1 point for the fulfillment of working principles. When these individual weightings were substituted in (1), an overall quality score of 1.5 was obtained for the concept.

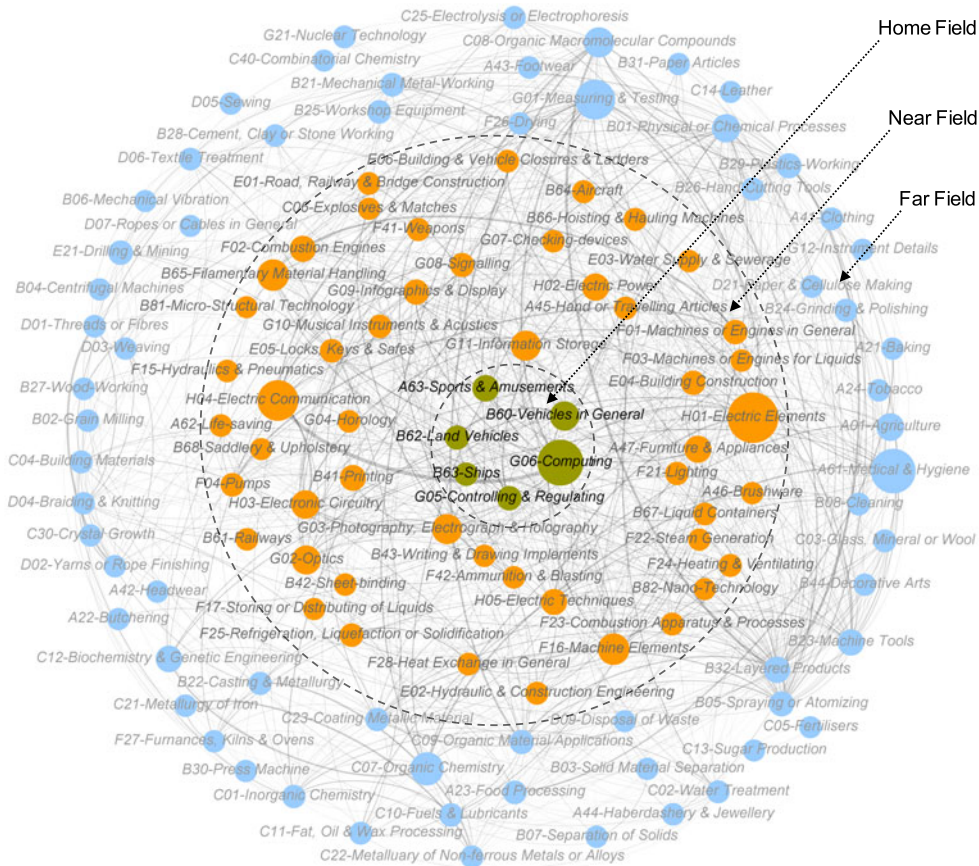
An inter-rater reliability test was conducted using three raters for 20 concepts. After two iterative rounds of analyzing, settling, reconciling differences and reaching Cohen's Kappa ratio of 0.86, the quality of the remaining concepts was rated based on the learning gained from the earlier iterations.

### 3.3. Locating patent stimuli in home, near and far fields within the patent technology network

To analyze the influence of stimulation distance of patent stimuli on ideation outcomes, we located the patents used as stimuli in the ideation exercise within the network of all technology classes. In the network, the stimulation distance of a patent to a design problem can be measured according to the knowledge proximity between the technology classes containing the patent and the technology classes that correspond to the designers' knowledge related to the design problem, i.e., the home field. To align with the theoretical lens of near and far analogies in the

<sup>1</sup> We carried out sensitivity analyses with altered weights for  $f$ ,  $w$  and  $s$  to investigate the robustness of the findings. For each of the tests, we held one of the three weights fixed, increased or decreased the second weight by 10%, and then adjusted the third one accordingly to ensure the weights sum up to 1. The findings regarding concept quality hold true in the tests.

<sup>2</sup> We alternatively experimented 5-point scale to assess  $f$ ,  $w$  and  $s$ , and found it was difficult for three raters to achieve a Cohen's Kappa ratio higher than 0.8 in the inter-rater reliability tests. By contrast, a 3-point scale evaluation enabled satisfactory inter-rater reliability.



**Figure 2.** Home, near and far fields of the SRR design in the technology class network. The size of each node is proportional to the number of patents in each technology class, and the thickness of a link is proportional to the knowledge proximity between the corresponding pair of technology classes. Nodes in the home field are highlighted in green, near field in orange, and far field in blue.

literature, we further located patents in home, near and far fields, which are groups of technology classes based on the latent community structure of the technology network.

**3.3.1. Construct a technology network**

First, we used the entire USPTO database from 1976 to 2016 to empirically create a patent technology network that approximates the total space of all known technologies to date (Alstott *et al.* 2017a; Yan & Luo 2017). In the network, 121 3-digit IPC classes, such as node F02 that represents a class of patents for combustion engine and node G06 for computing, are used to operationalize the nodes. Each network node representing a technology class can be viewed as a category of patents. The nodes are connected to each other according to the knowledge proximity between them, as shown in Figure 2.

If two technology classes have low knowledge proximity, i.e., design processes in two technology categories require relatively distinct design knowledge,

designers specializing in one technology category may find it difficult to understand or design using knowledge and technologies from the other (Luo 2015). On the contrary, if the design process in two technology categories requires similar knowledge pieces, designers in one category can easily understand and leverage design knowledge from the other. Prior patent data analysis has also statistically shown that inventors are more likely to succeed in filing patents in proximate categories in the technology space (Alstott *et al.* 2017a,b). Therefore, the information of knowledge proximity among technology classes will enable one to locate patents with different distances to a design problem in the technology space.

We utilized the reference-based cosine similarity index to calculate the knowledge proximity. Specifically, the distribution of references from patents in a technology class to unique patents is represented as a vector to characterize the design knowledge base of the technology class. The references of a patented technology are the proxy of the design knowledge used in the design of the technology. Then the knowledge proximity between a pair of technology classes is calculated as the cosine of the angle between their corresponding vectors (Yan & Luo 2017), as follows:

$$\text{Proximity} = \cosin(i, j) = \frac{\sum_k C_{ik} C_{jk}}{\sqrt{\sum_k C_{ik}^2} \sqrt{\sum_k C_{jk}^2}} \quad (2)$$

where  $C_{ik}$  or  $C_{jk}$  denotes the number of citations referred from patents in technology class  $i$  or  $j$  to the specific patent  $k$ ;  $k$  belongs to all the patents cited by patents in either technology class  $i$  or  $j$ . The cosine similarity index value is in the range  $[0, 1]$  and indicates the proximity of knowledge pieces required in designing technologies in both classes. In this study, the references of more than 6 million utility patents in the USPTO database were analyzed to calculate the cosine between each pair of the 121 IPC classes for the best possible empirical approximation of knowledge proximity between them.

This knowledge proximity measure is theoretically motivated by the design-by-analogy literature that has primarily focused on ‘similarity’, e.g., functional, structural and surface similarity, to define and measure analogical stimulation distance (Gentner & Markman 1997; Ward 1998; Christensen & Schunn 2007; Fu *et al.* 2013a; Murphy *et al.* 2014; Fu *et al.* 2015). In contrast to these prior studies addressing the similarity between specific designs or individual patent documents, our measure is formulated for the similarity between technology classes, i.e., categories of patents. At this level, a few studies (Kay *et al.* 2014; Leydesdorff *et al.* 2014) have used the cosine similarity of the vectors of patent references made from a pair of classes to other classes (i.e., class-to-class reference vectors). Our measure extends to class-to-patent vectors, concerning references to specific patents instead of aggregated classes, for a higher granularity. In addition, according to a recent study that compared 12 alternative knowledge proximity measures, our measure appears as one of the most correlated with and representative of other alternative knowledge proximity measures in the literature (Yan & Luo 2017).

### 3.3.2. Detect communities in the technology network

In the technology network, some groups of nodes are more cohesively connected internally and have a higher density of links within than between them. Such

dense groups of nodes are often called communities or clusters. In the network analysis and graph theory literature, various community detection algorithms have been developed to discover and analyze the latent community structures in networks (Clauset, Newman & Moore 2004; Newman 2006; Blondel *et al.* 2008; Chen *et al.* 2010; Browet, Absil & Van Dooren 2013; Wu *et al.* 2015). Specifically, in terms of the technology network in Figure 2, communities are cohesive groups of technologies (i.e., patent classes) with high knowledge proximity between them. Technologies in the same communities possess more common knowledge than technologies in different communities, and thus it is more likely to draw analogies across technology classes in the same community.

In this paper, we employed a hierarchical agglomeration algorithm proposed by Clauset *et al.* (2004) to detect the technology network's latent community structure. This algorithm was chosen because it is more efficient and faster than competing algorithms and returns a uniquely determined community partition rather than heuristic results. We assessed the community groupings of technology classes resulting from the algorithm and deemed them reasonable based on our engineering knowledge. Consequently, the 121 technology classes of the technology network were clustered into 5 communities. We also compared the community detection result with those from Louvain's greedy optimization method (Blondel *et al.* 2008) and found the results are consistent.

### 3.3.3. Locate 'home field'

To locate the 'home field' of SRRs in the total technology space, we first need to retrieve a set of US patents that can comprehensively represent the participant designers' knowledge base that is related to SRRs. As introduced in Section 3.1, prior to the ideation exercise, Sphero provided the students with the basic knowledge and understanding of SRRs, so we utilized the patents related to Sphero to define the students' SRR-related knowledge base. On this basis, we searched the patents of Sphero Inc. and obtained 16 patents as of 31st August 2016.

Then we used the classification information of the retrieved patents to identify the home field in the technology network. To do this, the following two steps are carried out: (1) identify the technology classes that contain the retrieved patents and sort them in descending order of the number of retrieved patents they contain; (2) successively identify the minimum set of classes required to cover all the retrieved patents. Such a procedure is unambiguous and reproducible. Specifically, the 16 patents are classified in 8 technology classes. Among them, the 6 technology classes – 'G05 Controlling & Regulating', 'A63 Sports & Amusements', 'B62 Land Vehicles', 'G06 Computing', 'B60 Vehicles in General' and 'B63 Ships' – constitute the smallest set of top classes that cover all the 16 patents.

In addition, the technologies used in the 15 exemplar SRRs presented to the students are also well covered by the 6 technology classes. Therefore, we considered these 6 technology classes as the 'home field' of the SRR design in the total technology space, which are located at the center and highlighted in green in Figure 2. We also tested the robustness of the choice of technology classes to represent the home field. First, we found each of the top 3 technology classes contains more than a half of the total set of 16 patents and has a much greater coverage than the other technology classes. We tested only using the top 3 classes to define the home field. The statistical results presented in Section 4 vary slightly and do not affect the general conclusions. In addition, we also compared the patent

set used in this paper with the patent set that resulted from an exhaustive patent search for SRRs and contained 153 SRR patents (Song & Luo 2017). The results show that only 1 of the top 6 technology classes differs for the two patent sets. These results show the robustness of the definition of the home field.

### 3.3.4. Identify 'near field' and 'far field'

Based on the network partition results from the algorithm of Clauset *et al.* (2004), the home-field classes G05, A63, B62, G06, B60 and B63 belong to 2 network communities. Then the technology classes other than these 6 in the same 2 communities were designated as the 'near field' of the SRR design, which are located at the inner ring and highlighted in orange in Figure 2. The near field surrounds the home field. The technology classes outside these 2 network communities were considered the 'far field' of the SRR design, which are located at the outer layer and highlighted in light blue in Figure 2. Thus, the technology class network, which represents the total technology space, was divided into three mega fields: home, near and far fields. On this basis, the patent stimuli used were assigned to one or multiple of the three fields in the technology space, according to their classification information. Note that a patent may belong to multiple fields if it is assigned in multiple technology classes.

In brief, the method of locating patents in home, near or far field involves three main procedures: (1) construct a network map of all technology classes in the patent database to represent the total technology space, (2) detect network communities, and (3) determine the home, near and far fields in the network. In turn, these three procedures respectively require: (1) a measure of the knowledge proximity between patent classes, (2) a community detection algorithm, and (3) a patent set representing the home of a design problem. In this sub-section, we have introduced our choices for each of the three elements. On this basis, we located the patent stimuli used in the ideation exercise in the home, near and far fields of the SRR design in the technology network, for further analysis.

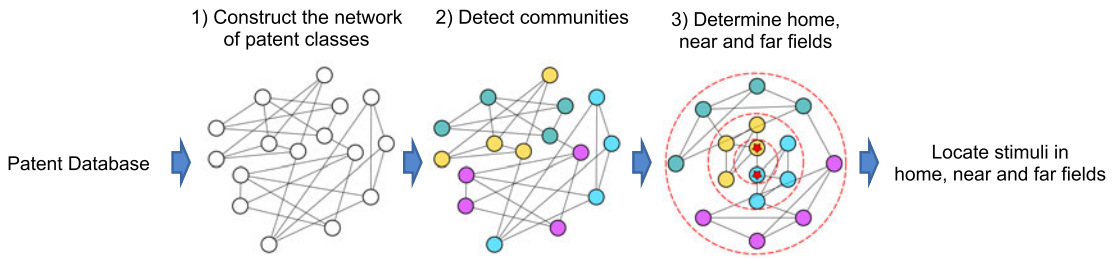
## 4. Findings

In this section, we report the frequencies and likelihood of the participants finding patent stimuli from home, near and far fields, and the novelty and quality of the concepts generated with patent stimuli from home, near and far fields.

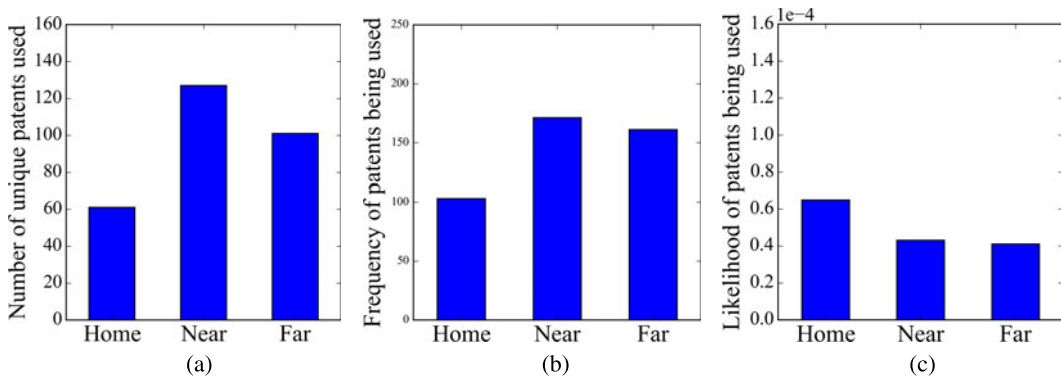
### 4.1. Where designers find patent stimuli in the technology space

Figure 4(a) shows the number of reported unique patent stimuli from home, near and far fields. The participant designers can use a patent as a stimulus for generating multiple concepts. Figure 4(b) shows the frequency of patents being used with multiple counting, i.e., it counts the use of a patent as stimulus for multiple concepts. Both the figures (Figure 4a,b) show a similar pattern: most patent stimuli used for concept generation are from the near field of the SRR design. We also calculated the likelihood for patent stimuli being identified from the various fields, as the number of patents used in a field to generate concepts divided by the total number of patents granted in the corresponding field from 1976 to 2016. As seen in Figure 4(c), patents in the home field are more likely to be used to generate concepts than those in the near and far fields, for which the likelihoods are almost the same.





**Figure 3.** The three procedures of our method to determine home, near and far fields.

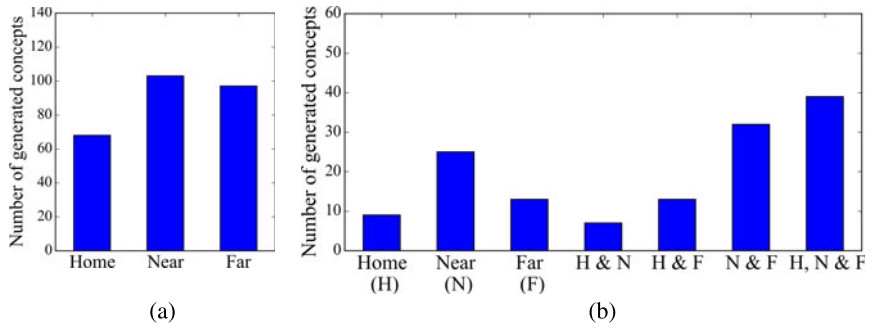


**Figure 4.** Patent stimuli in each field of the technology space: (a) number of unique patents used; (b) frequency of patents being used; (c) likelihood of patents being used.

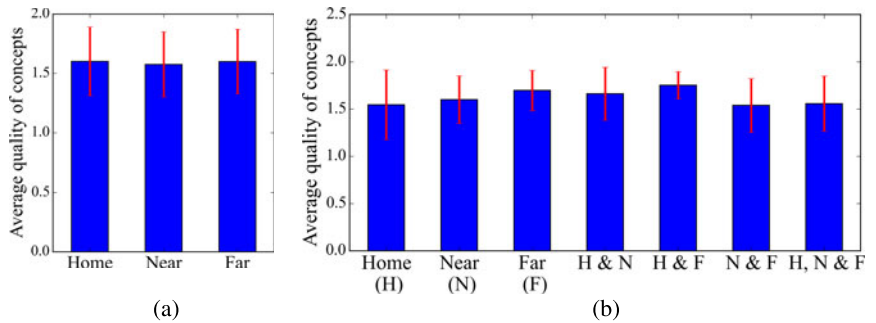
A concept can be stimulated by either a single patent or multiple patents, whose classes may fall into one or more of the home, near and far fields. Figure 5 shows the number of the concepts generated using patents from the individual fields and their combinations. In Figure 5(a), a concept stimulated by patents from multiple fields is counted multiple times, once for each field to which a patent stimulus belongs. Patents from the near and far fields help generate more concepts than patents from the home field. In order to present the results more unambiguously, we categorize a concept into only one of the individual fields or the combinations of multiple fields according to the sources of patent stimuli used by the concept. Under this setting, each concept is counted only once in one category. As shown in Figure 5(b), the highest number of the concepts is generated using patents from a combination of home, near and far fields (H, N & F), followed by the combination of near and far fields (N & F). The influence of patents from the near field either individually or in combination with other fields is prominent. As a single source, patents from the near field help generate most concepts.

#### 4.2. Implications of home-, near- and far-field stimuli on ideation outcomes

The average quality of the concepts generated using patents from the individual fields and their combinations is shown in Figure 6(a) and Figure 6(b), respectively. No significant difference in average quality across the individual fields is observed



**Figure 5.** Numbers of concepts generated using patents from individual fields and combinations of fields: (a) with patents from each field (one concept may be counted more than once); (b) with patents from combinations of the three fields (each concept is counted only once).



**Figure 6.** Average quality of generated concepts: (a) with patents from each field; (b) with patents from combinations of the three fields.

in Figure 6(a), suggesting that the three fields contribute almost identically to the quality of the concepts. When considering the combinations of the fields, we can see in Figure 6(b) that concepts stimulated with patents from the far field individually or the combination of home and far fields (H & F) have a significantly higher average quality than the rest. Such differences in concept quality are significant at 5% level in most cases based on pairwise 2-tailed t-tests, as shown in Table 1.

Figures 7(a) and figure 7(b) shows the distributions of the concepts generated using patents from the individual fields and their combinations by quality, respectively. In the figures, the low, medium and high quality categories correspond to the ranges of quality scores  $Q \leq 1.2$ ,  $1.2 < Q < 1.8$  and  $Q \geq 1.8$ , respectively, according to the multimodal frequency distribution of the generated concepts by quality. As observed in Figure 7(a), patents from the home or far field stimulate a higher percentage of high quality concepts than those from the near field. In Figure 7(b), a higher percentage of the concepts stimulated by patents from: (a) the far field and (b) the combination of the home and far fields (H & F) have high quality. Interestingly, no low-quality concepts are generated using patents from the far field individually, the combination of the home and far fields (H & F) and the combination of the home and near fields (H & N).



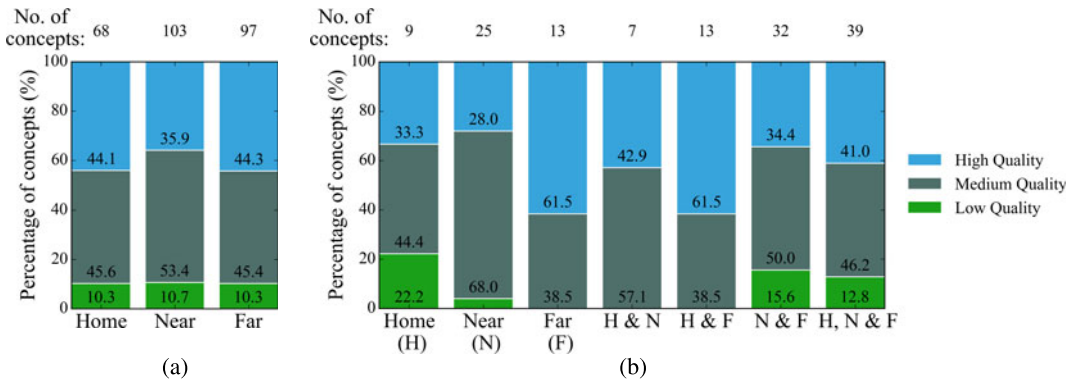


Figure 7. Distributions of concepts by quality: (a) with patents from each field; (b) with patents from combinations of the three fields.

Table 1. *t* statistics with *p*-values in parentheses for the pairwise comparison of the quality of concept sets as indicated by the row and column labels. Underlines denote significance at 5% level

	Far	Home and Far
Home	1.7303 (0.0909)	<u>3.5036</u> (0.0008)
Near	1.6473 (0.1039)	<u>3.8277</u> (0.1816)
Home and Near	0.4766 (0.6359)	1.8383 (0.0699)
Near and Far	<u>2.6887</u> (0.0080)	<u>5.3607</u> (<0.00001)
Home, Near and Far	<u>2.3611</u> (0.0189)	<u>4.8737</u> (<0.00001)

The average novelty of the concepts generated using patents from the individual fields and their combinations is shown in Figure 8(a) and Figure 8(b), respectively. The differences in average novelty of the concepts generated using patents from the individual fields are not statistically significant (see Figure 8a). Concepts stimulated by patents from the near field individually and all combinations containing the near field (N; H & N; N & F; H, N & F) have higher average novelty than concepts stimulated by patents from other individual fields or combinations (see Figure 8b). Specifically, the differences in average novelty are statistically significant at 5% level between concepts stimulated by patents from the combinations containing the near field (H & N; N & F; H, N & F) and those stimulated by patents from the home or far field individually or their combination (H; F; H & F), as shown in Table 2.

Figures 9(a) and figure 9(b) shows the distributions of the concepts generated using patents from the individual fields and their combinations by novelty, respectively. As mentioned earlier, the high, medium, low and no novelty corresponds to novelty score 3, 2, 1 and 0, respectively. As observed in Figure 9(a), patents from the near field contribute a slightly higher percentage of high novelty concepts than patents from the home and far fields. It is clear from Figure 9(b) that the concepts generated with patents from the near field individually and all combinations containing the near field (N; H & N; N & F; H, N & F) present a

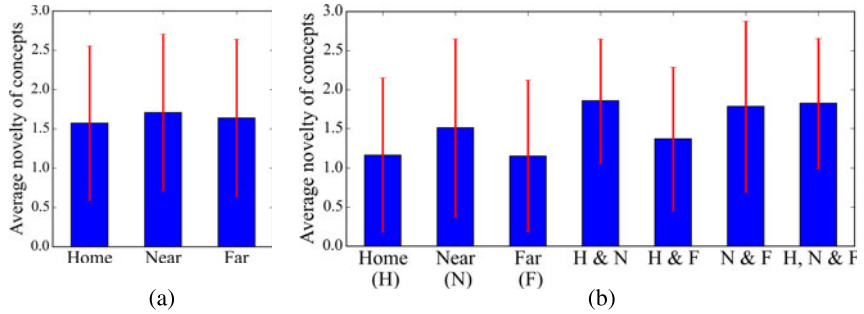


Figure 8. Average novelty of generated concepts: (a) with patents from each field; (b) with patents from combinations of the three fields.

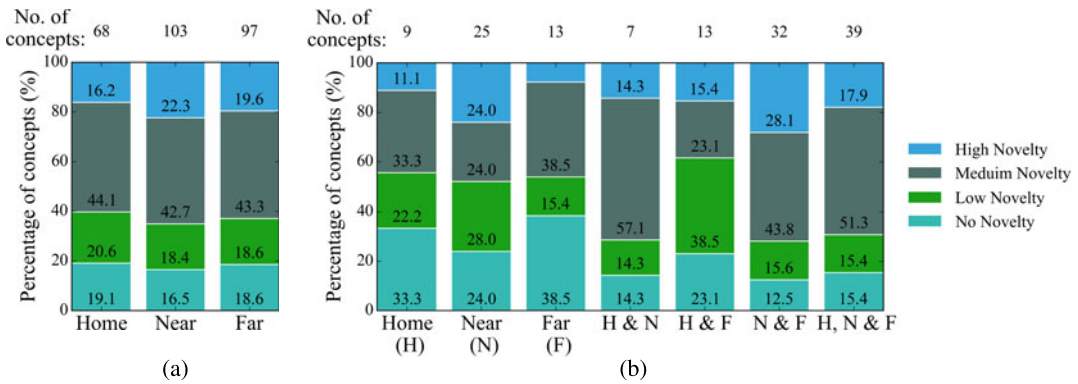


Figure 9. Distributions of concepts by novelty: (a) with patents from each field; (b) with patents from combinations of the three fields.

Table 2. *t* statistics with *p*-values in parentheses for the pairwise comparison of the novelty of concept sets as indicated by the row and column labels. Underlines denote significance at 5% level

	Near	Home and Near	Near and Far	Home, Near and Far
Home	1.1280 (0.2636)	<u>2.4251</u> (0.0203)	<u>2.2740</u> (0.0244)	<u>3.2022</u> (0.0015)
Far	1.3491 (0.1816)	<u>2.6817</u> (0.0102)	<u>2.7359</u> (0.0070)	<u>3.8571</u> (0.0001)
Home and Far	0.7047 (0.4826)	<u>2.1577</u> (0.0341)	<u>2.5120</u> (0.0129)	<u>3.6733</u> (0.0003)

higher percentage of high and medium novelty. These results are consistent with the findings in Figure 8(b). The results suggest that patents from the near field play a major role in stimulating concepts of higher novelty compared to patents from the home or far field.

## 4.3. Summary of findings

The following are observed in the data from the open concept generation exercise and using the technology space network-based definition of home, near and far fields of patent stimuli:

- (1) Patents from the home field are most likely to be used as stimuli for concept generation.
- (2) The near field contributes most patent stimuli for concept generation, and the most number of concepts are generated using patents from the near field.
- (3) The concepts generated with patent stimuli from the far field individually and the combination of the home and far fields have a higher average quality than concepts generated with patents from other individual fields or other combinations of the three fields.
- (4) The concepts generated with patent stimuli from the near field individually and the combinations of the near field with other fields have higher average novelty than concepts generated with patent stimuli from the home field, far field or their combination.

## 5. Discussion

### 5.1. Significance of findings

In this paper, we have presented a network-based approach to divide the technology space into three fields – home field, near field and far field – with varied knowledge proximity to a design problem based on the community detection within the network of all 3-digit IPC technology classes in the patent databases. The ‘home field’ entails technology classes that contain patents that are directly relevant to a design problem, the ‘near field’ comprises technology classes that are in the same network communities as those in the ‘home field,’ and the remaining technology classes from other communities in the technology network constitute the ‘far field.’ Each technology class can be further viewed as a category of patents. The data-driven definition of the home, near and far fields is motivated by the literature on design by analogy using patents as stimuli, which has used such terms as near or far analogies or stimuli, to characterize discrete stimuli or patents. Our approach provides a systemic and macro framework, which one can consistently use to position a patent in either home, near or far fields of design stimulation in the total technology space.

Based on the definitions of the home, near and far fields in the total technology space, our experiment shows that patents from the home field are more likely to be used than those from the near or far field. Since the home field is the area nearest to the defined design problem or opportunity, the result produces evidence in support of the argument that it is easier to analogize with stimuli from near domains than far domains to some extent. In the studied case, patents in the home field record technologies and processes having much more common knowledge basis with the SRR-related design knowledge of the student designers, and so the patents are much easier to be identified, understood and assimilated by them.

Results in this paper further suggest that in the given context, concepts generated using patents from the far field or the combination of the home and far fields have a higher average quality than the concepts stimulated by patents from other individual fields or combinations. This seems to contradict the findings of

Fu *et al.* (2013b) that solutions generated using stimuli from near sources had higher quality than those generated using far sources, but support the findings of Chan *et al.* (2011) that stimuli from far sources were more beneficial for developing solutions of higher quality. However, these studies and ours might not be comparable, primarily because the characterization of near and far is not the same. Fu *et al.* and Chan *et al.* used functional similarity to characterize the distance as near and far between individual patents and a specific design problem, while in this study knowledge base similarity between patent technology classes is used. Also, it must be noted that the context setting for this study is quite different from theirs. That is, the objective of the design problem is open and the participants are free to select patent stimuli from the given sets or search their own independently. In this case, it is understandable that where the patent stimuli are from has limited influence on feature creation. Although patents from different fields are used with different likelihoods, once they are identified and chosen as stimuli for concept generation, the designers would make effort to understand and make sense of the information provided in the patents and transform it into features in their own concepts.

The near field provides the most patent stimuli, which further stimulate the largest portion of the concepts. Concepts generated using patents from the near field have higher novelty on average than those generated without patents from the near field. On one hand, the near field is relatively nearer to the design problem than the far field, and contains design knowledge that is relatively easy to understand and make sense of. On the other hand, it is relatively more distant to the design problem than the home field, and probably provides stimuli with plenty of additional features for conceiving innovative attributes. Potentially, these attributes may contribute to novelty. In brief, it can be argued that patents from the near field are more beneficial for identifying stimuli and creating novel attributes in concepts compared to those from the other fields.

This research uses data that comprises several variables. 138 concepts were generated with 231 patents by student designers, who ideated individually in uncontrolled conditions. The student designers in this study, unlike other laboratory-based controlled ideation experiments, required more domain knowledge to accomplish the task. From among an alternative set of the concepts generated by individuals, one concept was chosen and if necessary, modified, and then prototyped and demonstrated by each team. So, the quality of the prototype depended on the concept set generated earlier. Therefore, the ideation was a fun exercise with lots of project stakes attached to it. The grades of student designers depended on their performance at every phase of the SRR development process. Some of the projects were further pursued toward entrepreneurial and co-curricular activities. So, the participants had adequate vested incentives to pursue this ideation exercise seriously. Therefore, the results of this study must be viewed taking into consideration the wide span of variables and the seriousness with which this exercise was pursued.

In addition to our experiment findings, this paper may have made a contribution to ideation methodology development, specifically regarding patent stimuli search. Although prior studies have suggested the efficacy of using patents as stimuli for concept generation, browsing through the huge patent database within a short period to identify relevant stimuli may be cumbersome. To address the problem, an efficient search-and-retrieve interface is required, through

which millions of patents can be searched through using defined keywords and relevant patents can be retrieved and ranked in the order of their appropriateness to the keywords. Fu *et al.* (2013a) and Murphy *et al.* (2014) have developed computational design tools to search and identify functionally relevant analogies from the US patent database. Results in this paper offer fundamental insights on designers' natural preferences for stimuli and the influence of stimulation distance on ideation outcomes. With these findings, the introduced method in Section 3.3, which is based on a technology class network and community detection to define the home field, near field and far field, might be a first step toward a data-driven computational tool for better-guided and more-informed search of patents as design stimuli. Such a tool is expected to allow designers to locate the home field of a design problem and be informed of the fields of their search for patent stimuli, and systematically guide them through the search for patent stimuli in either home field, near field, far field or their combinations.

In fact, the method we used in this experimental research is structured and repeatable. As introduced in Section 3.3 and depicted in Figure 3, the method generally involves three main procedures: (1) construct a network map of technology classes to represent the total technology space; (2) detect network communities; and, (3) determine the home, near and far fields in the network. In practice, once the home, near and far fields have been identified in the network, a designer can search, locate and use patent stimuli within the home, near and far fields, with the guidance of the understanding of the potential effectiveness of finding useful patent stimuli from different fields and the corresponding performance implications, as suggested from our experimental findings.

Meanwhile, the procedures require three key elements in practical implementation: (a) a knowledge proximity measure for constructing the network of technology classes, (b) a computer algorithm for partitioning the technology network into a few communities, and (c) a patent set for identifying the technology classes to determine the home field. For each of the elements, there are alternative implementation choices. In this paper, we have provided one superior choice for each element for our case study, but do not limit to them. For the first two elements, in this study, alternative knowledge proximity measures and alternative community detection algorithms result in quite similar community partitions of the technology network. In a different case, one can still pick out superior choices of the measures and algorithms by comparing the resulting community partitions with the expectation based on his or her knowledge and the specific context. Moreover, the patent set used to identify the home field can also be determined using an approach according to the context. For example, to find the home field of a designer, one can search the patents granted to the designer; to find the home field of a designer who has no patents, one can search patents using keywords that describe the technical expertise of the designer; one can also combine the search approaches. In brief, our method provides a structured but flexible framework, whose elements can be operationalized and calibrated according to the specific context and situation of its application.

## 5.2. Limitations

The results in this research have the following limitations. First, in this research the measure used to assess the knowledge proximity between 3-digit technology classes may not be directly related to the relevance of a stimulus to a problem.

Future research may explore knowledge proximity at alternative granularity levels, such as proximity between 4-digit IPC technology classes or between classes and patents. Second, the student designers received two sets of 121 patents (the most cited patent and a random patent from each of the 121 technology classes) for the ideation exercise. A patent's description text is often lengthy and written in a tedious and non-obvious manner, so it may be difficult for the student designers to browse through all the 242 patents to assimilate the information, identify relevant stimuli from them and use these for generating concepts all within a week. That is, coverage of all the 121 technology classes in the total technology space may not be guaranteed in practice. Future research may seek approaches for communicating the technical information and design knowledge in patent documents to designers more effectively and efficiently. Also, the ideation was an after-class exercise and thus uncontrolled. For concepts generated using both patents and other resources of inspiration, the influence of other resources was not accounted for locating the stimuli in the home, near and far fields in the technology space. This discounts a significant influence of other resources on the novelty or quality of the generated concepts.

Moreover, it should be noted that our findings result from a technology-driven design process with undergraduate designers as the participants. The findings may not hold true for experienced designers or in other design situations, such as the design of market-driven products. For example, when solving a design problem, experienced designers typically have extensive knowledge of the near field as well as the home field through their own learning and experience, which allows them to build their own 'feeling for near'. In such a situation, the experienced designers would not find patent stimuli from the home and near fields so useful as the student designers do. In addition, for the design of market-driven products, the goal is to identify users' value-based needs and convert them into product features successfully. In such a design process, the distance of stimuli from the design problem is largely determined by users' requirements but not the intrinsic proximity or distance between technologies. Moreover, the findings in this paper are based on the case of SRRs and may not hold true for other products, whereas the introduced method of locating patent stimuli is applicable to other contexts. This suggests a future research opportunity to apply our network-based methodology to more diverse products and contexts and potentially develop a contingency understanding on the influence of stimulation distance on ideation outcomes for different types of products.

## 6. Conclusions and future work

This research contributes to fundamental insights on designers' preferences for patent stimuli and the influence of stimulation distance on the ideation outcomes, as well as a network-based methodology for better-guided search of patents as design stimuli in concept generation practices. The objectives of this research are: (a) to identify where designers identify useful patent stimuli within the technology space: home, near or far fields, and (b) to study the implications of using patent stimuli from these fields on the novelty and quality of the concepts generated. It is observed that: (a) patents from the home field are more likely to be used as stimuli, (b) the near field contributes most patent stimuli, which further stimulates the most number of the concepts, (c) concepts generated with patents from the far field and the combination of the home and far fields have higher average

quality than concepts generated using patents from other individual fields or combinations of the three fields, and (d) the concepts generated with patents from the near field have higher average novelty than concepts generated without patents from the near field. The methodology based on a technology class network and community detection to define the home field, near field and far field might be a first step toward a tool for better-guided search of patents as design stimuli. Further efforts can be made to retrieve most useful patents for ideation at the field or class level.

## Acknowledgments

This research is supported partly by a grant from the SUTD-MIT International Design Centre (IDG31600105), and by Singapore Ministry of Education Tier 2 Academic Research Grant (T2MOE1403).

## References

- Alstott, J., Triulzi, G., Yan, B. & Luo, J. 2017a Mapping technology space by normalizing technology networks. *Scientometrics* **110** (1), 443–479.
- Alstott, J., Triulzi, G., Yan, B. & Luo, J. 2017b Inventor's explorations across technology domains. *Design Science* **3**; doi:[10.1017/dsj.2017.21](https://doi.org/10.1017/dsj.2017.21).
- Altshuller, G. S. & Shapiro, R. B. 1956 'О Психологии изобретательского творчества (On the psychology of inventive creation) (in Russian)', *Вопросы Психологии (The Psychological Issues)*, 6(Вопросы Психологии (The Psychological Issues)), pp. 37–49.
- Blondel, V. D., Guillaume, J.-L., Lambiotte, R. & Lefebvre, E. 2008 Fast unfolding of communities in large networks. *Journal of Statistical Mechanics: Theory and Experiment* **2008** (10), P10008.
- Breschi, S., Lissoni, F. & Malerba, F. 2003 Knowledge-relatedness in firm technological diversification. *Research Policy* **32**, 69–87.
- Browet, A., Absil, P. & Van Dooren, P. 2013 Fast community detection using local neighbourhood search, *Preprint*, pp. 1–10. Available at <http://arxiv.org/abs/1308.6276>.
- Cascini, G. & Russo, D. 2006 Computer-Aided analysis of patents and search for TRIZ contradictions. *International Journal of Product Development* **4** (1–2), 52–67.
- Chakrabarti, A., Sarkar, P., Leelavathamma, B. & Nataraju, B. S. 2005 A functional representation for aiding biomimetic and artificial inspiration of new ideas. *Artificial Intelligence for Engineering Design, Analysis and Manufacturing* **19** (2), 113–132.
- 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), 81004.
- Chan, J. & Schunn, C. 2015 The impact of analogies on creative concept generation: lessons from an in vivo study in engineering design. *Cognitive Science* **39** (1), 126–155.
- Chen, D., Wang, J., Chen, X. & Xu, X. 2010 A search algorithm for clusters in a network or graph. *International Journal of Digital Content Technology and its Applications* **4** (6), 115–122.
- Christensen, B. T. & Schunn, C. D. 2007 The relationship of analogical distance to analogical function and preinventive structure: the case of engineering design. *Memory and Cognition* **35** (1), 29–38.



- Clauset, A., Newman, M. E. J. & Moore, C.** 2004 Finding community structure in very large networks. *Physics* **70** (6), P066111.
- Dibiaggio, L., Nasiriyar, M. & Nesta, L.** 2014 Substitutability and complementarity of technological knowledge and the inventive performance of semiconductor companies. In *Research Policy*, pp. 1–12. Elsevier.
- Ejermo, O.** 2005 Technological diversity and Jacobs' externality hypothesis revisited. *Growth and Change* **36** (2), 167–195.
- Fantoni, G., Apreda, R., Orletta, F. D. & Monge, M.** 2013 Advanced engineering informatics automatic extraction of function – behaviour – state information from patents. *Advanced Engineering Informatics* **27** (3), 317–334.
- French, M. J.** 1985 *Conceptual Design for Engineers*. Springer.
- Fu, K., Cagan, J., Kotovsky, K. & Wood, K.** 2013a Discovering structure in design databases through functional and surface based mapping. *Journal of Mechanical Design* **135** (March 2013), 31006.
- Fu, K., Chan, J., Cagan, J., Kotovsky, K., Schunn, C. & Wood, K.** 2013b The meaning of 'near' and 'far': the impact of structuring design databases and the effect of distance of analogy on design output. *Journal of Mechanical Design* **135** (2), 21007.
- Fu, K., Murphy, J., Yang, M., Otto, K., Jensen, D. & Wood, K.** 2015 Design-by-analogy: experimental evaluation of a functional analogy search methodology for concept generation improvement. *Research in Engineering Design* **26** (1), 77–95.
- 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.
- Goel, A. K.** 1997 Design, Analogy, and Creativity. *IEEE Expert* **12** (3), 0–25.
- Hall, B. H., Jaffe, A. B. & Trajtenberg, M.** 2000 *Market Value and Patent Citations: A First Look*. NBER Working Paper No. 7741 (National Bureau of Economic Research); doi:[10.3386/w7741](https://doi.org/10.3386/w7741).
- Jaccard, P.** 1901 Distribution de la flore alpine dans le bassin des Dranses et dans quelques régions voisines. *Bulletin de la Société vaudoise des sciences naturelles* **37**, 241–272.
- Jaffe, A. B.** 1986 Technological opportunity and spillovers of R & D: evidence from firms' patents, profits and market value. *The American Economic Review* **76** (5), 984–1001.
- Jansson, D. G. & Smith, S. M.** 1991 Design-Fixation. *Design Studies* **12** (1), 3–11.
- Jensen, D., Weaver, J., Wood, K., Linsey, J. & Wood, J.** 2009 Techniques to enhance concept generation and develop creativity. In *Annual Conference, Austin, TX*.
- Kay, L., Newman, N., Youtie, J., Porter, A. L. & Rafols, I.** 2014 Patent overlay mapping: visualizing technological distance. *Journal of the American Society for Information Science and Technology* **65** (12), 2432–2443.
- Kogler, D. F., Rigby, D. L. & Tucker, I.** 2013 Mapping knowledge space and technological relatedness in US cities. *European Planning Studies* **21** (9), 1374–1391.
- Leydesdorff, L., Kushnir, D. & Rafols, I.** 2014 Interactive overlay maps for US patent (USPTO) data based on international patent classifications (IPC). *Scientometrics* **98** (3), 1583–1599.
- Linsey, J. S.** 2007 Design-by-analogy and representation in innovative engineering concept generation.
- Linsey, J. S., Tseng, I., Fu, K., Cagan, J., Wood, K. L. & Schunn, C.** 2010 A study of design fixation, its mitigation and perception in engineering design faculty. *Journal of Mechanical Design* **132** (4), 41003.

- Luo, J.** 2015 The united innovation process: integrating science, design, and entrepreneurship as sub-processes. *Design Science* 1; doi:[10.1017/dsj.2015.2](https://doi.org/10.1017/dsj.2015.2).
- Mcadams, D. A. & Wood, K. L.** 2002 A quantitative similarity metric for design-by-analogy. *Journal of Mechanical Design* 124 (2), 173–182.
- Mukherjea, S., Bamba, 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., Fu, K., Otto, K., Yang, M., Jensen, D. & Wood, K.** 2014 Function based design-by-analogy: a functional vector approach to analogical search. *Journal of Mechanical Design* 136 (10), 101102.
- Newman, M. E.** 2006 Modularity and community structure in networks. *Proceedings of National Academy of Sciences, USA* 103 (23), 8577–8582.
- Pahl, G. & Beitz, W.** 2013 *Engineering Design: A Systematic Approach* (ed. K. Wallace). Springer Science & Business Media.
- Qian, L. & Gero, J. S.** 1996 Function-behavior-structure paths and their role in analogy-based design. In *Artificial Intelligence for Engineering Design, Analysis and Manufacturing*, vol. 10(4), pp. 289–312.
- Sarkar, P. & Chakrabarti, A.** 2011 Assessing design creativity. In *Design Studies*, pp. 1–36. Elsevier.
- Shah, J. M., Li, Y. L., Gessmann, T. & Schubert, E. F.** 2003 Experimental analysis and theoretical model for anomalously high ideality factors (n2.0) in AlGaIn/GaN p-n junction diodes. *Journal of Applied Physics* 94 (4), 2627–2630.
- Small, H.** 1973 Co-citation in the scientific literature: a new measure of the relationship between two documents. *Journal of the Association for Information Science and Technology* 24 (4), 265–169.
- Song, B. & Luo, J.** 2017 Mining patent precedents for data-driven design: the case of spherical rolling robots. *Journal of Mechanical Design* 139 (11), 111420.
- Souili, A., Cavallucci, D., Rousselot, F. & Zanni, C.** 2015 Starting from patents to find inputs to the problem graph model of IDM-TRIZ. *Procedia Engineering* 131, 150–161.
- Srinivasan, V., Song, B., Luo, J., Subburaj, K. & Rajesh, M.** 2017a Investigating Effects of Stimuli on Ideation Outcomes. In *DS 87-8 Proceedings of the 21st International Conference on Engineering Design (ICED 17) Vol 8: Human Behaviour in Design, Vancouver, Canada, Aug 21-25, 2017*.
- Srinivasan, V., Song, B., Luo, J., Subburaj, K. & Rajesh, M.** 2017b Understanding effects of analogical distance on performance of. In *ASME 2017 International Design Engineering Technical Conferences and Computers and Information in Engineering Conference*. American Society of Mechanical Engineers.
- Taura, T. & Yukari, N.** 2012 *Concept Generation for Design Creativity: A Systematized Theory and Methodology*. Springer Science & Business Media; doi:[10.1007/978-1-4471-4081-8](https://doi.org/10.1007/978-1-4471-4081-8).
- Tece, D. J., Rumelt, R., Dosi, G. & Winter, S.** 1994 Understanding corporate coherence: theory and evidence. *Journal of Economic Behavior and Organization* 23 (1), 1–30.
- Trajtenberg, M.** 1990 A penny for your quotes: patent citations and the value of innovations. *The RAND Journal of Economics* 21 (1 (Spring, 1990)), 172–187; Published by: Wiley on behalf of RAND Corporation.

- Ward, T. B.** 1998 Analogical distance and purpose in creative thought: mental leaps versus mental hops. In *Advances in Analogy Research: Integration of Theory and Data from the Cognitive, Computational, and Neural Sciences*, pp. 221–230. New Bulgarian University Press, Bulgaria.
- Weisberg, R. W.** 2006 *Creativity: Understanding Innovation in Problem Solving, Science, Invention, and the Arts*. John Wiley & Sons.
- Wilson, J. O., Nelson, B. A., Rosen, D. & Yen, J.** 2010 The effects of biological examples in idea generation. *Design Studies* 31 (2), 169–186.
- Wu, Y., Jin, R., Li, J. & Zhang, X.** 2015 Robust local community detection: on free rider effect and its elimination. *Proceedings of the International Conference on Very Large Data Bases (VLDB)* 8 (7), 798–809.
- Yan, B. & Luo, J.** 2017 Measuring technological distance for patent mapping. *Journal of the Association for Information Science and Technology* 68 (2), 423–437.