FUNDAMENTAL STUDIES IN DESIGN-BY-ANALOGY: A FOCUS ON DOMAIN-KNOWLEDGE EXPERTS AND APPLICATIONS TO TRANSACTIONAL DESIGN PROBLEMS

Analogy is the process of association between situations from one domain (source) to another (target) made possible through the establishment of relations or representations (Gentner, 1983). Designs are analogous if they share at least one function or behavior, but not necessarily similar structures (Qian & Gero, 1996; Visser, 1996). Analogy association processes promote new inferences and problem understanding. Analogical association and retrieval in human cognition depend on how a problem is represented, where previous research shows that multiple representations facilitate analogical reasoning through the retrieval of effective and novel analogies stored in designers’ long-term memory (Anderson, 1983; Blanchette & Dunbar, 2000; Brown, 1989; Linsey, Murphy, Wood, Markman, & Kurtoglu, 2006; Linsey, Wood, & Markman, 2008b; McKoy, Vargas-Hernández, Summers, & Shah, 2001; Roediger, Marsh, & Lee, 2002; Vattam, Helms, & Goel, 2008).

Numerous examples of innovative systems and products based on analogies may be found in practice and in the literature, like bio-inspired products such as flippers (aquatic bird legs) or Velcro (Arctium plants). Design-by-Analogy (DbA) is an area that seeks to assist designers in identifying and developing examples, related cases and scenarios, and connected experiences (i.e., analogies) to solve design problems (Goldschmidt, 2001; Leclercq & Heylighen, 2002; Linsey, Clauss, Wood, Laux, & Markman, 2007; Linsey, Laux, Clauss, Wood, & Markman, 2007). DbA is a potentially powerful tool in idea generation (ideation), in a number of knowledge domains such as engineering design. The research reported in DbA underscore the intensity of research into creativity at the interface of cognitive science, social psychology, and knowledge domains such as engineering design (Schunn, Paulus, Cagan, & Wood, 2006; Christensen & Schunn, 2007; Tseng, Moss, Cagan, & Kotovsky, 2008b). A careful consideration of the literature indicates the need for more in-depth studies of ideation methods, the theoretical basis of these methods, and the variables or factors involved in executing these methods, especially for different knowledge domains and creative problem-solving scenarios (Weisberg, 1993; Weisberg, 2009; Jensen, Weaver, Wood, Linsey, & Wood, 2009; Jensen, et al., 2012).

Design process and method development, such as ideation, for the area of services (e.g., transactional processes) are an important and growing area of research. The importance of analyzing idea generation as part of the design process in service companies (defined in OCDE, 2010 as “retail and wholesale trade; transport and communications; real estate, finance, insurance and business services; education, health and other personal services; public administration; and defense”) lies in the fact that services as an economic activity has increased by 10% compared to products and agriculture during the last three
decades. By 2008, services accounted for more than 65% of the economic activity reported by the Organization for Economic Cooperation and Development (OECD, 2010; OECD, 2011), while in the US in 2009 services comprised more than 77% (Chesbrough, 2011; WorldBank, 2011). Based on this growth, suitable design approaches for services, and in particular transactional processes, are needed to ensure competitiveness and the development of innovation processes for this economic sector.

To understand the meaning of a transactional process or problem in these types of economies, consider a banking institution. A bank does not provide a physical system or product per se, but instead it provides an experience, a transaction, a service, which for the customer may be expressed as a fast loan approval, the reliable retrieval of money from an ATM, an easy way to access real time account balance information through the internet or phone applications, or an efficient way to pay for purchases. Transactional processes are different from products and physical systems, and, because of the socio-economic implications of these processes, the need exists to develop a deeper understanding of innovation methods to support these processes.

A number of newly developed ideation techniques and methods are emerging with supporting cognitive studies. These include directed methods and techniques such as Design-by-Analogy. Similar techniques and methods are needed to understand creative cognition in the area of service innovation, building upon the laboratory work in cognitive science and the knowledge domain studies in engineering and architecture. For these types of studies, knowledge-domain experts are preferred given the characteristics, experiences, and perspectives experts provide, especially across knowledge domains (Ball, Ormerod, & Morley, 2004; Björklund, 2012; Bonnardel & Marmèche, 2004; Casakin, 2004; Christensen & Schunn, 2007; Cross, 2004 a & b). However, studies with experts are far less prevalent in creative cognition and analogy research due to the difficulty in networking, connecting, and preparing studies within industrial or professional settings (Dixon & Johnson, 2011; Guidon, 1990; Kim, Kim, & Jin, 2005; Ozkan & Dogan, 2012). Nonetheless, the research reported here seeks to engage experts in a professional setting, focus on transactional type problems, and investigate analogical reasoning in terms of semantic word-based approaches. To the best of our knowledge, and based on the literature cited above, no such research has been published with this focus, especially considering the knowledge domain at the intersection of engineering design, business, and management.

Therefore, considering that previous research has shown effectiveness of ideation methods for improving the generation of concepts in engineering and architectural design, an important research question is:

*Can a Design-by-Analogy approach, in particular, a word-based ideation method, provide transactional domain experts the ability to increase quantity, novelty and quality while*
reducing design fixation of solutions generated for transactional design problems, compared with no intervention or use of such methods?

1. Background and Context

1.1. Services and Physical Products

There is a wide range of definitions for services (Cook, Goh, & Chung, 1999; Gadrey, Gallouj, & Weinstein, 1995; Grönroos, 1990; DISR, 1999). Authors such as De Jong, et al. (2003) have studied services and concluded that they appear to be: “intangible, simultaneously produced and consumed; and often customized to a client’s needs.” Such definitions trigger the following questions: What makes a transactional problem different when compared to a physical product or artifact problem? Are these differences significant to call for an alternative or adapted innovation process? Alternatively, do methods for product design translate directly to transactional problems?

Vermeulen (2001) presented four features that services have when they are contrasted to products: intangibility, simultaneity, heterogeneity and perishability. We adapt Vermeulen’s original comparison of services and products, as shown in Table 1, with additional data attributes for services and products to provide a deeper understanding and comparison of these two terms. We include these two rows to express in part the complexity and abstraction associated with services as opposed to products. Services are often analyzed through qualitative attributes and usually the behavior of such data does not follow a normal distribution which adds a level of complexity to its statistical analysis and data collection.

<table>
<thead>
<tr>
<th>Services</th>
<th>Products</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intangible</td>
<td>Tangible</td>
</tr>
<tr>
<td>Simultaneous production and consumption:</td>
<td>Separation of production and consumption:</td>
</tr>
<tr>
<td>costumers participate in production</td>
<td>costumers do not normally participate in production</td>
</tr>
<tr>
<td>Heterogeneous</td>
<td>Homogeneous</td>
</tr>
<tr>
<td>Perishable: cannot be kept in stock</td>
<td>Can be kept in stock</td>
</tr>
<tr>
<td>Typically require non-parametric statistics</td>
<td>Typically may be represented by normal</td>
</tr>
<tr>
<td></td>
<td>distributions (parametric)</td>
</tr>
<tr>
<td>Typically expressed through qualitative data</td>
<td>Typically expressed with quantitative data</td>
</tr>
</tbody>
</table>

De Jong, et al. (2003) and Vermeulen (2001) explore the characteristics of services and products in a “pure state,” so it seems inevitable to think of them as opposite environments and develop specific approaches catering to each one in isolation. This dichotomous view of services and products is the subject of considerable discussion and debate (Ennew, Wong, & Wright, 1992; Levitt, 1981; Vermeulen & Dankbaar, 2002). However, transactional and physical systems should not be considered as absolute
states. Instead, especially in contemporary economies, it is frequent to find that services and products are interconnected (example: service: transportation, products: GPS, vehicles, and containers); therefore, they should be considered as part of a continuum (De Jong, Bruins, Dolfsma, & Meijaard, 2003; Johne & Storey, 1998) as in Fig. 1.

![Figure 1. Illustration of the continuum of transactional to physical systems](image)

This continuum implies the potential for tools and methods for conceptual design from the domain-knowledge fields of engineering and architecture to be transferred and assist idea generation in transactional fields. Some authors have stated that the early stage development for services is no different than the conceptual design of physical products, but rather that it is at the detailed design phase where the methods diverge (Cagan & Vogel, 2013), therefore, idea generation and development of transactional and physical systems may require similar approaches.

### 1.2. Analogy

Analogy is a central concept in human cognition and creative thinking (Itkonen, 2005; Dunbar & Schunn, 1990; Dunbar, 2001). Past work indicates clear relationships between analogical reasoning and the cognitive processes associated with linguistics, long term memory retrieval, and categorization (Kalogerakis, Lüthje, & Herstatt, 2010; Schunn & Dunbar, 1996; Smith, Ward, & Schumacher, 1993; Chiu & Shu, 2007 a & b).

Previous studies in cognitive science lead to basic definitions and views of analogy as a concept. Definitions and relationships between analogy and metaphor, for example, are presented by Gentner and Markman (1997) on a coordinate design space where the axes are: “relations shared” and “attributes shared”. Hey, et al. (2007) describe the spatial area of analogy and metaphor as: “Analogous items share relational and structural similarity, while metaphors span the spectrum of relational similarity at one end, and appearance similarity at the other.”

Hey, et al. (2007) later introduced a third dimension, “purpose”, to expand the understanding of analogy and metaphor relationships within the design context. Some ideas or concepts may be represented as both an analogy and a metaphor. Metaphors can be used to understand and frame design problems by considering relevant problem context, for example, customer needs and customer feedback of a system,
product or process (Bowdle & Gentner, 2005; Casakin, 2007; Gentner, Bowdle, Wolff, & Boronat, 2001; Gentner, et al., 1997; Osterloh & Von Wartburg, 1997; Tourangeau & Sternberg, 1982). Analogy associates causal structures (system’s, device’s or process’ functional relationships, behavior, geometry or component configuration) between design problems as domain sources and possible solution target domains (Ball & Christensen, 2009; Bonnardel, 2000; Linsey, Clauss, Wood, Laux, & Markman, 2007; Linsey, Wood, & Markman, 2008b; Markman, Wood, Linsey, Murphy, & Laux, 2009).

The dimensional spectrum and relationships between analogy and metaphor provide a basis for understanding and investigating Design-by-Analogy methods for transactional problems considering that in a transactional environment, some ideas and processes may have relational similarities that can be both metaphorical and analogical. In this context, analogical reasoning may assist designers in using causal structures to enable the identification of analogous domains or particular analogies for solving a transactional problem.

1.3. Design-by-Analogy Methods

A range of DbA methods have been developed, and their sources of analogous inspirations vary form answering direct questions that allow exploration of analogical categories as in Synectics (Gordon, 1961), taking inspiration from the natural world (French, 1988; French, 1996), developing biomimetic and bio-inspired concepts (Chakrabarti & Shu, 2010; Cheong, Hallihan, & Shu, 2012; Helms, Vattam, & Goel, 2009; Mak & Shu, 2008; Nagel, Nagel, Stone, & McAdams, 2010; Singh, et al., 2009; Tinsley, Stone, McAdams, & Shu, 2008), developing analogous solutions from abstractions of functional models and flows (Hirtz, Stone, McAdams, Szykman, & Wood, 2002; Chakrabarti, et al., 2011), and exploring analogous domains by means of design problem re-representation and semantic mappings (Linsey, Markman, & Wood, 2008; Linsey, Wood, & Markman, 2008a; Smith & Linsey, 2011; Verhaegen, D’hondt, Vandevenne, Dewulf, & Duflo, 2011; Segers, De Vries, & Achten, 2005). These approaches to DbA motivate the study reported in this paper, and express a commonality in the use of linguistics and semantic transfer, either explicitly or implicitly, as a foundation for analogical reasoning.

Recent advancements in Design-by-Analogy ideation methods also include the development of analogical search approaches and search engines to identify potential analogies from digital sources, databases, and repositories (e.g., Verhaegen et al., 2011; Wu et al., 2010). One approach transforms a design problem into a functional representation and then searches patent databases using a mapped functional basis to identify near- and far-field analogies for designers to use as a basis for ideation (Murphy, 2011). Likewise, Fu et al. (2011, 2012, 2013) developed an advanced approach using a combination of Latent Semantic Analysis and a Bayesian based algorithm for discovering structural relationships of analogies, resulting in clusters of source analogies, connected by their relative similarity. Even with these particular
approaches to analogical search, the foundation is again in linguistics and semantic transfer for analogical reasoning.

1.4. Semantic Memory Retrieval

Semantic memory refers to the organization accumulation of meaningful information, and in cognitive psychological literature, is often conceptualized as a network of concepts that are associated with each other, such as through categories (Anderson, 1983; Roediger, Marsh, & Lee, 2002; Linsey, Laux, Clauss, Wood, & Markman, 2007).

For example, in Fig. 2, the concept of food storage is represented by associations as a web with nodes and links. When one thinks about food, a node becomes active, and this activation travels across other linked nodes. The activation reach weakens with distance from an activation source node.

Based on this basic model for semantic memory, a concept node will be remembered more easily if the distance (i.e. number of links traversed) shortens, or if multiple active paths converge to a specific concept node. More general concept nodes, such as “container” in Fig. 2, tend to be connected to a larger number of nodes, thus becoming hubs in the network. Linking new concepts through hubs increases the probability of being retrieved via distance shortening (Anderson, 1983; Ball, Ormerod, & Morley, 2004; Roediger, Marsh, & Lee, 2002).

This model for semantic memory and retrieval has implications on the process of developing concepts and how designers access long term semantic memory. A goal is to develop methods that increase the likelihood of retrieving solutions or ideas they have previously experienced and stored. These methods should take advantage of semantic network hubs and alternative representations that shorten the distance of accessing concept nodes.

Figure 2. Example semantic network food storage
1.5. WordTree Design-by-Analogy Method

Building from the concepts of metaphor, analogy, semantic memory retrieval, the WordTree Method is a Design-by-Analogy method that has been developed, researched and applied with successful results in engineering design. The method provides a structured approach for re-representing design problems and identifying potential analogies and analogous domains (Linsey, Markman, & Wood, 2008; Linsey, Wood, & Markman, 2008a; Linsey, et al., 2011; Linsey, Markman, & Wood, 2012; Verhaegen, D’hondt, Vandevenne, Dewulf, & Duflou, 2011).

The first step of the method consists of identifying “key problem descriptors (KPDs).” KPDs can be a design problem’s key functions, customer needs, user activities, and clarifying descriptions that, after being identified, are then linguistically re-represented in a diagram, known as a WordTree. The diagram is populated using hypernyms and troponyms of the KPDs. From this WordTree diagram, two main outcomes are identified: first, potential analogies that can be further researched, and second, analogous domains that are used to find sets of solutions in near-field or distant regions, known as far-field analogies (Chan, et al., 2011 a & b; Fu, et al., 2012). The next step is developing alternative problem statements (domain specific or general statements). The last step consists of an individual group idea generation where identified results (analogies, patents, analogous domains and problem statements) are used to refine and develop concept solutions, inspired both from the experience set and long term memory of the designer(s) and the identification and research of analogies outside this experience set such as troponyms from the WordTree diagram.

A similar method to WordTree with applications in architecture is known as the Idea Space System (ISS) (Segers & De Vries, 2003; Segers, De Vries, & Achten, 2005). This computational ideation approach was developed to support architects’ design processes. ISS captures a range of design data such as textual descriptions, sketches, and images, and uses this information to generate semantic associations by means of Princeton’s WordNet.

1.6. Divergent Tree Method

Extentics (Cai, Yang, & Lin, 2003), was developed as a method to solve contradictory problems using fuzzy sets extension methods such as the Divergent Tree Method. The method intends to expand original solution domain by using divergence. Divergence can be executed considering that artifacts and systems have and may share characteristics that can be described by associating the characteristics with a value or qualitative factors (Cai, Yang, & Lin, 2003). This concept resembles Hey et al.’s (2007) ideas for analogy and metaphor.
Extentics and the Divergent Tree Method have been applied in the following areas: (1) knowledge management and data mining, (2) product design and product innovation, (3) detection and control, (4) architectural design, (5) engineering and business management, (6) sustainable social development, (7) identification, search and diagnoses, and (8) complex system modeling (Research Institute of Extentics and Innovation Methods, 2011; Li & Yang, 2008; Zhu, Nagalingam, & Hsu, 2008).

These connections of the concept of divergence to analogy and metaphor, as well as to the WordTree Method, serves as inspiration and a clear approach to enhance the WordTree method (Linsey, Markman, & Wood, 2012) applying semantic divergence (that we defined as the use of not only hypernyms and troponyms, but antonyms, nouns, adverbs, adjectives) to establish new associations through different paths or levels of abstraction.

1.7. Summary

The use of analogical prose (such as action verbs through hypernyms and troponyms) suggests that methods such as the WordTree Method may be suitable for transactional processes due to their intangible and functional nature. This suggestion is supported by other research in linguistics (Chiu & Shu, 2007 b; Ivey & Shu, 2007).

In a transactional environment, or more generally services, some ideas and processes have relational similarities that can be both metaphorical and analogical. To harness analogical reasoning for transactional problems, the development of word/prose methods should assist designers in the mapping of knowledge and causal structures from problem domain to analogous solution domains.

The research frontier of analogical reasoning with transactional problems may now be visualized based on research from cognitive psychology, business-management research, engineering and product design, architectural design, and transactional systems.

2. Experimental Approach

Building from the literature foundation in Section 1, a set of experiments with groups of domain knowledge experts in transactional problems from 22 product and 14 service companies was conducted in Mexico, to understand the influence of a word-based ideation method on transactional design problems. The word-based ideation method is a combination of the WordTree and Divergent Tree Methods. The experiments consider a transactional design problem and focus on innovative solution generation.
2.1. Experiment Design

The experimental study reported here uses a transactional design problem with domain knowledge participants to compare a control scenario (no structured method) to an assisted scenario (method assisted generation).

Table 2. Participants' demographics and background

<table>
<thead>
<tr>
<th></th>
<th>Non-Engineering, Economics and Business related</th>
<th>Engineering</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Female</td>
<td>18</td>
<td>5</td>
<td>23</td>
</tr>
<tr>
<td>Male</td>
<td>23</td>
<td>27</td>
<td>50</td>
</tr>
<tr>
<td>Total</td>
<td>41</td>
<td>32</td>
<td>73</td>
</tr>
</tbody>
</table>

Domain knowledge participants were selected based on their professional background (transactional related) and the role of transactional processes in their companies. A group of domain knowledge experts \((n=73)\) in transactional problems from product and service industries was recruited from Lean Six Sigma training programs held in Mexico, where Table 2 displays their professional backgrounds and gender distribution. All participants have significant managerial experience, as desired for this study, and are immersed in transactional problems on a daily basis, due to their roles in their respective companies.

2.2. Transactional Design Problems

The transactional design problem statement for the study was selected in coordination with a CEO of a Lean Six Sigma consulting program. The criteria for problem statement selection was (1) the recurrence in which such problems were presented by black and green belt projects within professional development programs and consulting interactions over a number of years, and (2) transactional problems that were considered difficult to solve through traditional approaches. Fig. 3 shows a subset of the cumulative distribution of 126 transactional problems existent in the Lean Six Sigma consulting program data base from 2004 to 2012.
The highest recurrent transactional problem relates to overdue accounts/credit problems, therefore, the present study posed the following transactional problem: “Reduce overdue accounts/unpaid credits.” The purpose of the study with this problem is to concentrate only on the ideation aspect of the overall problem. Of course, a limitation of this approach is the overall representation and characterization of the process details, affordances, and rich context for a particular company or particular types of financial accounts. Future research may of course extend the findings of the study considered here, investigating other related factors (Kim Y., Lee, Maeng, & Cho, 2010; Kim K., et al., 2011).

2.3. Selection of Key Problem Descriptors

For the chosen design problem, key problem descriptors (KPD), such as functional requirements, customer requirements, user activities, and key words from the problem statement, were selected by the authors to be presented to participants during an ideation phase (method assisted generation). Participants were provided with KPDs in order to explore analogical prose aids such as hypernyms, troponyms, synonyms, and antonyms; extracted from WordTree tools such as Princeton’s WordNet® and Thinkmap’s Visualthesaurus© (Linsey, Markman, & Wood, 2008; Linsey, Wood, & Markman, 2008a; Weaver, et al., 2009; Linsey, et al., 2012). The alternative KPDs that participants developed where intended to stimulate the memory retrieval process, as the first step of the WordTree Method, for analogical reasoning. Selected KPDs are presented in Table 3.
Table 3. Selected KPDs for Transactional Design Problem

<table>
<thead>
<tr>
<th>Transactional Design Problem</th>
<th>Reduce overdue accounts</th>
</tr>
</thead>
<tbody>
<tr>
<td>• Assure</td>
<td></td>
</tr>
<tr>
<td>• Increase</td>
<td></td>
</tr>
<tr>
<td>• Guarantee</td>
<td></td>
</tr>
<tr>
<td>• Pay /payment</td>
<td></td>
</tr>
<tr>
<td>• Reward</td>
<td></td>
</tr>
</tbody>
</table>

2.4. Experiment Execution

The overall experiment execution is shown in Fig. 4. Participants were enrolled in a 5 day Lean Six Sigma training program and participated in the experiment as an additional activity. During the first phase of the experiment (intuitive generation), all participants, working individually, were provided with the following instructions:

“Consider the problem given below. You will be given up to 3 minutes to read this information, followed by 15 minutes to create solutions to the problem. Please do not start to write down solutions until the 15 minute period has started. Your goal is to create as many solutions to the problem as possible. Present your solutions as a phrase, written description, and / or sketch / diagram as you desire. You will be given a five minute and a one minute warning before your time is up.

Problem: Reduce Overdue Accounts / Unpaid Credits

Please feel free to record any thoughts or comments that you might have as you develop each solution. In particular, please record any motivating related problems, similar solutions, or example that may have motivated you, if they exist and can be articulated.”

For this first phase, participants were only allowed to use their own knowledge and creativity (no other tool or software was allowed at this point). Participants had 15 minutes to complete this task.
Figure 4. Experimental execution diagram

A second phase was programmed two days later to activate long term memory analogy retrieval, where the participants were divided into an experimental group and control group (Table 4), and their members distributed according to demographics, such as gender, professional degree and employment role. Before the start of the second phase, participants in the experimental group, referred to as WT “With Technique” ($n_{WT}=37$), were taken to a different room and taught the combined WordTree and Divergent Tree Method, as well as the associated software, Princeton’s WordNet® and Thinkmap’s Visualthesaurus® (Linsey, Markman, & Wood, 2008; Linsey, Wood, & Markman, 2008a; Linsey, Markman, & Wood, 2012; Weaver, et al., 2009). The training of the method and software was carried out in 15 minutes, while the control group for the experiment, referred to as NT “No Technique” ($n_{NT}=36$), continued their Lean Six Sigma training program.

<table>
<thead>
<tr>
<th></th>
<th>WT</th>
<th>NT</th>
</tr>
</thead>
<tbody>
<tr>
<td>Female</td>
<td>12</td>
<td>11</td>
</tr>
<tr>
<td>Male</td>
<td>25</td>
<td>25</td>
</tr>
<tr>
<td>Total</td>
<td>37</td>
<td>36</td>
</tr>
</tbody>
</table>

Table 4. Experimental group’s gender distribution

During the second phase, the control group (NT), continued solving the same problem, under the same conditions as the first phase. The experimental group (WT) group was asked to continue solving the problem with the provided method and software. During this task, experimental group participants were asked to explore at least three levels away from their initially selected nouns, adverbs, verbs, adjectives or antonyms related to the KPDs, customer needs and functional requirements that were presented in the template materials that allowed them to re-represent and develop analogical solutions for the design
problem. These participants were also asked to generate as many solutions as possible while extracting useful information from the method and software. Both experimental and control groups were given an additional 15 minutes to generate solutions to the transactional problem. Independence between both WT and NT groups was achieved due to the physical separation during phase II.

3. Ideation Metrics

Once phase I and II concluded, all ideas recorded by participants were extracted from the provided materials, and a total of 1,133 ideas were identified. The purpose of the study is to explore the effects in ideation performance after introducing a semantic word-based ideation method to solve transactional design problems. Five ideation metrics were chosen to evaluate the results of the study: (1) semantic solution transfer, (2) quantity of ideation, (3) design fixation, (4) novelty, and (5) quality of solution concepts. These ideation metrics have been previously used to formally study ideation in the engineering knowledge domain (Girotra, Terwiesch, & Ulrich, 2010; Linsey, et al., 2011; Linsey, Markman, & Wood, 2012; McAdams & Wood, 2002; Oman, Tumer, Wood, & Seepersad, 2013; Shah, Kulkarni, & Vargas-Hernandez, 2000; Shah, Smith, & Vargas-Hernandez, 2003; Srivathsavai, Genco, Höttä-Otto, & Seepersad, 2010). We first revisit previous approaches and definitions, so we can then build upon and adapt them to the particularities of the transactional domain problems and semantic structure of the collected data. Through this process, we will have robust and comparable results to the previously reported findings in the engineering and related knowledge domains.

The listed metrics, with the exception of the third, were previously explored by Chan et al. (2011a) for an engineering design problem. When transferred to transactional problems, they provide insights about the way participants: (1) process the information using the imparted method, (2) use the prescribed method to generate solutions (ideation process), and (3) embed intrinsic value (meeting customer and/or process requirements, breakthrough concept, originality, etc.) to the solutions generated.

Analyzing semantic solution transfer of word re-representations to explore analogies is important for developing a better comprehension of the stimulus provided by the technique. Quantity of ideation is a metric that enables quantification of the exploration level of the design space (Chan, 2011a). Fixation provides a measure to evaluate the effect of the method in preventing design fixation from appearing in the participants, by contrasting the number of times ideas generated are repeated (reappearing) with the total of ideas developed (Linsey, et al., 2010). Novelty provides a measure of the creativity/originality level of a given solution (Markman & Wood, 2009), i.e. its uniqueness or originality within a context. Finally, we also included a quality evaluation because no matter how novel an idea might be, if it does not meet customer needs, process specifications, and technical and economic feasibility, it will be discarded by the customer or solution implementer (Markman & Wood, 2009).
3.1. Data setup

Participants recorded their ideas by means of bulleted lists, flow or process diagrams, charts, storyboards, or combinations of textual and graphical illustrations. Fig. 5 shows selected examples of participants’ recorded ideas.

Figure 5. Selected examples of participants that generated large amount of concept solutions by experimental group

A procedure for organizing and formatting the data was executed for post-experiment analysis. At the end of each phase, participants’ solutions sheets were collected, and the recorded ideas were evaluated by two different domain knowledge expert raters (one from a business background and the other from industrial engineering, both with extensive banking experience). To accomplish this task, the raters independently grouped the solution statements into bins of distinctive ideas resulting in 129 bins generated by the two raters with a calculated Cohen’s kappa (Von Eye & Mun, 2005) of 0.79, considered an “excellent” level (Robson, 2002), where the data and objectivity of the evaluation approach should be trusted.
After independently recognizing the initial 129 bins, the raters resolved the few remaining disagreements through discussion, resulting in a final 117 distinctive bins (ideas). Fig. 6 shows a frequency plot of the resulting bins, where it may be observed that there is no differential frequency “jump” between ideas, and that 43% (50/117) of the bins have a relatively low frequency ($F \leq 4$).

To assess if the 117 bins completely define the solution space of ideas, a plot of the cumulative chronological contribution made by the participant groups of the study (i.e., groups of participants that were included in the study at different chronological time) is presented in Fig. 7, where it is shown that the largest increments occurred at the beginning and became marginal at the end of the study.

### 3.2. Semantic solution transfer

Semantic solution transfer is here defined as the extent to which participants transferred semantic representations to analogical solutions. There is no unique quantitative score for this metric, because only a subset of the participants explicitly recorded the word or path of words followed to develop their solutions.

Therefore, a mixed quantitative and qualitative analysis for the semantic solution transfer was developed by pairing participants’ recorded list of inspirational words (for the subset of participants who did record them) with their corresponding solution ideas. The ideas were then mapped to their corresponding bin and frequency. The total number of ideas generated in phase II of the experiment was also recorded.

Participants’ listed words provide insights about the participants’ stimuli received by the technique (experimental group) and their retrieval process, i.e., if the solutions were developed using domain distant
words or were the same as originally provided. Frequency is an indicator of the relative novelty of the ideas. A ratio of the paired “selected word” to “idea developed” with respect to the total ideas generated in phase II provides insights about the relative efficiency on the method’s use.

3.3. Quantity of ideation

Some definitions and procedures to calculate quantity of concepts have been developed in the domain of engineering design (Bouchard & Hare, 1970; Dean, Hender, Rodgers, & Santanen, 2006; Linsey, et al., 2011; Oman, Tumer, Wood, & Seepersad, 2013; Shah, Smith, & Vargas-Hernandez, 2003). Building upon these definitions, we define an “idea” as any form of statement and/or diagram that provides an operational method or solution for the transactional design system to be solved (to accomplish the goal).

Two primary mathematical representations for quantity of ideas are considered in the study reported here: (1) Quantity of Total ideas \((Q_{\text{Total}})\), and (2) Quantity of Non-Repeated ideas \((Q_{\text{NR}})\). Eq. 1 shows the interrelationship between them:

\[
Q_{\text{Total}} = \sum \text{all ideas generated} = Q_{\text{NR}} + \text{Repeated ideas} \tag{1}
\]

Quantity of total ideas generated is expressed as the summation of all ideas generated (Eq. 1), at different levels, such as in phase (I, II), across experimental groups (WT, NT), and per participant. Thus, no specific indices are listed for the summation operator of Eq. 1.

From Eq. 1 we also have an alternative definition for \(Q_{\text{Total}}\) that breaks it down into two main components: Quantity of Non-Repeated ideas \((Q_{\text{NR}})\) and repeated ideas \((Q_{\text{R}})\). Quantity of non-repeated ideas \((Q_{\text{NR}})\) corresponds to filtered data, since it takes into account all ideas generated at different levels that were not repeated. A repeated idea occurs when a participant states an idea more than once (usually due to a slight variation or rewording of the idea). A more detailed description and classification on repeated ideas for this study is provided in Section 4.4.

3.4. Fixation

Jansson and Smith (1991) define design fixation as “a blind adherence to a set of ideas or concepts limiting the output of conceptual design.” Building on this definition, research in the area of fixation indicates root causes of design fixation and the foundational elements to develop associated metrics. Exemplar types of fixation may be due to unawareness of technological advances and conformity due to supporting technologies of an existing solution (Luchins & Luchins, 1959), design expertise (Linsey, et al., 2010), cognitive processes (Smith & Blankenship, 1991), as well as conceptual and knowledge fixation at
different levels of consciousness (Youmans & Arciszewski, 2012). Since repeating ideas was an observed phenomenon that occurred while participants generated solutions for the transactional design problem, we decided to assess fixation, introducing a quantity metric based on the procedure outlined by Linsey, et al. (2010) and further developed by Vimal & Linsey (2012), i.e., based on quantity of repeated ideas.

We define that a repeated idea occurs when a given participant states, on more than one occasion, an idea (entirely, as a slight variant, or by rewording it). There are two different sources for repeated ideas in the study:

- Repeated ideas within a phase \( (R_w) \): defined as the summation of all repeated ideas across all participants that have a frequency \( (F) \) greater than 1 as shown in Eq. 2.

\[
R_w = \sum_{i=1}^{b} \sum_{k=1}^{n} F_{ijk} - 1 \quad \forall \quad F_{ijk} > 1
\]  

where \( F_{ijk} \) = frequency of repeated ideas for the \( i \)th phase, \( j \)th bin, and \( k \)th participant; \( i \) = phase number (1, 2); \( b \) = number of bins (117); \( n \) = number of participants. A unit is subtracted from \( F_{ijk} \), to maintain accountability of the total of ideas generated.

- Repeated ideas between phases \( (R_B) \): for bin and participant levels, \( R_B \) takes into account all ideas that were repeated in phase II when compared to phase I. Eq. 3 includes the condition expressed previously that the bin-idea had to be previously generated in phase I and that it has to appear in phase II.

\[
R_B = \sum_{i=1}^{b} \sum_{k=1}^{n} F_{2jk} \quad \forall \quad F_{1jk} > 1 \quad \text{AND} \quad F_{2jk} > 0
\]  

where \( F_{ijk} \) = frequency of repeated ideas for the \( i \)th phase, \( j \)th bin, and \( k \)th participant; \( i \) = phase number (1, 2); \( b \) = number of bins (117); \( n \) = number of participants.

Using the definitions and results for quantity of repeated ideas, an operational fixation definition is implemented as shown in Eq. 4:

\[
\text{Fixation} = \frac{\text{Total \# of repeated ideas}}{\text{Total \# of generated ideas}} = \frac{R_w + R_B}{Q_{\text{Total}}}
\]  

The mathematical definition of fixation, as stated in Eq. 4, provides the ability to perform statistical comparisons between both the experimental and control groups in order to discover the usefulness and
effectiveness of the method to overcome fixation when developing solutions for a transactional design problem.

3.5. Novelty

Novelty provides a measure of the uniqueness or originality of a given solution when contrasted/compared with others in the design space of possible solutions. Some novelty metrics, such as the one developed by Jansson and Smith (1991), referred to as “Originality” (Eq. 5) and another adapted by Chan (2011a), measure the non-similar or non-related concepts with respect to the total of concept ideas generated.

\[
Originality = 1 - \frac{\text{number of similar ideas (designs) generated by participants}}{\text{total number of ideas (designs) for all participants}}
\]  

(5)

Inspired by these definitions, novelty metrics were computed through three approaches considering phase I as defining the design space baseline and using only the total quantity of non-repeated \((Q_{nr})\) ideas:

1. Define novelty as the space of all bins that were uniquely generated in phase II (shadow area in Fig. 8). This definition is adopted because these uniquely generated bins constitute a tangible expansion of the design space.

2. Define novelty as a function of the space composed of all ideas (not bins) that were uniquely generated in phase II as a function of the experimental group level over the total phase II space design as shown in Eq. 6.

\[
Novelty_k = \frac{\sum_{p=1}^{n} F_{2jk} \land F_{1jk}=0 \text{ AND } F_{2jk}>0}{\sum_{p=1}^{n} F_{2jk}}
\]  

(6)
where \( F_{ijk} \) = frequency of ideas for the \( i \)th phase, \( j \)th participant, and \( k \)th group (control, NT; or experimental, WT); \( i \) = phase number (1, 2); \( k \) = group (NT, WT); \( n \) = number of participants.

3. Define novelty as a function of the space composed of all ideas (not bins) that were uniquely generated by a participant in phase II over the participant’s total phase II design space of ideas.

\[
Novelty_{j,k} = \frac{\sum_{b=1}^{b_{\text{total}}} F_{2jkl} \quad \forall \ F_{1jkl}=0 \quad \text{AND} \quad F_{2jkl}>0}{\sum_{b=1}^{b_{\text{total}}} F_{2jkl}}
\] (7)

where \( F_{ijk} \) = frequency of ideas for the \( i \)th phase, \( j \)th participant, \( k \)th group, and \( l \)th bin; \( i \) = phase number (1, 2); \( j \) = participant number (1,...,73); \( k \) = group (NT, WT); \( b \) = number of bins (117)

3.6. Quality

There are different ways to evaluate the quality of solution ideas generated for engineering design problems (Verhaegen, D’hondt, Vandevenne, Dewulf, & Duflou, 2011) by means of criterion or dimensions such as: technical feasibility and conformance to specifications (Shah, Kulkarni, & Vargas-Hernandez, 2000; Shah, Smith, & Vargas-Hernandez, 2003), workability, relevance, specificity, and novelty (Dean, Hender, Rodgers, & Santanen, 2006), or implementability scales (Linsey, et al., 2011; Linsey, Clauss, Wood, Laux, & Markman, 2007; Linsey, Laux, Clauss, Wood, & Markman, 2007).

However, due to the specific characteristics of transactional problems and the open-ended nature of the generated solutions, an alternative approach to perform a quality analysis for transactional design problem solutions is proposed here via Benchmarking, i.e. comparing the study results with recent, notable ideas that have been published in the innovative banking industry (and that were unknown to participants).

The qualitative evaluation for quality of solutions generated was developed comparing publicly available innovative and award winning solutions of Citibank (Citigroup Inc., 2006; Citigroup Inc., 2011; National Infocomm Awards, 2012), and of Westpac Bank (Westpac Banking Corporation, 2012).

4. Results

4.1. Statistical data validation

The number of chosen expert participants as a sample for this study was a controlled variable that depended on the number of participants enrolled in specific lean six sigma trainings. As a means to
understand and validate the power of the tests performed, we developed a retrospective power and sample size study.

The more powerful the statistical test, the more likely a Type II error (accepting the null hypothesis when it is false) may be avoided (Clark-Carter, 2010). There are factors that influence power such as (1) the probability $\alpha$ of a Type I error (level of significance), (2) the variability $\sigma$ in the population, (3) the minimum difference (corresponding to the difference between population means the study will be able to detect), and (4) sample size.

For two-sample t-tests, the sample size, as a function of chosen power, was first evaluated by setting the significance level as $\alpha=0.05$ (as typically chosen for similar cognitive studies), the power as 0.8 (as typically reasonable value within social science studies), the variance depending on the metric being evaluated, and the minimum difference by using a low (L) and a high (H) value for consideration. Results for sample size are displayed in the left section of Table 5.

<table>
<thead>
<tr>
<th>$\beta$=0.8, $\alpha$=0.05</th>
<th>Standard Deviation</th>
<th>Difference</th>
<th>Prospective Sample Size</th>
</tr>
</thead>
<tbody>
<tr>
<td>$Q_{Total}$</td>
<td>2.5</td>
<td>$L=2$</td>
<td>26</td>
</tr>
<tr>
<td></td>
<td></td>
<td>$H=3$</td>
<td>12</td>
</tr>
<tr>
<td>$Q_{NR}$</td>
<td>3</td>
<td>$L=2$</td>
<td>37</td>
</tr>
<tr>
<td></td>
<td></td>
<td>$H=3$</td>
<td>17</td>
</tr>
<tr>
<td>Fixation</td>
<td>0.25</td>
<td>$L=0.18$</td>
<td>32</td>
</tr>
<tr>
<td></td>
<td></td>
<td>$H=0.3$</td>
<td>26</td>
</tr>
<tr>
<td>Novelty</td>
<td>0.5</td>
<td>$L=0.35$</td>
<td>34</td>
</tr>
<tr>
<td></td>
<td></td>
<td>$H=0.45$</td>
<td>21</td>
</tr>
</tbody>
</table>

Table 5. Power and sample size analysis

<table>
<thead>
<tr>
<th>Difference</th>
<th>Actual Sample Size</th>
<th>Retrospective Power</th>
</tr>
</thead>
<tbody>
<tr>
<td>2.5</td>
<td>$NT=36$</td>
<td>0.98</td>
</tr>
<tr>
<td></td>
<td>$WT=37$</td>
<td>0.98</td>
</tr>
<tr>
<td>2.5</td>
<td>$NT=36$</td>
<td>0.93</td>
</tr>
<tr>
<td></td>
<td>$WT=37$</td>
<td>0.94</td>
</tr>
<tr>
<td>0.2</td>
<td>$NT=36$</td>
<td>0.91</td>
</tr>
<tr>
<td></td>
<td>$WT=37$</td>
<td>0.92</td>
</tr>
<tr>
<td>0.4</td>
<td>$NT=36$</td>
<td>0.91</td>
</tr>
<tr>
<td></td>
<td>$WT=37$</td>
<td>0.92</td>
</tr>
</tbody>
</table>

Based on this analysis, it is found that for the desired differences, the study’s actual sample sizes are sufficient to assure at more than an 80% power. When performing the analysis using the study’s actual sample sizes (L=sample size of NT group, H=sample size of WT group), it is found that all power values are higher than 90% for all metrics (Table 5) at a confidence level of 95% with the desired difference in outcome.

The assumptions for conducting relevant comparative sample tests, i.e., normality of data across techniques were met and evaluated by means of Anderson Darling Normality Test. For the cases where its value was less than 0.05, the experimental study produced data that reasonably fits a log-normal distribution; therefore, statistical analysis can be performed without transforming the data, especially considering that ANOVA has a degree of robustness for departures from normality (Table 6).
Table 6. Normality test

<table>
<thead>
<tr>
<th></th>
<th>$Q_{\text{Total}}$</th>
<th>Repeated ideas</th>
<th>Fixation</th>
<th>Novelty</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>NT</strong></td>
<td>~N (p-value=0.248)</td>
<td>~N (p-value=0.022)</td>
<td>~N (p-value=0.141)</td>
<td>~N (p-value&lt;0.05)</td>
</tr>
<tr>
<td></td>
<td>~LogN (p-value=0.248)</td>
<td>~LogN (p-value=0.245)</td>
<td>~LogN (p-value=0.255)</td>
<td>~LogN (p-value=0.577)</td>
</tr>
<tr>
<td><strong>WT</strong></td>
<td>~N (p-value=0.018)</td>
<td>~N (p-value&lt;0.05)</td>
<td>~N (p-value=0.05)</td>
<td>~N (p-value&lt;0.05)</td>
</tr>
<tr>
<td></td>
<td>~LogN (p-value=0.117)</td>
<td>~LogN (p-value&lt;0.05)</td>
<td>~LogN (p-value=0.425)</td>
<td>~LogN (p-value=0.385)</td>
</tr>
</tbody>
</table>

4.2. Semantic solution transfer

We now present semantic solution transfer results of the mixed quantitative and qualitative analysis (Table 7) for the participants that recorded each instance of the word or path of words followed to develop their solutions. Method efficiency results are greater than or equal to 80% for all participants. We thus infer that, in general, ideas generated in phase II were indeed fostered by the method.
<table>
<thead>
<tr>
<th>Participant</th>
<th>Method Efficiency</th>
<th>Bin #</th>
<th>Freq</th>
<th>Words Recorded</th>
</tr>
</thead>
<tbody>
<tr>
<td>P1</td>
<td>4/5=0.8</td>
<td>110</td>
<td>7</td>
<td>Teach/train</td>
</tr>
<tr>
<td></td>
<td></td>
<td>35</td>
<td>18</td>
<td>Motivate</td>
</tr>
<tr>
<td></td>
<td></td>
<td>76</td>
<td>15</td>
<td>Prepayment</td>
</tr>
<tr>
<td></td>
<td></td>
<td>98</td>
<td>32</td>
<td>Reward</td>
</tr>
<tr>
<td></td>
<td></td>
<td>41</td>
<td>6</td>
<td>Train</td>
</tr>
<tr>
<td></td>
<td></td>
<td>96</td>
<td>10</td>
<td>Modify</td>
</tr>
<tr>
<td></td>
<td></td>
<td>79</td>
<td>6</td>
<td>Consequence</td>
</tr>
<tr>
<td></td>
<td></td>
<td>5</td>
<td>23</td>
<td></td>
</tr>
<tr>
<td>P2</td>
<td>5/5=1</td>
<td>57</td>
<td>5</td>
<td>Reward</td>
</tr>
<tr>
<td></td>
<td></td>
<td>40</td>
<td>7</td>
<td>Train</td>
</tr>
<tr>
<td></td>
<td></td>
<td>89</td>
<td>14</td>
<td>Guarantor</td>
</tr>
<tr>
<td></td>
<td></td>
<td>98</td>
<td>32</td>
<td>Reward</td>
</tr>
<tr>
<td></td>
<td></td>
<td>104</td>
<td>1</td>
<td>Simplify, persuade, motivate</td>
</tr>
<tr>
<td>P3</td>
<td>4/4=1</td>
<td>89</td>
<td>14</td>
<td>Guarantee</td>
</tr>
<tr>
<td></td>
<td></td>
<td>59</td>
<td>27</td>
<td>Compensate, split</td>
</tr>
<tr>
<td></td>
<td></td>
<td>72</td>
<td>28</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>6</td>
<td>5</td>
<td>Change</td>
</tr>
<tr>
<td></td>
<td></td>
<td>21</td>
<td>5</td>
<td>Change</td>
</tr>
<tr>
<td>P4</td>
<td>5/5=1</td>
<td>90</td>
<td>12</td>
<td>Change</td>
</tr>
<tr>
<td></td>
<td></td>
<td>72</td>
<td>28</td>
<td>Penalize</td>
</tr>
<tr>
<td></td>
<td></td>
<td>98</td>
<td>32</td>
<td>Reward</td>
</tr>
<tr>
<td></td>
<td></td>
<td>114</td>
<td>2</td>
<td>Teach/train</td>
</tr>
<tr>
<td></td>
<td></td>
<td>91</td>
<td>8</td>
<td>Design</td>
</tr>
<tr>
<td></td>
<td></td>
<td>76</td>
<td>15</td>
<td>Prepayment</td>
</tr>
<tr>
<td></td>
<td></td>
<td>76</td>
<td>15</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>71</td>
<td>1</td>
<td>Penalize</td>
</tr>
<tr>
<td></td>
<td></td>
<td>59</td>
<td>27</td>
<td>Settle</td>
</tr>
<tr>
<td></td>
<td></td>
<td>34</td>
<td>2</td>
<td>Compensate</td>
</tr>
<tr>
<td>P5</td>
<td>10/11=0.9</td>
<td>110</td>
<td>7</td>
<td>Teach/train</td>
</tr>
<tr>
<td></td>
<td></td>
<td>35</td>
<td>18</td>
<td>Motivate</td>
</tr>
<tr>
<td></td>
<td></td>
<td>76</td>
<td>15</td>
<td>Prepayment</td>
</tr>
<tr>
<td></td>
<td></td>
<td>98</td>
<td>32</td>
<td>Reward</td>
</tr>
<tr>
<td></td>
<td></td>
<td>41</td>
<td>6</td>
<td>Train</td>
</tr>
<tr>
<td></td>
<td></td>
<td>96</td>
<td>10</td>
<td>Modify</td>
</tr>
<tr>
<td></td>
<td></td>
<td>79</td>
<td>6</td>
<td>Consequence</td>
</tr>
<tr>
<td></td>
<td></td>
<td>5</td>
<td>23</td>
<td></td>
</tr>
</tbody>
</table>

When considering the frequency of the bin-idea associated with a solution, it is clear that four of the bins have a frequency of one (highlighted cells in columns 3 and 4). These unique, novel ideas were thus inferred to be developed using the method.

An interesting result is that out of the 11 KPDs provided, only five (5) were explicitly recorded as the ones used to generate solution (highlighted cells in column 5). A total of 55 words were recorded as presented.
in Table 7, and excluding the ones matching the KPDs, the remaining 50 (91%) of the words were the result of the semantic search from the combined WordTree and Divergent Tree method.

Some examples of the semantic search carried out by means of software tools such as Thinkmap’s Visualthesaurus© are displayed in Fig. 9. The red circles show some of the information participants selected to develop solutions starting from KPDs or other words (placed in the middle).

4.3. Quantity of ideation

Quantity of total ideas generated \( (Q_{Total}) \) as well as quantity of non-repeated ideas \( (Q_{NR}) \) for each phase and group type (WT or NT) are shown in Table 8. \( Q_{NR} \) constitutes the filtered data used for metric calculations such as novelty. \( Q_{Total} \) has a total of 1,133 ideas, while \( Q_{NR} \) has a total of 817 ideas, distributed between the experimental phases and groups as presented in Table 8.

<table>
<thead>
<tr>
<th>Table 8. Quantity of generated ideas</th>
</tr>
</thead>
<tbody>
<tr>
<td>( Q_{Total} )</td>
</tr>
<tr>
<td>286</td>
</tr>
<tr>
<td>( Q_{NR} )</td>
</tr>
<tr>
<td>239</td>
</tr>
</tbody>
</table>

An ANOVA shows no statistically significant difference in the quantity of ideas generated in Phase I for both groups WT and NT \( (Q_{Total}: F=1.82, \text{p-value}=0.182, \text{and } Q_{NR}: F=2.75, \text{p-value}=0.102) \) which indicates a level of consistency in the performance of the groups in a non-assisted scenario. A paired \( t \)-test for \( Q_{Total} \), phase II compared to phase I for the NT (control) group shows no statistically significant difference \( (t\text{-value}= 0.08 \text{ p-value}= 0.940) \), which is to be expected considering that phase II for this group is also
non-assisted. A paired t-test for $Q_{Total}$ phase II compared to phase I for the WT (experimental) group shows a statistical significant difference ($t$-value = -3.37, $p$-value = 0.002). These results indicate that there is a distinct quantity difference between the phases for the WT group. In terms of $Q_{Total}$, this result is explained by the existence of a cognitive load on phase II for WT, leading to a lower quantity of total ideas. This result and interpretation are consistent with previous cognitive studies where the intervention adds significant time due to cognitive processing (Chan, 2011a & b; Tseng, Moss, Cagan, & Kotovsky, 2008a & b). Likewise, for $Q_{NR}$, a paired t-test of phase II compared to phase I for the NT and WT groups shows a statistical significant difference in the quantity of ideas for both scenarios (NT: $t$-value = -4.97, $p$-value = 0.000, and WT: $t$-value = -4.19, $p$-value = 0.000).

4.4. Fixation

The resulting distribution of repeated ideas within and between phases over both groups and phases is summarized in Table 9.

<table>
<thead>
<tr>
<th>Table 9. Repeated ideas by group, source and phase</th>
</tr>
</thead>
<tbody>
<tr>
<td>[</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>Total repeats within</td>
</tr>
<tr>
<td>Total repeats between</td>
</tr>
<tr>
<td>TOTAL</td>
</tr>
<tr>
<td>Ratio by participant</td>
</tr>
</tbody>
</table>

A set of t-tests identify statistically significant differences in the quantity of repeated ideas. When comparing the results obtained in phase I and II for the WT group, no statistically significant difference is found, i.e., the number of repeated ideas was the same ($t$-value = 0.45, $p$-value = 0.658). This is not the case for NT group, where a significant difference in the quantity of repeated ideas is found ($t$-value = 6.63, $p$-value = 0.000). Finally, a two sample t-test comparing phase I of both experimental groups (NT, WT) shows no statistical significant difference in the quantity of repeated ideas ($t$-value = 0.06, $p$-value = 0.953).

After applying the mathematical expression for Fixation (Eq. 4) to each group, we have the results illustrated in Table 10:

<table>
<thead>
<tr>
<th>Table 10. Fixation (%) by Phases of both Groups</th>
</tr>
</thead>
<tbody>
<tr>
<td>[</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>Group</td>
</tr>
</tbody>
</table>
A two sample $t$-test between phase I of both the experimental (WT) and control (NT) groups shows no statistical significant difference in fixation as defined by Eq. 4. ($t$-value= 0.89, $p$-value=0.376), which means that there is a certain base level of fixation in non-assisted scenarios. A two sample $t$-test between phase II of both the experimental (WT) and control (NT) groups shows a statistical significant difference in fixation ($t$-value=-4.33, $p$-value=0.000), which is an indicator of a significant change (% reduction in fixation by the WT experimental group) when applying the method. This implies that the introduced method assists in mitigating fixation.

4.5. Novelty

For the first novelty approach, we have a total of 10 bins uniquely generated in phase II, distributed as NT=3 and WT=7.

Table 11 presents the total number of non-repeated ideas uniquely generated in phase II ($Q_{NR}$) using the second approach for evaluating novelty, as well as the respective novelty calculated value applying Eq. 6.

<table>
<thead>
<tr>
<th></th>
<th>Total # of Ideas ($Q_{NR}$)</th>
<th>Novelty</th>
</tr>
</thead>
<tbody>
<tr>
<td>NT</td>
<td>4</td>
<td>1.35%</td>
</tr>
<tr>
<td>WT</td>
<td>9</td>
<td>3.03%</td>
</tr>
</tbody>
</table>

Applying the expression to evaluate novelty in Eq. 7, by participant, we may calculate average and standard deviation results as shown in Table 12 for a total number of non-repeated ideas uniquely generated in phase II ($Q_{NR}$) and novelty of both groups.

<table>
<thead>
<tr>
<th></th>
<th>Total # of Ideas ($Q_{NR}$)</th>
<th>Novelty</th>
</tr>
</thead>
<tbody>
<tr>
<td>WT</td>
<td>Total # of Ideas ($Q_{NR}$)</td>
<td>Novelty</td>
</tr>
<tr>
<td>Average</td>
<td>0.24</td>
<td>4.5%</td>
</tr>
<tr>
<td>StDev</td>
<td>0.55</td>
<td>0.11</td>
</tr>
<tr>
<td>NT</td>
<td>Total # of Ideas ($Q_{NR}$)</td>
<td>Novelty</td>
</tr>
<tr>
<td>Average</td>
<td>0.11</td>
<td>0.9%</td>
</tr>
<tr>
<td>StDev</td>
<td>0.32</td>
<td>0.03</td>
</tr>
</tbody>
</table>

An ANOVA is applied to determine if statistically significant differences exist in the quantity of ideas uniquely generated in phase II for both groups, and the novelty calculated value (Eq. 7). There is no statistically significant difference between the total number of ideas generated by both groups ($F=1.57$, $p$-value=0.214). The ANOVA comparing the calculated novelty values of both groups shows a statistically significant difference ($F=3.71$, $p$-value=0.058).
4.6. Qualitative Approach for Evaluating Quality

The qualitative analysis for evaluating quality of the generated solutions is presented in Fig. 10, where the first and third column present solutions developed and presented as innovations by leading banking corporations (Citigroup Inc., 2006; Citigroup Inc., 2011; Westpac Banking Corporation, 2012). The middle column and blue arrows represent a mapping from the bin-ideas developed in the study. The solutions under the study column are highlighted in green if they are not novel according to study definitions but can be mapped to the innovative solutions of the leading banking corporations or yellow if according to the study they are considered novel. In Fig. 10 it is shown that there are seven (7) yellow boxes that correspond to the novel bins presented in Section 4.5. In the study boxes, the assigned bin number is also listed according to the labels generated as part of the study.

Figure 10. Qualitative analysis of innovative transactional solutions for problem used (Citigroup Inc., 2006; Citigroup Inc., 2011; Westpac Banking Corporation, 2012)
The analysis shows that solutions proposed by study's participants are very closely related to innovative results recently implemented by leading banking corporations. These banking innovative solutions were published outside of the participants’ industry domain, have been noted as innovations in their field, and are very recent developments. These results provide validation of the potential that this Design-by-Analogy approach has for generating valuable (high quality) ideas for transactional problems.

The solutions listed under the “Study” column (Fig. 10) are, in general, more tangible-specific than those presented by the corporations. The lower seven bins in this column are novel according to the study, however, some of them have no immediate correspondence with those implemented by corporations and may have the potential for solving the design problem in an innovative way, such is the case of “change staff (role swap, rotation),” where this radical approach consists of exchanging sales personnel with debt collection staff at specified time intervals, so all staff are more aware of the actual difficulties of collecting debt payments when customers are not carefully selected. Likewise, the idea for “specialists for sales (innovative selection and training),” where the twist will consist in training employees, instead of customers, using more innovative training programs that will make sales staff aware of the system implications of their work beyond commissions.

**4.7. Relationships between metrics**

After performing the evaluation of the selected ideation metrics and considering that the purpose of the study is to explore the effects in ideation performance after introducing a semantic word-based ideation method to solve transactional design problems, we developed a correlation analysis between the quantitative ideation metrics to account for possible interactions among the metrics (Chan, et al., 2011 a & b).

An initial aggregated relational model that summarizes the study results (including data sets from both the control and experimental groups) is semantically described below and presented in Fig. 11:

- The higher the novelty level, the lower the level of fixation.
- High quantity of ideation may lead to lower levels of fixation, i.e., less variants of a base idea or repetition of the same idea.
- Semantic solution transfer appears to be positively correlated to quantity for its efficiency results, where the ideas generated are directly associated with semantic exploration.
- Quality appears to be positively correlated with novelty and indirectly with quantity. Benchmarking results showed that instances of the banking market’s innovative solutions for the problem could have been achieved using novel solutions derived from the method.
Analyzing the control and experimental groups separately, and isolating the effects of qualitative metrics, the initial model expands to the model shown in Fig. 12. Notice in this figure that the general model is transferable for both groups with the exception of the correlation between novelty (third approach) and fixation for the control group NT. This difference in results is due to the significant and large novelty difference between both the groups.

Fig. 12 contains bar charts of the metrics presented in Section 5 for both the control and experimental groups and phases. Based on these results, there is no significant difference for quantity between both
groups in each phase. On the other hand, for novelty and fixation in phase II, there is a significant difference between the control and experimental groups.

4.8. Survey and debrief

After completing each phase, participants of both experimental and control group were given a brief survey to gather demographic information, feedback about the ideation activity, the selected problem and their performance.

The age of the study’s participants ranged from 28 to 49 with an average of 36 years. Participants rated the questions using a Likert Scale, where participants selected their level of agreement to a statement by indicating a position along a segment of five (5) boxes labeled -2, -1, 0, 1, and 2 to identify participant’s average agreement with the statements presented in the survey. A value of zero denotes indifference (or neutrality) and -2 and 2 represent total disagreement and total agreement, respectively. Results for phase I and II results are presented in Table 13.

Table 13. Survey results

<table>
<thead>
<tr>
<th>Phase I</th>
<th>Phase II</th>
</tr>
</thead>
<tbody>
<tr>
<td>Boring - Fun</td>
<td>0.778</td>
</tr>
<tr>
<td>De-motivating - Motivating</td>
<td>0.385</td>
</tr>
<tr>
<td>Frustrating - Inspiring</td>
<td>0.462</td>
</tr>
<tr>
<td>Difficult - Easy</td>
<td>0.821</td>
</tr>
<tr>
<td>Uninteresting - Interesting</td>
<td>0.037</td>
</tr>
<tr>
<td>Difficult - Easy</td>
<td>0.500</td>
</tr>
<tr>
<td>Frustrating - Inspiring</td>
<td>0.640</td>
</tr>
<tr>
<td>Minimal effort - Worked hard</td>
<td>1.071</td>
</tr>
<tr>
<td>Unfocused - Focused</td>
<td>0.930</td>
</tr>
<tr>
<td>Unsatisfactory - Successful</td>
<td>-0.033</td>
</tr>
<tr>
<td>Time than ideas - Ideas than time</td>
<td>1.855</td>
</tr>
<tr>
<td>I had more</td>
<td>1.385</td>
</tr>
</tbody>
</table>

For the question “The ideation activity was…”, the mean scores for both phases and groups were equal or higher than zero, therefore, closer to “Fun, Motivating, Inspiring, Easy” attributes. A similar trend is
observed for “The problem was…”, which means that participants’ perceptions were correlated with “Interesting, Easy and Inspiring” attributes. For the “performance on the activity”, participants’ averages were again greater than zero which meant that they considered themselves as being “Worked hard, Focused and Successful.” Lastly, for the last question of the survey “I had more… Time than ideas - Ideas than time,” average scores for both phases were close to zero, expressing indifference, i.e. the proportion of time to generate ideas was sufficient. For the WT group in phase I the perception was closer to having more time than ideas, and in phase II shifted to zero which may be interpreted as that the initially perceived extra time in phase I was being used in phase II to apply the technique.

5. Discussion

The experimental study produced a number of intriguing and thought-provoking contributions. The bin distribution shown in Fig. 6 is not what is usually expected compared to ideation results for engineering knowledge domain design problems. Many previous studies show distinct shifts in frequencies indicating transitions from more novel ideas to those which are based on more standard or historical approaches (Chan, et al., 2011 a & b; Fu, et al., 2012; 2013; Linsey, Markman, & Wood, 2008; Weaver, et al., 2009). Fig. 6 of this study shows a gradual exponential rise of idea frequency, where nearly 50% of the unique ideas (bins) have relatively low frequency. No clear shifts or breaks are observed in the frequency of ideas generated. Further studies are needed to explore more transactional problems to determine if this pattern repeats, especially with participants that are knowledge-domain experts, as in this study. However, the data pattern suggests that the experts spanned a variety of industries and experience sets, adding solutions from these varied perspectives. The pattern also suggests that many different foundational approaches exist to solving transactional problems such as personnel-based solutions, customer or user based solutions, environmental solutions, technology solutions, and permeations of these approaches.

Semantic solution transfer provides insights about the participants’ stimuli received as a consequence of introducing the DbA technique and the associated analogy retrieval process. For the listed words and solutions, the efficiency indicator, i.e. the ratio of solutions created with the analogic semantic search compared to the total quantity of ideas generated is greater than 80%. This indicates significant potential for the combined WordTree and Divergent Tree method for assisting in the generation of ideas.

Focusing on the participants’ analogical retrieval process, it is interesting to find that 91% of the words listed as idea triggers did not originate from the provided KPDs, but from the method, i.e. from semantic exploration performed by participants. There are also a number of bin-ideas stated only once, which indicates the methods enablement of novelty, at least when it comes to unusual ideas (low frequency).
The WordTree and Divergent Tree method enable generation and decomposition of multiple linguistic representations for a design problem from different perspectives; thus it can be applied to overcome the implicit intangible nature of the solutions for transactional design problems that are better expressed verbally rather than through physical models. Such linguistic approaches promotes metaphorical reasoning to seek means to understand and re-represent the design problem at hand, they also address designers’ lack of access to transactional solutions databases (unlike physical products) and a direct means to identify potential analogies sources for their current problem to another solution domain. Not having means to relate to other solutions makes it necessary to add divergence in order to exploit new representations of the design problem, that previous literature have shown to be effective for creating new relations and associations.

From the analogical memory retrieval perspective, since the WordTree and Divergence Tree method provide both divergent and convergent capabilities, new representations of the problem and key components identified from these methods may activate wide and diverse cues for semantic memory retrieval, increasing the possibility of finding solutions and exploring analogous solution domains for transactional problems. It is also possible that having an unsolved problem i.e. open goal and being provided with analogical hints that came out of the divergent WordTree method enabled the participants to have better or new ways to solve the problem as it has been shown in other previous open goal studies (Moss, Cagan, & Kotovsky, 2007; Moss, Kotovsky, & Cagan, 2007; Tseng, Moss, Cagan, & Kotovsky, 2008a).

During phase II, participants were presented with customer needs and functional requirements that may have assisted them to frame or re-represent the design problem (metaphor reasoning), and it is also found that participants while exploring KPDs to re-represent the transactional design problem were simultaneously developing forms of solutions, which they explained later as means to not forget the targeted solution triggered by the selected words. Participants also commented that in some cases some of the displayed words helped them to retrieve another word or situation not directly related with the generated or displayed words that helped them to develop alternative solutions. For example the word “reinforce” led a participant to think of “educate/teach,” and develop solutions based on financial education models. Comments such as these support the quantitative results and confirm that semantic memory retrieval and/or divergence enables exploring an analogous domain to develop innovative transactional solutions.

The metric for quantity produced intriguing results. It is found that when both the control and experimental groups generated ideas without using any particular tool (phase I), they performed at an equivalent level. The NT control group was able to produce the same quantity of ideas in both experimental phases, whereas for the WT experimental group, phase II resulted in less ideas compared with phase I. These
results indicate that for total quantity of ideas, there is an expected cognitive load during phase II for the WT group. But what is more intriguing is that after removing repeated ideas, the quantity performance of both the control and experimental groups was statistically equivalent. Thus, what is originally perceived as cognitive load for the WT experimental group has no actual consequence on generating a final quantity of ideas. Instead, with experience and proficiency (the experimental group received only 15 minutes of training with the DbA method), it is very possible that the method could lead to an increase in quantity compared with a group without the use of such methods. Of course, this conjecture would need to be validated through further experimentation.

Considering the quantity of repeated ideas, Table 9 shows that the total number of repeated ideas in phase I is not statistically significantly different for both of the study’s groups (NT=45, WT=47). However, and quite distinctly, the total quantity of repeated ideas for the NT control group is statistically significantly different (172) and more than three times the amount of the other phases. These results clearly demonstrate the high-level of fixation level encountered by the NT control group.

When analyzing repeated ideas by source (within and between phases), the number of repeated ideas within phases is very homogeneous, with the exception of the phase II results of the NT control group. This result implies that the introduced method has a positive impact on reducing the number of slight variants or repeated ideas. The number of repeated ideas between phases for the NT control group is almost five times greater than the WT experimental group, confirming again the beneficial effects of introducing the DbA technique.

The study results indicate that a large quantity of the participants’ time is spent creating variants, or very similar ideas; however, this effect appears to be greatly reduced when participants perform directed analogical semantic search and problem re-representation. This statement is validated with the semantic transfer results and efficiency of the method, i.e., participants are developing their ideas from the explored words.

Considering the experimental phases, participants from the NT control group in phase II quickly encountered problems generating new ideas (approximately at the midpoint of the allotted time). When asked about this situation once the experiment concluded, many participants stated that it was difficult to generate solutions when they were not able to find different perspectives (a reference, an analogous solution). Some participants also expressed the desire to talk to customers to find different ways to approach the design problem. These anecdotal results support the quantitative findings. There is a tendency for the non-assisted scenario (control) participants to seek small variations to previously generated ideas when alternative categories of ideas are not apparent.
The novelty results indicate, for the first and second evaluation approaches, that approximately 70% of the unique ideas are developed by the WT experimental group compared to a baseline design space from phase I. For the third approach to evaluate novelty, i.e., calculation of uniquely generated ideas in phase II relative to the total ideas generated by each participant, there is a statistical significant difference in the order of magnitude between the NT control and WT experimental group novelty percentages (five times). This result is remarkable, considering that participants have only been trained for 15 minutes, which may imply that as participants become more proficient with the technique and explore divergent words in terms of synonyms, antonyms, hypernyms and adverbs, an even wider range and higher quantity of novel results might be developed from analogies. Further studies would be needed to explore and validate this conjecture.

When mapping the backgrounds of the participants with novel ideas generated during phase II of the experiment, no trend is identifiable in the field, educational background, or industry domain. These results provide two interesting insights: first, that for developing novel solutions, general background is not relevant, but instead the expertise or proximity to the type of problem, and second, that the ideation phases are relatively background independent (except with general transactional problems), i.e., that with or without a directed ideation method, participants can be either stimulated or not, respectively. In the case of using the method, participants are clearly more stimulated or freed to develop novel solutions compared to the non-technique scenario.

The qualitative analysis of generated solution quality shows that solutions generated by participants with knowledge-domain expertise include indicators of being feasible for implementation by banking corporations. This result shows the potential to develop a high quality of results when generating ideas with techniques such as the introduced Design-by-Analogy approach, especially when in the hands of experienced professionals for transactional systems.

A transaction is something that is not physical; it is more socioeconomic related and has a high level of abstraction. The design-by-analogy method introduced as part of this experimental study shows great potential to work in this abstract space, especially with respect to mitigating fixation and improving the novelty of solutions. The experimental results demonstrate that design-by-analogy methods developed in engineering artifact fields are appropriate and have great potentially for success in generating ideas in transactional fields.

The survey and debrief results indicate that the use of DbA methods shows clear differences for an assisted versus non-assisted scenario. Introduction of semantic-based analogical approaches assist in maintaining interest and performance level, and, at a personal level, they also help to provide an improved motivational state for the ideation activity. Frustration and idea blocking appear to be greatly reduced.
6. Implications and future work

This experiment explores the important area of innovation processes and idea generation for transactional problems. Previous research shows the effectiveness and robustness of idea generation methods in engineering artifact fields (manufacture, products and tangible objects), but there exists significant opportunities for the adaptation of these techniques to transactional problems. The experimental study reported here demonstrates that problem representation clearly matters in the generation of analogies and analogous domains. The use of analogical semantic exploration allowed participants to re-represent a transactional design problem to explore different domains through new and different sets of problem statements and solution spaces. It is likewise demonstrated that problem fixation may be mitigated significantly through analogue semantic exploration. The possibilities appear to be extensive and exciting.

An interesting avenue to be explored beyond this study is connection of sketching and diagramming as part of ideation processes for transactional problems. It was noticed that participants not only expressed their ideas through sentences for re-representing a problem, but also developed solutions using flow diagrams and abstract sketches. There are unexplored possibilities for sketching and related ideation techniques for transactional design problems.

The present experiment focused on idea generation to address transactional problems; however, exploration of other ideation techniques and problem types will provide a better understanding of the process behind idea generation across any domain physical and transactional.

7. Acknowledgements

The authors wish to acknowledge support provided by the ITESM-BMGI Lean Six Sigma Program (http://6sigma.mty.itesm.mx), without whose participants, this study would not have been possible. We wish to acknowledge the SUTD-MIT International Design Centre (IDC) (http://www.sutd.edu.sg/idc.aspx) for its support and encouragement to take the research to new levels. Finally, to the Postdoctoral Graduate Fellows program between MIT and SUTD for creating a nurturing environment where collaboration from multiple backgrounds and expertise is possible and encouraged. Any opinions, findings, or recommendations are those of the authors and do not necessarily reflect the views of the sponsors.

8. References


FIGURE CAPTIONS

Figure 1. Illustration of the continuum of transactional to physical systems
Figure 2. Example semantic network food storage
Figure 3. Transactional problem recurrence distribution
Figure 4. Experimental execution diagram
Figure 5. Selected examples of participants that generated large amount of concept solutions by experimental group
Figure 6. Bin-ideas frequency distribution
Figure 7. Chronological bin-idea cumulate contribution by group of participants
Figure 8. Design space by phase
Figure 9. Key problem descriptors displayed by Thinkmap’s Visualthesaurus©
Figure 10. Qualitative analysis of innovative transactional solutions for problem used (Citigroup Inc., 2006; Citigroup Inc., 2011; Westpac Banking Corporation, 2012)
Figure 11. Aggregated metrics interrelationship model
Figure 12. Disaggregated model for metrics interrelationship

TABLE CAPTIONS

Table 1. Differences between services and products (modified from Vermeulen, 2001)
Table 2. Participants’ demographics and background
Table 3. Selected KPDs for Transactional Design Problem
Table 4. Experimental group’s gender distribution
Table 5. Power and sample size analysis
Table 6. Normality test
Table 7. Semantic word transfer
Table 8. Quantity of generated ideas
Table 9. Repeated ideas by group, source and phase
Table 10. Fixation (%) by Phases of both Groups
Table 11. Novelty Metric’s second approach and associated number of ideas
Table 12. Average and standard deviation for novelty
Table 13. Survey results