A Study on an Assessment Framework for The Novelty Of Ideas Generated By Analogical Thinking

Eunyoung Kim*, Hideyuki Horii*

*Department of Civil Engineering, The University of Tokyo, Tokyo, 113-8656, Japan

Abstract

Although there have been many educational programs focusing on the creation of new ideas, the assessment of novelty is still a controversial issue. As an ideation tool, analogical thinking enables conceptual change, which is seen as a crucial aspect of creativity. In this regard, the use of analogy can be an important instrument to facilitate novel idea generation. Analogies are generated by superficial or structural similarities from the memory. For creating a new idea by analogy, this study regards novelty as the domain-changing influenced by structural consistency with the source ideas. Consequently, we designed an assessment framework based on the latent semantic analysis of the domains and the consistency of the underlying mechanism between the source and the new ideas. Data was collected from the 14 subjects who participated in the workshop for this study. The workshop consists of three tasks: 1) Pre-task: All subjects were asked to read the 25 cases of the collective intelligence services, which is a business model creating value from large and loosely organized groups of people working together electronically e.g. Amazon.com, Google Japanese input; 2) Categorization task: Subjects were asked to categorize each case based on the underlying mechanism of the business through group discussion; 3) Generation task: Subjects were asked to create a new service idea individually using analogical thinking. As a result, 12 ideas were created, 6 of which were assessed as novel according to our assessment framework. The remaining 6 ideas, 4 were assessed as having high superficial similarity in terms of the idea domain, and 2 as having neither superficial nor structural similarity with the source ideas. Although our findings suggest that the proposed assessment framework for novelty evaluation is unable to provide a ‘one-size fits-all method’, it does enable us to overcome some of the limitations of current evaluation methods which depend on subjective judgements for rating.

Keywords: analogical thinking, novelty of ideas, assessment frameworks, innovation workshops, ideation for innovation

* Corresponding author. Tel.: +81-3-5841-6091; fax: +81-3-5841-7496.
E-mail address: horii@civil.t.u-tokyo.ac.jp
1. Introduction

Recently, much attention has been given to increasing innovation in various fields. According to the social science citation index of the Thomson Reuters Web of Science, the number of scholarly articles with innovation in their titles per 10,000 social science articles has increased from 10% shares in 1990s to 20% shares in early 2000s. In addition, world class companies as well as governments have paid profound attention to innovation to tackle the current economic and social challenges. Creating novel solutions to problems has fascinated many researchers in diverse disciplines and has encouraged a greater focus on this important area. As a result, the subject has been approached from a range of perspectives: cognitive, biological, clinical, social, organizational, educational, historical, cultural, and personal. In accordance with the increasing need to facilitate innovation in society, many academic institutions provide workshop programs.

The i.school at the University of Tokyo was established in 2009 with the aim of fostering innovative leaders, and is tasked with the development and provision of human-centered innovation educational programs. ‘Innovation’ here means not only conventional conceptions of technology-driven innovation but also the creation of any kind of new values: adding values through introducing new novel ideas, methods, directions, opportunities, and solutions that meet new requirements, through more effective products, processes, services, and technologies that are readily available to users. The innovation workshop program at i.School encourages participants to generate new ideas based on analogical thinking as leverage into creating new ideas.

Despite all of this attention to the innovativeness, the assessment of novelty is still problematic. It is possible that certain characteristics of the judges may moderate attribute preferences under consideration. Thus, highly creative people (Basadur, Runco, & Vega, 2000) and people who have substantial expertise working in a domain (Weisberg, 1999) may apply different standards in evaluating ideas. Many researchers have discussed the relevance of idea evaluation, only a few studies have expressly examined how idea evaluation operates (Lonergan, Scott & Mumford, 2004). In addition, Roger and Adhikarya (1979) observed what kind of standards people routinely apply in evaluating new ideas and found that low implementation cost, consistency of the idea with extant systems, and rate of return influenced the adoption of new ideas. Prior studies indicate that people actively seek to eliminate original high risk ideas (De Dreu, 2003; Suri & Monro, 2003). Since Schumpeter (1934) has introduced theoretical definition of innovation as new combinations of existing resources, which are economically more viable than the old way of doing things, people have put more emphasis on commercialization role of innovation which differentiating itself from invention. However, it is important to recognize these pragmatic, economic standards are not the only attributes people might consider in appraising new ideas (Blair & Mumford, 2007).

To design a workshop process that enhances innovation and creativity skills, an objective assessment framework for novelty needs to be developed. In response to this, we suggest an assessment framework for the novelty of ideas created through analogical thinking.

2. Literature Review

2.1. Analogical thinking as a tool for idea generation

Numerous previous studies have supported analogical thinking as a key mechanism for novel idea creation (Finke et al., 1992; Gentner et al., 1997; Hesse, 1966; Holyoak & Thagard, 1989; Koestler, 1964; Perkins, 1997). According to several existing studies of new idea creation methods, analogical thinking has the greatest theoretical support as the driver of innovative thought beyond doubt. As a consequence, researchers in major disciplines accept the premise of Goel (1997); Clement (2008); Hofstadter (2008); Holyoak, K. J. (1996) that analogy plays a central role in innovation and creativity.

Clement (1988) examined sources of creative scientific theory formation in the domain of non-formal reasoning. He provided scientific problems to ten professionals in technical fields and videotaped their problem solving process
with think-aloud. For the first task, all participants selected correct answer and eight people generated at least one analogy. Furthermore, with analysis of the think aloud protocol transcript, he concluded that there were at least three types of analogy generation methods: 1) generation via a principle, 2) generation via a transformation, and 3) generation via an association. Among these three methods, the generation via transformation was observed as the most frequently used with 18 out of 31 analogies.

People solve problems better if they have experienced associated or similar problems (Holyoak and Koh, 1987; Novick, 1988; Ross, 1987). In this context, Analogical thinking has been suggested as a basic mechanism inspiring creative tasks, in which people transfer information from well-known, existing categories (i.e., base or source domains) to utilize it in constructing their new idea (i.e., the target domain) (e.g., Finke et al., 1992; Gentner et al., 1997; Perkins, 1997). Analogies can play an important role in conceptual change, which is a crucial aspect of creativity (Perkins, 1997). The use of analogies is an important instrument for theory formation, design and construction (Sarlemijn and Kroes, 1988). Novelty is the degree to which an idea is original and modifies a paradigm: paradigm preserving (PP) or paradigm modifying (PM). PM ideas are sometimes radical or transformational (McCrimmon and Wagner, 1994). In addition, Bingham & Kahl (2013) noted that analogy is a highly effective tool for companies and organizations to cope with significant change and innovation.

In this study, we agree to the premises of previous studies that analogical thinking is a powerful tool for generating novel ideas by presenting the novel framework to assessing the novelty of idea generated by analogical thinking.

2.2. Superficial similarity and structural similarity in analogical thinking

Analogies are generated by superficial or structural similarities from the memory (Blanchette and Dunbar, 2000). To evaluate novelty, it is necessary to examine the similarities and differences, compatibilities and incompatibilities between elements (Berlyne, 1960).

Superficial similarity refers to the resemblance between the objects in the source and target and their properties (Keane et al., 1994). Superficial similarity refers to perceptual, contextual, and semantic attributes and features that are irrelevant to the solution principle but which are preserved in both source and target problems (Daehler & Chen, 1993). For example, when people examined two pictures in the figure 1, and were asked to indicate which object in the picture b could be match to the boy in the picture a), they are more likely to map it to the man in the picture b) based on superficial similarity while people who consider the relation among the objects and of higher-order relations between relations, i.e. structural similarities, map it to the tree in the picture b).

Blanchette and Dunbar (2000) assessed superficial similarity by the range of the domain: in their experiment, target was fixed to the deficit problem, analogies coded as within-domain were analogies for which the source was taken from the domains of either politics, economics, or personal finances. Sources from these three domains have a high degree of superficial similarity with the target problem. Analogies with sources from any other domain, such as animal, farming, or eating, were coded as other-domain analogies. Green et al. (2006; 2010) claimed that the superficial similarity is able to be measured with the Latent Semantic Analysis. LSA is a method for quantifying the similarity between words (or even whole passages) on the basis of statistical analyses of a large corpus of text (Prabhakaran & Gray, 2014). LSA is based on the vector space model from information retrieval (Salton and McGill, 1983). Here, a given corpus of text is first transformed into a term × context matrix, displaying the occurrences of each word in each context (Wandmacher, & Alexandrov, 2008).
On the other hand, Holyoak et al. (2010) defined analogy as focusing on abstract relational categories. There have been lots of studies that suggest categorization may be importantly related to analogical reasoning (Bowlde & Gentner, 2005; Gentner & Markman, 1997; Hesse, 1966; Holyoak & Thagard, 1997; Sternberg, 1977). Besides, many studies (Falkenhainer et al. 1989; Forbus et al. 1994; Forbus & Ohlinger, 1990; Genter, 1983) suggest that the structural similarity could be presented by matching the relation of each element in one idea to the other idea by introducing the structure-mapping framework. In a study of Blanchette & Dunbar (2000), they asked 49 participants to produce as many analogies as possible on the topic of pro-zero deficit and anti-zero deficit and they identified structural similarity as the resemblance in the underlying systems of relations between the elements of the sources and the target: Ten different categories of underlying structures were identified through the grouping process, five for each viewpoint (Anti and Pro-zero deficit). For instance, in the anti-zero deficit condition, the most frequently used statement structure was "If cut X, then negative consequence". In this case, the specific objects representing both "X" and "negative consequence" varied in the sources. Likewise, in the pro-zero deficit condition, the most frequently used statement structure was "If Y is not solved, and then negative consequences", and the objects matching to "Y" and "negative consequence" varied in the sources.

3. Objectives and Methodology

3.1. Research Goal

We aim to identify the novelty of ideas from analogical thinking by designing an assessment framework based on the measurement of superficial and structural similarities between the source ideas and the new idea. Three workshops and a questionnaire survey were conducted and the results were analyzed based on the proposed framework.

3.2. Innovation workshop in the i.school : Data sourcing

The workshops for this study were conducted three times on 14 international students with 11 different nationalities in the University of Tokyo, between the years of 2014-2015: N=5, N=3, N=6 for each workshop. Pre-task was given to all participants to read the 25 collective intelligence business cases study, which is a business model creating value from large and loosely organized groups of people working together electronically. The cases consist of well-known services such as Amazon.com, or Google Japanese input, as well as unfamiliar services to the participants but popular in Japan, such as Tabelog, atCosme.

During the workshop, participants were asked to categorize each case based on the underlying mechanism of the business by group discussion and confer the title of each category (see the picture a) and b) in figure 2). Consequently, participants were instructed to use the analogy table for the idea generation task. Each participant was asked to create new service idea by analogical thinking sourcing from the title of the category as well as selected case(s) upon his or her preference (see the picture c) in figure 2).
All the workshop process were recorded in video files as well as the text-note format with time lapse by the APISNOTE, which is shared sticky note software developed for the workshop (http://apisnote.com/). Participants created 25 notes in average for generation task; min=12, max=41, stdev=9.3.

Twelve ideas were created in total during the 3 workshops: 10 ideas from individual ideation and 2 ideas from collaborative ideation by two participants each. After the workshop, we selected a couple of outcomes from the workshop and conducted a questionnaire survey on the novelty perception to the third party (N=11), the respondents were asked to answer how much they could feel the idea is novel on a scale of 1-5.

3.3. Methods and Results

According to the Encyclopedia of creativity (Runco & Pritzker, 1999), the novelty is defined as original, innovative, or creative, in other words, it is also described as unusual, new and useful or domain-changing (Stokes, 1999). In this study, the novelty of idea generated from the analogical thinking is defined by superficial or structural similarities between the source cases which the participant referred and the new idea they created. To measure superficial similarities, the latent semantic analysis (LSA; Landauer, Foltz, & Laham, 1998) is used. To judge structural similarity, the comparative analysis was conducted based on the structural features between the source and the new idea they created.

- The assessment framework for novelty: the matrix of superficial × structural similarity

Many authors have argued that structural similarity is the crucial defining feature of analogy. Creativity is best realized with deeply structured representations that are relatively firm, but that admit limited, structurally guided alterations. Analogy’s power to reveal common structure and to import structure from a well-articulated domain into a less coherent domain makes it the foremost instrument of major theory change (Gentner et al., 1997). As a consequence, we propose the matrix of superficial x structural similarity for assessing the novelty of idea as shown in the figure 3.
Quadrant I, which is labelled as the novel area, is for the idea generated with representing the structure of source cases deeply in a less coherent domain, in other words it has low superficial similarity and high structural similarity from the source cases. This area is for the new ideas of participants who could generate novel ideas by importing structural features from the example cases and applying it in different domain. Quadrant II is for the idea generated with representing the structural features of source cases deeply but applying it in similar domain. Analogical thinking does not always promote domain-changing. The ideas of participants able to make matching the structural features from source case, however, unable to expand it to the new domain, are labelled as quadrant II. Quadrant III is for the ideas generated by high superficial similarity but low structural similarity. Along with quadrant II, this happens mainly due to the participants’ long-term memory. Similarity-based access to long-term memory most often produces mundane literal, i.e. superficial similarity matches (Gentner et al., 1993; Reeves & Weisberg, 1994; Ross, 1989). Quadrant IV is for the idea which generated by neither superficial nor structural similarity. It could be assumed that participants generated the idea not from during the workshop using analogical thinking as instructed, but from the inherent idea in long-term memory and failed to match the analogue with the source cases when they had to present it in an analogy table.

According to the matrix of superficial x structural similarity, six out of twelve ideas were assessed as novel ideas. Among the other six ideas, three are in quadrant II, one in quadrant III, and two are in quadrant IV. The criteria for judging the high and low of superficial or structural similarity are explained in following sub-sections.

- **The judgement of structural similarity: the categorization by mechanism**

During the workshop, categorization task was given to participants and it allows them to learn similarities between the given cases in terms of underlying mechanism, i.e. structural features. Nonetheless, the results of categorization task are not always appropriately classified because of limitations in time, knowledge, or group dynamics. It requires to present exemplary categorization to judge structural similarity, after all, four professional researchers: 1 assistant professor and 3 PhD candidates of the innovation science research group in our laboratory categorized the source cases individually. As a result, 27 category labels were created but 6 labels, which describe only single case, were removed for the cluster analysis (see the table 1). Each rater creates 6.75 labels in average (stdev=2.06) for 25 cases, after the removal of the 6 labels of single case category, the average number of labels per rater reduced to 5.25(stdev=1.26).
Table 1. The samples of categories presented by four raters

<table>
<thead>
<tr>
<th>No</th>
<th>Rater</th>
<th>Category label</th>
<th>Amazon.com</th>
<th>Kurasushi</th>
<th>Innocentive</th>
<th>atCosme</th>
<th>Rakuten Travel</th>
<th>Kopernik</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>A</td>
<td>Matching demand and supply</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>B</td>
<td>Matching parties</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>A</td>
<td>Review and Ranking</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>9</td>
<td>D</td>
<td>Using AI to forecast and enhancing</td>
<td>1</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>18</td>
<td>D</td>
<td>A platform where people share</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>20</td>
<td>C</td>
<td>Competition of creators</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Subsequently, cluster analysis was conducted to present exemplary categorization (method = median, distance = squared-Euclidian) and 21 labels were clustered into three groups at large and the structure mapping could be shown as follows (see the figure 4). Sub categories are possible within the category with specific mechanism, e.g. in the intermediation category, subcategories exist such as intermediation by competition, or matching to professional/knowledgeable.

![Cluster Analysis Diagram](image)

Finally, the new ideas generated by participants and referenced source cases or category labels were re-categorized based on the result above. Accordingly, structural similarity was assessed if the new idea is derived within the same category or not (see the table 2).

Table 2. The examples of structural similarity between the ideas generated by participants and the referenced source ideas

<table>
<thead>
<tr>
<th>The title of idea generated</th>
<th>Structure of the idea (A)</th>
<th>Referenced case (B)</th>
<th>Structure of B (C)</th>
<th>Structure written in analogy table (D)</th>
<th>Structural similarity (A ( \subseteq ) C or D)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ultimate movie recommendation</td>
<td>II-1</td>
<td>Amazon.com</td>
<td>II-1</td>
<td>III-1</td>
<td>High</td>
</tr>
<tr>
<td>Quick geek fix</td>
<td>I-1</td>
<td>Innocentive</td>
<td>I-1</td>
<td>III-1, III-3, I</td>
<td>High</td>
</tr>
<tr>
<td>Soft loan for educating the poor</td>
<td>I</td>
<td>Kopernik</td>
<td>I</td>
<td>I</td>
<td>High</td>
</tr>
<tr>
<td>Beauty map</td>
<td>II-1</td>
<td>Bike lovers’ map /at Cosme</td>
<td>III-2/III-3</td>
<td>III-1</td>
<td>Low</td>
</tr>
</tbody>
</table>
The measurement of superficial similarity: the latent semantic analysis (LSA)

As mentioned above, superficial similarity is measured by the semantic similarity between the domains of source cases referred and the created idea. Here, we have a question what the domain is. Hirschfeld and Gelman (1994) attempt to give a definition: “A domain is a body of knowledge that identifies and interprets a class of phenomena assumed to share certain properties and to be of a distinct and general type” (p. 21). As a consequence, it is also questionable that how to decide a domain. There are general domains as well as particular domains and classifying the domain in proper way may differ in a context or a point of view.

To compute similarity by the latent semantic analysis, less ambiguous terms of domain need to be selected. Furthermore, business cases are by nature domain-specific. Every business is targeting the specific users with certain products or services. Domains are usually defined in terms of content (Miller, 1996). As a consequence, we extracted words for domain using noun format in higher concept as well as in more specific concept. (see the table 3)

Table 3. Examples of the words selection for the domain

<table>
<thead>
<tr>
<th>Source cases</th>
<th>Amazon.com</th>
<th>Kura sushi</th>
<th>Innocentive</th>
<th>atCosme</th>
<th>Rakuten Travel</th>
<th>Kopernik</th>
</tr>
</thead>
<tbody>
<tr>
<td>Domain A</td>
<td>Shopping</td>
<td>Food</td>
<td>Innovation</td>
<td>Beauty</td>
<td>Travel</td>
<td>Poor</td>
</tr>
<tr>
<td>Domain B</td>
<td>Book</td>
<td>Restaurant</td>
<td>N/A*</td>
<td>Cosmetic</td>
<td>Hotel</td>
<td>Technology</td>
</tr>
</tbody>
</table>

*Innocentive looks for innovative solutions in many different fields such as medicine, agriculture, education, society, international issues.

In this study, novelty is defined as domain-changing by applying similar structural feature from the source ideas. Therefore, we compute semantic similarity between the domains of idea created by participant and all the other domains using similar structures (see the figure 4). For instance, a participant created new idea, a matching service for the computer repair, mentioning that he made reference from the Innocentive case (which provides matching clients who look for the solution by the knowledgeable people) as a source of the idea. Another example is called ‘soft loan for educating the poor’. This participant created this idea using analogy from the Kopernik case (which provides matching with donors to fund the product development and deployment project using the appropriate technology for the poor). Correspondingly, all domains in the subcategory of intermediation (see the figure 4), which is matching to professional/knowledgeable were selected to be computed using the latent semantic analysis (http://lsa.colorado.edu) website (see the figure 5). For computation option, we selected the topic space of “general reading up to first-year college (300 factors)” and term-to-term matrix comparison type. Technically, this measure of semantic similarity corresponds to the cosine of the angle between vectors corresponding to the terms within a given semantic space, which is derived through analyses of all of the contexts in which the word tends to be present or absent in that topic space (Landauer et al., 1998; see also Laham, 1997; Landauer & Dumais, 1997).

![Fig. 5. Examples of superficial similarity of two ideas calculated by LSA](image-url)
Accordingly, the superficial similarity in new idea is decided as shown in the table 4, and it is found that the former idea has low superficial similarities with source ideas while the latter has higher similarity.

Table 4. The examples of superficial similarity between the ideas generated by participants and the referenced source ideas

<table>
<thead>
<tr>
<th>Generated idea</th>
<th>Domains (A)</th>
<th>Referenced case 1</th>
<th>Referenced case 2</th>
<th>Domains (C)</th>
<th>Calculated similarity by LSA : (A) with (B) (C) Other domains in a structure*</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Quick geek</td>
<td>computer</td>
<td>Innocentive</td>
<td>Tabelog</td>
<td>food</td>
<td>0.11 0.01 0.05 0.17</td>
<td>0.17</td>
</tr>
<tr>
<td>fix</td>
<td>repair</td>
<td>innovation</td>
<td>n/a</td>
<td>restaurant</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Soft loan for</td>
<td>poor</td>
<td>Kopernik</td>
<td>n/a</td>
<td>n/a</td>
<td>1.00 n/a 0.09 1.09</td>
<td>1.09</td>
</tr>
<tr>
<td>educating</td>
<td>scholarship</td>
<td>poor technology</td>
<td>n/a</td>
<td>n/a</td>
<td>0.18 n/a 0.13 0.31</td>
<td>0.31</td>
</tr>
</tbody>
</table>

* Average value of LSA with all the other domains in the same structural category

These two sample ideas were assessed in terms of novelty by the third party (N=11, 10 students and 1 professor) using 5-scales questionnaire survey. We provided the written description of new ideas and asked how much they feel it is novel (1: not new at all to 5: very new). As a results, the former idea was rated as 3.36 in average scale (min=2, max=5, stdev=0.8) and the latter idea was rated as 2.54 (min=1, max=4, stdev=1.0). The average value of superficial similarity and average scales of novelty perception on the new idea was analyzed by single-factor ANOVA, and significant relationship was founded between the superficial similarity and people’s perception on novelty (F=27.75, P<0.05).

4. Conclusion and Discussion

This study proposes a novel assessment framework for the ideas generated using analogical thinking. The results from the study and implementation of the matrix of superficial x structural similarity for assessment provide insights into the effectiveness and benefits of our assessment framework, especially for those who study on the educational programs for promoting novel ideas. Obviously, this framework is unable to be one-size fits-all method. We used only 12 sample ideas to be assessed, and the assessing criteria are confined to an ideation tool, which is analogical thinking. In our workshop, subjects were free to select the domain of problem to be solved after attaining knowledge by reading case studies. In our real lives, in contrast, there are many cases that the domain of problem to be solved is given and a range of knowledge is limited.

In spite of this, our approach is important not only because it allows us to overcome a weakness in current assessment methods which depend on subjective judgement but also it enables further studies of how people generates novelty ideas by observing all the ideation process. In past studies, numerous kinds of measures have been used to evaluate ideas, and each measure has its own set of limitations. Specifically, in single-dimension measure of idea evaluation, raters may consciously or unconsciously include multiple constructs in a single rating (Dean, et al. 2006). Similarly, in multidimensional measures, raters also may consciously or unconsciously be influenced by other dimensions or sub-dimensions. In addition, most of past research evaluates outcomes, in other words, new ideas itself, however, our study is able to trace the source of idea as well as personal thinking process.

The ideas generated from our ideation workshops showed various level of novelty: 50% of participants could generate novel ideas but the others could not. Therefore, next step in our research is to identify the factors that enable participants to generate novel idea, such as specific skills shown during the workshop, regardless of the personal intelligence, types of thinking process for idea generation, and performance in the categorization task, prior to the idea generation task. Finally, the ultimate goal of our investigation is to design workshop processes which can enhance novelty in idea generation.
References


