Fifth International Conference on Design Creativity (ICDC2018)
Bath, UK, January 31st – February 2nd 2018

TESTING DESIGN STIMULI FOR DESIGN-BY-ANALOGY ON A LARGE SET OF DESIGNERS

L. Jia[1]; N. Becattini[1]; G. Cascini[1]; R. Tan[2]

[1]Dept. of Mech. Eng., Politecnico di Milano, Milan, Italy
[2]Dept. of Mech. Eng, Hebei University of Technology, Tianjin, China

Abstract: This paper presents evidence supporting the hypothesis that, for designers not specifically trained in designing-by-analogy, the sources of inspiration that share the same (sub-functions) and context of the target system lead to ideas having higher novelty and quality. The exploration of the design space gets positively affected as well. These evidence emerge after the statistical analysis of the results of an experiment that involved 84 graduate students in Mechanical Engineering, with typical competencies on engineering design, but without any specific skill on analogy-based idea generation.

Keywords: *creative stimuli, idea generation, design-by-analogy, experiment*

1. Motivation and background

Analogies, in design, play a relevant role in the activation of creative thinking as they can help to figure out meaning in front of a situation as well as to suggest potential solution strategies in problematic ones. Analogy-based design (or Design-by-analogy) is a practice in which analogy is applied in the design process for helping the designers to get inspired to solve the target problems (Christensen and Schunn 2007; Fu et al. 2013; Goel 1997). Engineering design literature explores the role of analogies by distinguishing different factors that might affect designers' performance. Chan et al. (2011), for instance, explored how different modalities of analogical stimuli influence the design outcomes. Behavioural differences between experts and novices in the way they use analogies have been recently discussed in Chai et al. (2015) and Ozkan and Dogan (2013). Distance (from a target domain or situation) for analogical stimuli has been considered a key factor to effectiveness in design. Dunbar (1997) divides analogies with a biological perspective into: within-organism (systems are from the same organism), other-organism (systems are from different organisms) and distant analogies (systems are from a non-biological domain). Chai et al. (2015), similarly, classify near analogy as the analogical association between the same product category, medium analogy as association between different product categories and distant analogy as the association between the product and a non-product category. Chan et al. (2011) argue that the near-field analogy appears when the target and analogical source are from the same or very similar problem domain, while far-field analogy appears when they are from different problem domains. Considering semantic analysis and structure of patents, Fu et al. (2013) distinguish the near and far analogy based on the contextual similarity of patents. The effect of analogical distance on design outcomes is also controversial. Some empirical studies revealed that higher the distance from source domain to target domain was, greater the novelty and quality of ideation would be (Chiu and Shu 2012;

Gentner and Markman 1997). However, other studies questioned the advantages of distant analogy, because it could be cognitively challenging to retrieve distant information from memory. It was also negatively associated with the probability of reducing development time (Casakin and Goldschmidt 1999; Weisberg 2009). A further issue that might explain the contradictory outcomes of apparently similar studies in this field is the impact of specific training on the capability of designers, either novices or experts: the recognition of analogies and the generation of ideas through the transposition of some elements of the analogous concept into a new idea is not necessarily intuitive and can be improved by practice.

Within the above mentioned research thread, this research investigates how different sources of inspiration, working as creative stimuli for the generation of novel and feasible ideas during a design task affect designers without specific skills on analogy-based idea generation.

Through the combination of extant perspectives on analogical stimuli, the authors propose to organize them into sources of near-field, middle-field and far-field analogies. Wherein, sources of near analogies are regarded to be stimuli that share the same (sub)function with the target and are applied in the same industry context. For middle distance analogies the sources still share the same (sub-)function with the target, but applied in different industry contexts. Those whose (sub-)functions are not directly relevant with the target, but could be potentially useful, whether the industry application contexts are similar or not, go under the category of far sources.

The authors believe that it is now necessary to extend such studies to a wider set of designers, also with different background and skills, in order to provide new evidence and more definitive conclusions about what still remains controversial. 84 novice mechanical designers, subdivided in small design teams, participated in the experiment here presented: they were assigned a task of conceiving new ideas for a customer product and received stimuli characterized by different analogical distance (near, medium, far) to the target technical field. The discussion focuses on both the design performance obtained with different stimuli and the limitations emerged in the application of the experimental protocol.

2. Experimental protocol

The experiment was conducted at Hebei University of Technology in China, with 84 postgraduate students (17F-67M – MS in mechanical engineering). Randomly, 21 participants were assigned to the "Control Group". The same amount of testers also composed the groups exposed to near-field analogies ("Near field"), middle-field analogies ("Middle Field") and far-field sources of analogy ("Far Field"). Each group counted 7 design teams 3 members each to recreate a typical collaborative design session. The experiment has 5 stages, 10 minutes each. The first 10 minute-stage comes with no stimuli at all, to verify if groups are homogenous (effectiveness of the randomization in group composition). Then, two different stimuli are introduced every 10 minutes during stages 2 to 5, to compare the effects of different analogical distances.

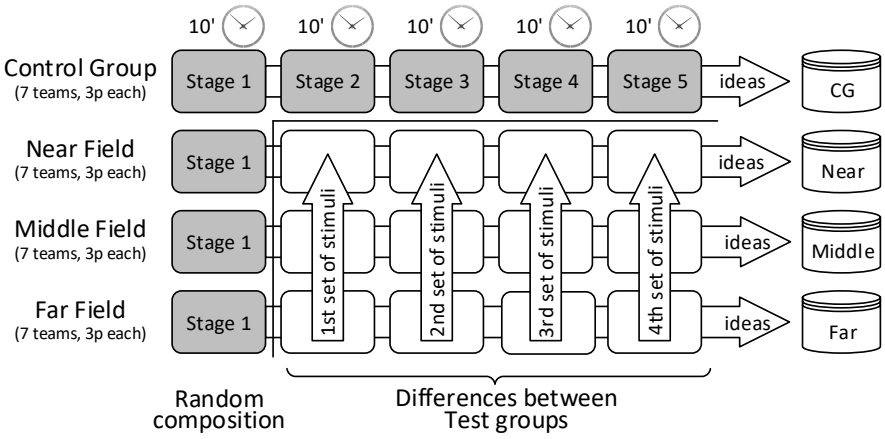

Figure1. Experiment setup

With a short presentation before the test, the participants were asked to propose ideas for the next generation of vacuum cleaning robots. The presentation discussed some of the most common problems of these device as a preliminary design brief:

1. The quality and efficiency of cleaning are low, especially for corners and edges of the room.
2. The collection device is hard to clean.
3. The robot easily gets stuck and the wheels are also easily twined by string, cable, et al.
4. Other problems, such as high noise, insufficient energy, et al.

During the presentation of the case study, students were encouraged to generate also ideas beyond the scope of the proposed problems.

The experiment took place in a large classroom to allow the communication within every team and prevent any between-team interference. During the experiment, surfing the internet was forbidden and participants had to write down their ideas as text and sketches on the ideation template handed out at the beginning of the experiment. The participants also had to specify the problem they focused on, the stimulus that inspired them and the team name.

For what concerns the treatments for the three test groups receiving stimuli, they have been generated according to the classification of analogical distance shown in Table 1. The authors identified eight exemplary devices or products for each test group, each supposedly suitable as stimulus to address the problems of cleaning, moving, saving energy and decreasing noise. With respect to the suggestions from Goldschmidt and Sever (2011) and Paivio (1990), every stimulus is provided as a combination of picture and text. Table 2 shows an example. The whole set of stimuli used along the experiment is available at this URL: https://www.dropbox.com/s/ufln2qqb29qmndz/Appendix_creative_stimuli.pdf?dl=0.

Table 1. Summary of selection guidance for analogical stimuli

Analogy kind	The stimulus share with the target system (or one of its subsystem)…	
	…same function	…same context
Near-field	Yes	Yes
Middle-field	Yes	No
Far-field	No	No

Table 2. Stimuli from different analogical distance

Type	Example	Content of the stimulus
(-Near OR -Middle OR -Far) field stimuli		Description of the function, of the working principle and/or the context of use, together with a web-retrieved picture of the product used as source of analogy

The ideation performance of each team have been measured according to four metrics: Quantity, Variety, Novelty and Quality (Shah et al, 2003; Sarkar and Chakrabarti, 2011). Figure 2 presents the scores for each metrics. The assessment of ideas according to the metrics have been carried out by two raters, whose worst degree of agreement is 0,92 (Kendall's W, $p<0,01$). Results got clustered for each treatment group. The 7 teams per group contribute to generate a descriptive statistic for the population administered with the same treatment. Such data is then used in order to evaluate:

- The uniformity among the different groups before any treatment (results of Stage 1). This check on the goodness of the randomization for group composition is a necessary pre-condition for the meaningfulness of the second part of the experiment.
- The differences between the different kinds of stimuli (results Stage 2-5).

To evaluate the significance of different ideation performance between groups, Kruskal-Wallis is chosen against ANOVA because of the unknown nature of the distribution of the population, thus, of its variance. Then, differences due to the effect of stimuli having different analogical distances are also explored by means of one-to-one comparison between different groups. These differences are every time measured through meaningful statistical estimators to the effectiveness of ideation performance.

Quantity per each team counts the overall amount of generated ideas, while the equation of Figure 2 yields the value for Variety. Both are computed considering the whole set of ideas generated by teams.

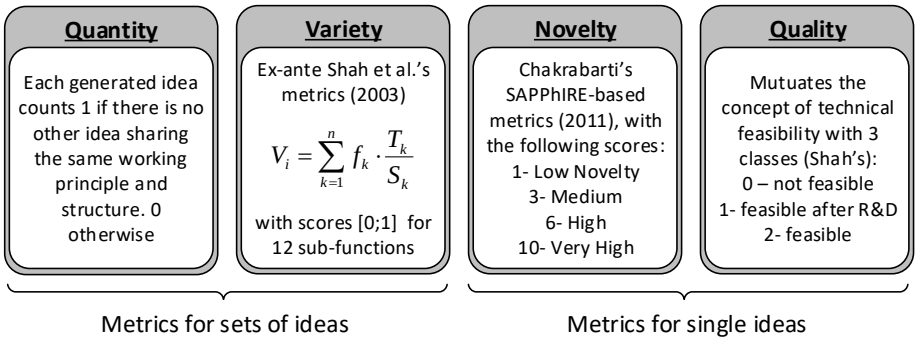

Figure 2. Metrics for the evaluation of ideation performance

For what concerns Novelty and Quality, whose assessment regards single ideas, the following bullet list summarizes the statistical estimators used for the assessment of differences.

- Average values: each idea gets its score. The sum of each score is then divided by the overall amount of generated ideas.

$$\text{Average}_{\text{ Novelty (Quality)}} = \sum_{1}^{\text{Quantity}} \frac{\text{Novelty (Quality) score}_{\text{i}}}{\text{Quantity}} \quad (1)$$

- Percentages: each idea gets its score. The percentage is calculated as the ratio between the ideas whose score overcomes a given threshold value and the overall amount of generated ideas

$$\%_{\text{ Novelty (Quality)}} = \frac{\#\text{ of ideas with Novelty (Quality) score} > threshold}{\text{Quantity}} \quad (2)$$

- Maximum values: each idea gets its score. The maximum value for each team is computed by adding the scores of the N top-rated ideas. The choice of N depends on the minimum Quantity recorded among the 28 different teams.

$$\text{Max}_{\text{ Novelty (Quality)}} = \sum_{1}^{\text{N}} \text{Novelty (Quality) score}_{\text{i}} \quad (3)$$

$$\text{Novelty (Quality) score}_{\text{ i}} \in \text{ TopRated} = \{\text{ideas with Top} - \text{Rated scores}\}$$

With reference to the evaluation of maximum values, it is worth mentioning that the selected index allows for a uniform evaluation of ideas among the groups. In fact, on the one hand this allows removing a flattening effect among teams and groups if the set collects just the idea getting the highest score in the whole set is considered. On the other hand, limiting such assessment to the N top-rated scores removes the bias due to the different values of Quantity among teams and groups. This prevents that more groups characterized by higher Quantity will benefit from having a larger number of low-scored addends, if compared to groups that were not equally fluent but that performed better on Novelty and/or Quality. Furthermore, one-to-one homogeneity comparisons allow to estimate the statistical significance of differences between groups receiving different treatments (therefore, these results will be only presented in terms of p-values). Such comparisons are carried out through Mann-Whitney tests, considering the results come from ordinal scales, for which it is not ensured the normal distribution of the population. Both the mentioned non-parametric statistical tests also appear to be more reliable for small samples (Rice, 2006). This analysis, however, is not sufficient to clarify which of the groups performed better under different creative stimulations. The descriptive statistics about the idea generation process, then, are considered with their actual values in order to draw conclusions and clarify the meaning behind the analysis of significance for the comparison.

3. Results of the experiment and related analysis

This section presents the analysis of the data gathered along the previously described experiment. Still with reference to Figure 1, the first subsection deals with the results of the first stage of the experiment.

The second one, then, summarizes the differences among the different kinds of stimuli according to the results obtained after idea generation (second phase of the experiment, stages 2 to 5).

3.1. Experiment Stage 1 - Analysis of homogeneity among groups

As stated in Section 3, the group composition was determined by chance, thus following a randomization process. The results of the Kruskall-Wallis test show that there are no significant effects, due to group composition, on the quantity of generated ideas ($p=0.904>0.05$). The same can be said for what concerns the novelty of the generated ideas, that in this case has been computed as the average score of novelty for each of the groups ($p=0.228>0.05$). As well, the differences in average quality of ideas among groups ($p=0.692>0.05$) and the related variety of what they ideated ($p=0.838>0.05$) show that groups have been properly created through randomization. It can be concluded then that emerging differences between groups in the second phase, if any, depend on the effect of analogical stimuli the groups are exposed to.

3.2. Experiment Stages 2-5: Analysis of the effects of near-, medium- , and far-field analogies

The results of the Kruskal-Wallis test shows that differences among groups are significant for the average score of novelty ($p=0.004<0.01$), quality ($p=0.000<0.01$) and for the variety of ideas ($p=0.016<0.05$). This implies that the analogical stimuli have a statistically evident influence on the design outcome. On the contrary, the results show that the differences among quantity of ideas among the groups does not (statistically) depend on the stimuli ($p=0.116>0.05$).

For what concerns the Novelty of the ideas, the results of the Mann-Whitney test support that the outcomes of the idea generation are significantly different between the Near-field group and all the others (max p=0,013). The boxplots of Figure 3 (first row) on the average values (left) and on the ratio (center) clearly witnesses that the distribution of the results in the Near-group is way less dispersed (except one outlier) than the others. Moreover, the same distribution shows a significantly higher median. The data are more dispersed for the maximum value (right), but in that case the median is higher than the extreme values of the other treatments. Thus, it clearly emerges that the stimulation through near-field analogies is the most effective in supporting the generation of ideas of higher novelty for this sample of testers. The ratios for treated subjects is generally higher than for the control group, confirming the positive impact of creative stimulation on idea generation effectiveness.

In terms of quality of ideas, the performance of the group dealing with Near-field analogies are statistically different from both the control group (max p=0,006) and the Far-field group (max p= 0,002). The Middle- and the Far-group are statistically significant as well (max p=0,016). The Boxplots of Figure 3 (central row) highlight that there is a negative relationship between the quality of ideas and the distance of analogies used as sources of inspiration. The highest percentage of ideas of high quality is achieved through near-field analogies, while the smallest with far-field analogies. The same is true for both the average and maximum values. This trend is so marked that is also clear that far field analogies play a counteracting role in creative stimulation also if they are compared to no treatment at all, while medium ones might have a light, but positive impact.

To evaluate the influence of analogical distance on variety, as a measure of the exploration of the design space, the equation of Figure 2 provides the values for each of the 7 teams in each group. The Mann-Whitney test displays that, similarly to Quality, the only significant differences of Variety are between the Near group, respectively against the Control Group and the Far-Group. The distributions of the treated groups (Figure 3, third row) also confirm the negative relationship between the variety of outcomes of the idea generation process and the analogical distance. Shorter the distance, higher the variety and vice versa. The distribution of the Near-Group is the less dispersed and it also presents a median which is bigger than the largest majority of the data points of other groups. Differently from the case of quality, here the performance of the Control Group and the Far-Group do not appear to be markedly different. This means that whatever is the analogical distance of creative stimuli, their effect is positive or, in the worst case, null. In other words, the results of the experiments show that the examples provided as sources of inspiration for design-by-analogy do not present any specific effect of design fixation on the tested sample.

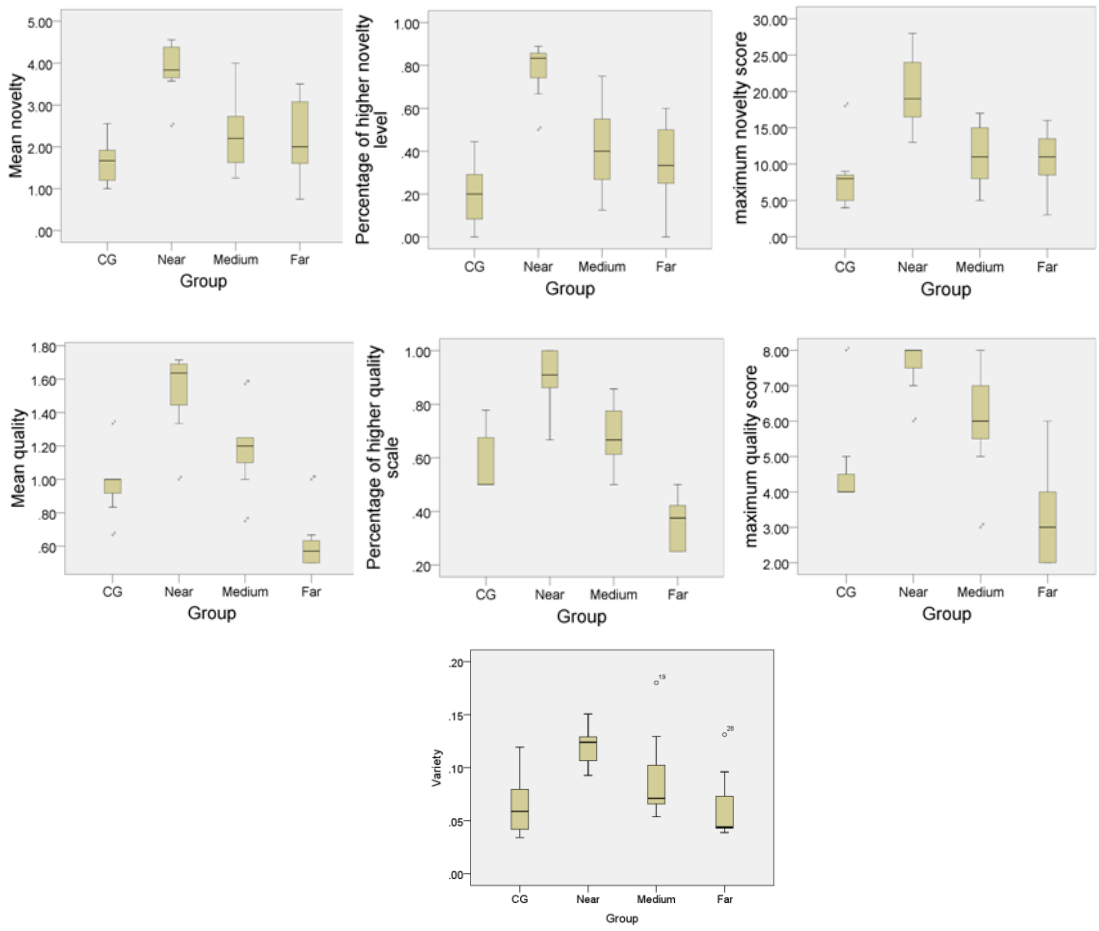

Figure 3. Boxplots. First row: Novelty. Second row: Quality. Third row: Variety

4. Discussion

As depicted by the results presented in section 3.2, the design outcomes from all the test groups with analogical examples were not significantly different from each other for what concerns the quantity of generated ideas (p=0.116). The setup of the experiment in stages can at least partially explain the homogeneity of results for quantity. After the experiment, the participants were interviewed. Some of them said that they faced the experiment as an examination (even if participants received explicit instructions saying that it was not an exam session) and handed in the "answer" every 10 minutes. Overall, even though the quantity of ideas is approximately stable in every stage and the novelty of ideas in Control group and Near-field group keeps increasing slightly, the worsening of average quality, as the experiment proceeds, could suggest that all the groups get closer to saturation. This effect might depend on the depletion of ideas or on the tiredness of participants. Further investigation are required to clarify this issue. For novelty, all the chosen criteria (average and maximum novelty, percentage of novelty of ideas) display that the ideas inspired by near-field analogies are significantly more novel, on average, than the ones generated with different treatments, including the Control Group. In addition, Figure 3 implies that, compared to the control condition without any stimuli, both medium-field and far-field analogy are tending towards higher novelty design output, even if the effect is not fully statistically significant. With regard to quality, although there is slight discrepancy among average quality, maximum quality and the percentage of ideas having higher quality, the conclusion is in accordance with Chan, Dow, and Schunn (2015): the quality of ideas inspired by near-field analogies is better than

the ones inspired by the far-field analogy and control conditions. Figure 3 shows that stimuli of closer analogical distance generally correspond to a better quality of ideas. Unlike for the influence on novelty, the far-field analogy negatively affects the percentage of ideas having higher quality. This suggests that distant analogies may require designers to retrieve, map and transfer a hardly manageable amount of information for cognition and this appears as more difficult without a dedicated training. About variety, the results only allow inferencing that near-field analogies help designers generate ideas in a broader design space than what control condition, medium-field and far-field analogy do. This could also reinforce the former statement about the effectiveness of near-field analogy, as they can also inspire better ideas with higher novelty and quality.
In summary, this experiment provides evidence that near-field analogies play the most beneficial role in the ideation at least when proposed to designers not specifically trained on this practice.

5. Conclusions

This paper explores the effects of the distance between analogical stimuli and the design target (using near-field, medium-field and far-field analogies) on the idea generation process. The experiment here presented aims at providing further evidence to clarify the controversial results emerged in the design literature. The estimation of such effects is carried out through statistical tests aiming at highlighting the significance of differences behind treatments using diverse sources of creative stimulation. Analogical stimuli have been reclassified into three categories based on the combination of their function/purpose and context of application. Near-field analogical stimuli share both a sub-function and the context with the target system. Medium ones share with the target just the function, not the context. Far ones, independently from the context, entail a non-shared sub-function. Relevant metrics to evaluate the outcomes of the ideation process are adopted from literature, together with a tailored approach developed for comparing the results obtained with design stimuli characterized by different analogical distance. For the statistical significance of the experimental results, the experiment took place on a large scale (compared to similar design investigations). In fact, it involved 84 novice designers with a background in mechanical engineering but no specific training on analogy-based idea generation, subdivided in 28 design teams organized in 4 treatment groups. The teams were asked to generate ideas for an innovative robot for vacuum cleaning. Each group was exposed to a different treatment: the control group received no treatment, while the others were exposed to design stimuli with, respectively, near-, medium.- and far-analogies. The main findings of this experiment show that:

- Novelty: Near-field analogies are the most effective to generate novel ideas. The effect of medium- and far- field analogies is comparable to the outcomes of the control group.
- Quality/Feasibility: there is a negative relationship with the distance of analogies used as sources of inspiration. Far field analogies work against quality as the results are worse, compared with the performance of the control group.
- Variety: there is a negative relationship between the outcomes of the idea generation process and the analogical distance. Whatever the analogical distance of creative stimuli is, their effect is positive or null, as the performance of far-field and no analogies are similar.
- There is no statistical significance that the creative stimulation by analogy produces an effect on the quantity of generated ideas.

With reference to the classification of analogies by distance presented above, the better outcomes obtained with near-field analogies suggest both the shared (sub-)function and the context help generating more promising ideas. Moreover, the trends among near, medium and far field analogies also suggest that sources of inspiration which do not share the same context with the target are less effective, especially if the source of inspiration do not share the same (sub-)function either. This effect can depend on the difficulty for people not specifically trained on design-by-analogy to make proper cognitive associations when more features needs to be matched at the same time, in order to solve a constrained problem (in terms of functions to be carried out and operational context).
Some limitations emerged with the analysis of the experimental results. Participants would require a more effective introduction to the design task that spans over a longer time, to reduce the density of

information they have to process. Such introduction, however, should not narrow the exploration of the design space by constraining the design space (as the set of problems of the current robot for vacuum cleaning presented before the experiments might have exerted this effect).
The results of the experiment show that design novices, differently from the conclusions in Bonnardel and Marmèche (2004), tend to benefit from inner-domain sources as the near-field analogies are positively correlated to good ideation performances. On the other hand, this seems contradictory with the findings in Ozkan and Dogan (2013), as, in that case, distant sources of inspiration appeared to be more effective on novices. This suggests investing further efforts to properly characterize the profile of the experimental sample as yet unconsidered factors might be playing a role in exploring creative stimuli effectiveness, including cultural issues and the target application field. From this perspective, the authors plan to extend the experiment to a broader sample of designers with a controlled profile.

6. References

Bonnardel, N., and E. Marmèche. 2004. "Evocation processes by novice and expert designers: Towards stimulating analogical thinking." Creativity and Innovation Management 13(3): 176-186. doi: 10.1111/j.0963-1690.2004.00307.x.

Casakin, H., and G. Goldschmidt. 1999. "Expertise and the use of visual analogy: Implications for design education." Design Studies 20(2): 153-175. doi: 10.1016/S0142-694X(98)00032-5.

Chai, C., F. Cen, W. Ruan, C. Yang, and H. Li. 2015. "Behavioral analysis of analogical reasoning in design: Differences among designers with different expertise levels." Design Studies 36: 3-30. doi: 10.1016/j.destud.2014.07.001.

Chan, J., S. P. Dow, and C. D. Schunn. 2015. "Do the best design ideas (really) come from conceptually distant sources of inspiration?" Design Studies 36: 31-58. doi: 10.1016/j.destud.2014.08.001.

Chan, J., K. Fu, C. Schunn, J. Cagan, K. Wood, and K. Kotovsky. 2011. "On the benefits and pitfalls of analogies for innovative design: Ideation performance based on analogical distance, commonness, and modality of examples." Journal of Mechanical Design 133(8): 081004. doi: 10.1115/1.4004396.

Chiu, I., and L. H. Shu. 2012, "Investigating effects of oppositely related semantic stimuli on design concept creativity." Journal of Engineering Design 23(4): 271-296. doi: 10.1080/09544828.2011.603298.

Christensen, B. T., and C. D. Schunn. 2007, "The relationship of analogical distance to analogical function and preinventive structure: The case of engineering design." Memory & cognition 35(1): 29-38. doi: 10.3758/BF03195939.

Dunbar, K. 1997. "How scientists think: On-line creativity and conceptual change in science." In Conceptual structures and processes: Emergence, discovery and change, edited by T. B. Ward, S. M. Smith, and S. Vaid, 461-493. Washington, D.C: American Psychological Association press.

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

Gentner, D., and A. B. Markman. 1997. "Structure mapping in analogy and similarity." American Psychologist 52(1): 45-56. doi: 10.1037/0003-066X.52.1.45.

Goel, A. K. 1997, "Design, analogy, and creativity," IEEE expert 12(3): 62-70. doi: 10.1109/64.590078.

Goldschmidt G., and A. L. Sever. 2011. "Inspiring design ideas with texts." Design Studies 32(2): 139-155. doi: 10.1016/j.destud.2010.09.006.

Ozkan, O., and F. Dogan. 2013. "Cognitive strategies of analogical reasoning in design: Differences between expert and novice designers." Design Studies 34(2): 161-192. doi: 10.1016/j.destud.2012.11.006.

Paivio, A. 1990. Mental representations: A dual coding approach. New York: Oxford University Press.

Rice, J. A. 2006. Mathematical statistics and data analysis. Belmont: Thomson Higher Education.

Sarkar, P., and A. Chakrabarti. 2011. "Assessing design creativity." Design Studies 32(4): 348-383. doi:10.1016/j.destud.2011.01.002.

Shah, J. J., S. M. Smith, and N. Vargas-Hernandez. 2003. "Metrics for measuring ideation effectiveness." Design Studies 24(2): 111-134. doi: 10.1016/S0142-694X(02)00034-0

Weisberg, R. W. 2009. "On 'out-of-the-box' thinking in creativity." Tools for innovation: 23-47. doi: 10.1093/acprof:oso/9780195381634.001.0001.