A scientific framework for testing creativity enhancing techniques

Gaetano Cascini; Niccolò Becattini
Politecnico di Milano, Dept. of Mechanical Engineering
Corresponding Author Email: gaetno.cascini@polimi.it

SESSION: S2: Educating the Edisons of the 21st Century

CONTEXT Among the hot topics of Engineering Design research, creativity enhancing techniques play a key role with abundancy of approaches, methods and tools proposed by different authors to address a range of objectives such as exploring new product opportunities, suitable framing of ill-defined problems, overcoming mental inertia and design fixation phenomena, improving co-creative activities etc. Typically, those research activities claim some advantages with respect to the existing literature through case studies or sometimes through controlled experiments. However, the impact of the application context, the attitude and background of participants, but even the variability of the design task of the experiment have a dramatic impact on the performance of the subjects. Consequently, it is very hard identifying regularities and common aspects to draw some general conclusions and make a step ahead towards the definition of comprehensive models and related educational approaches. A further limitation is that variability of testing conditions and observation means make it difficult to replicate those experiments to scale-up the investigation with more significant data sets.

PURPOSE This study aims at proposing a reference set-up for design experiments, suitable to test creativity-enhancing techniques under controlled conditions.

APPROACH The proposed testing framework can involve either students or practitioners, working in teams or as individuals, divided in different groups of treatment, one of which receives no treatment and plays the role of control group. Each group should count a sufficient number of individuals or teams receiving the same treatment to allow for analysis of statistical significance of the results. The experiment consists of two rounds: the first does not involve the application of any treatment to check whether the testing groups show homogeneous performance in terms of idea generation capabilities. In the second round, all groups but the control one, receive some specific instruction or set of stimuli, suitably differentiated, to observe the impact of the creativity-enhancing techniques under study. A panel of experts assesses groups' performance by means of the evaluation of all the generated ideas, according to four reference metrics.

RESULTS The proposed experimental set-up proved to be applicable in several contexts, to compare the impact of different creativity enhancing techniques. So far, it allowed to compare the effects of: *(i)* creative stimuli based on analogical thinking; *(ii)* information extracted from patents as a trigger for idea generation; *(iii)* the introduction of external knowledge from biology represented with different functional/causal modelling techniques.

CONCLUSIONS The paper proposes a structured experimental approach to test creativity-enhancing techniques in terms of impact on design performance and usability, as well as to monitor the learning process of target groups. The tests carried out so far have mostly explored the applicability of the experimental set up varying different creative techniques and operational environments; the structured definition of the experiment enables the replication of a test by different researchers with controlled changes on the experimental conditions.

KEYWORDS Design creativity, creativity stimuli, creativity metrics.

Introduction

Design creativity has become a hot topic in Engineering Design for at least a decade, and significant advancements have been done in understanding how to activate creative thinking while performing design tasks. Creative stimuli leverage several mechanisms such as the identification of analogies (e.g., Christensen & Schunn 2007; Goel 1997) or the guided exploration of resources available in the system or easily accessible (e.g. Becattini et al, 2012). Even though some experiments show that also “internal” stimuli, i.e. retrieved from the company's information repository, can bring advantages to idea generation in brainstorming sessions (e.g., Howard et al., 2011), more extensive studies are dedicated to the injection of relevant pieces of information from other fields of application (e.g. through patents, as in Fu et al, 2015), or from other domains (e.g. Bar-Cohen, 2006; Bonser & Vincent, 2007). The research in this domain investigates several further aspects, such as the degree of interactions in brainstorming activities (Faure, 2004) or the means used to propose creative stimuli, e.g. in textual, graphical or mixed form (Gonçalves et al., 2016).

Overall, this multitude of studies has increased the knowledge on influential factors, but several fundamental limitations remain:

- in many cases, the conclusions of proposed experiments are not robust enough, due to the difficulty of involving large samples of subjects, especially when these are practitioners from industry;
- nor it is possible to scale-up complementary studies by comparing their results, since many uncontrolled variables change in the set-up of the experiments (e.g., the application context, the attitude and background of participants, the design task of the experiment), thus affecting the interpretation of the observed phenomena.

Not surprisingly, contradictory outcomes emerge from apparently similar studies, and this makes it very difficult the identification of regularities and common aspects to draw some general conclusions and make a step ahead towards the definition of comprehensive models. In turn, this limits the efficient advancement of the research in the field.

This study proposes a reference set-up for design experiments, suitable to test creativity-enhancing techniques under controlled conditions. The proposed testing framework exploits the experience of previous experimental activities carried out by the authors and by other scholars in the engineering design domain. As such, more researchers might organize their studies according to the same framework, thus allowing regular and structured replications of experiments, as for instance in the pioneer experience proposed by Belski et al. (2016).

The paper is structured as follows: the second section describes the proposed framework and its rationale; then, two exemplary applications of the same are proposed, to show the adaptability of the framework to different contexts and specific objectives. Eventually, the last section proposes the conclusions of the authors based on the evidences emerged so far.

The testing framework

The testing framework has to be adaptable to a large variety of experimental conditions, so as to allow the observation of the influence of diverse controllable factors, such as:

- the profile of the subjects (e.g., distinguished by age, gender, education and background, cultural environment, etc.);
- the design task (e.g., constrained problem solving task, exploration of the design space through divergent thinking, etc.);
- the designing conditions (e.g., duration of the design activity, work in teams or as individuals, the features of the test location etc.);
- the creative stimuli differentiated in terms of source of information, expected mechanism of creativity enhancement, representation form of the stimuli, etc.

At least two different groups of treatment should be set up, one of which receives no

treatment and plays the role of control group. In the authors' experience, typically four groups are organized, i.e. a control group and three different groups of treatment, so as to allow a broader exploration of the impact of influential factors on idea generation.

Each group should count a sufficient number of individuals or teams receiving the same treatment to allow for analysis of statistical significance of the results. The basic elements that constitute a modular unit of the experimental set-up are the followings: several individuals or teams (subjects) belonging to a certain group receive the same treatment; the controlled input is a design task to be approached by the subjects; the observed output is the collection of all the ideas generated by the subjects; other influential factors, such as the designing conditions and the creative stimuli, are controlled so as to ensure the same treatment to the entire group.

The experiment consists of two rounds (figure 1): the first one does not involve any induced method or techniques, nor it introduces any external stimulus to the subjects. The rationale of this round is to check whether the testing groups show homogeneous performance in terms of idea generation capabilities. Therefore, the outputs produced by the groups in this phase are compared through the ideation metrics described below and the randomization of the subjects is considered adequate if no significant statistical differences emerge.

In the second round, instead, all groups but the control one (Adair et al., 1990) receive some specific treatment, i.e. they are exposed to different instructions or set of stimuli, or they are invited to work in different controlled operating conditions. The treatment of the test groups should be suitably differentiated so that:

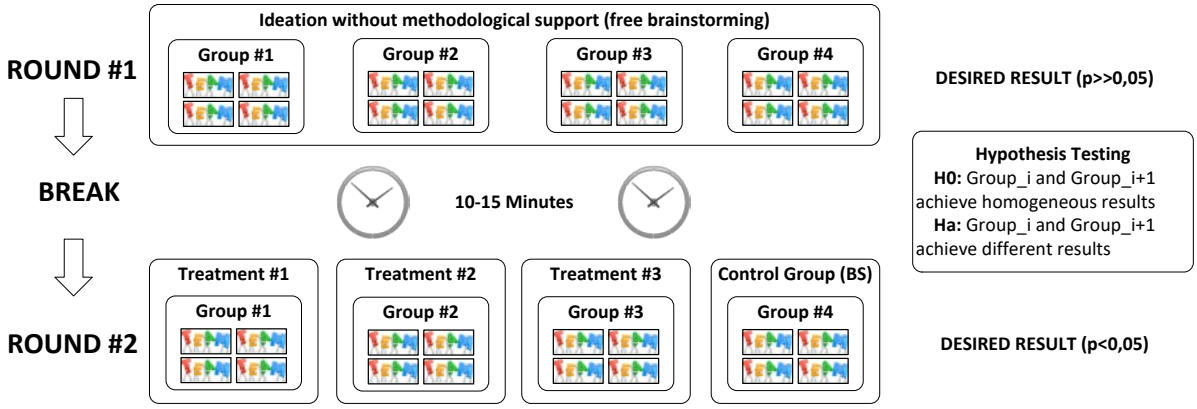

Figure 1: Organization of the two rounds of the experiment

- statistically significant differences can be observed between the output produced by the test groups;
- the controlled variations of the testing conditions are coherent with the objectives of the specific study.

All the generated ideas (of either rounds of the experiment) are evaluated by a panel of experts who assess groups' performance by means of four well-established metrics in the field literature (Shah et al., 2003):

- Fluency, as a measure of quantity of generated ideas;
- Novelty, as the difference between the generated idea and the initial state of the solution (ex-ante or a-priori), or in terms of originality with respect to the rest of the cohort of subjects (ex-post or a-posteriori);
- Variety, as a measure of the diversity among a set of generated ideas, representing the capability to explore the solution space;
- Quality, as a measure of the goodness/viability of a generated idea (if possible) or completeness of the idea description (under the assumption that more detailed descriptions are more likely to be based on supporting arguments).

More in detail, the assessment of ideas according to the above metrics performed by the

panel of experts should be checked in terms of degree of agreement among the raters, e.g. by means of the Kendall's W coefficient of concordance. The results of the individuals/teams constituting each group are clustered to generate a descriptive statistic for the population administered with the same treatment. The resulting data are compared to evaluate:

- the uniformity among the different groups before any treatment (Round 1): this check on the goodness of the randomization for group composition is a necessary pre-condition for the meaningfulness of the second round of the experiment;
- the differences between the different behaviour of the subjects exposed to a different treatment (Round 2).

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 different treatments are 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.

Exemplary applications of the testing framework

This section briefly describes two different experiments based on the proposed testing framework, so as to show its applicability in different situations and with different specific objectives, despite sharing the same overall goal to observe the impact of different creative stimuli on the ideation performance.

Testing creative stimuli for design-by-analogy

Design-by-analogy is a well-known practice in which analogy is applied in the design process for helping the designers to get inspired to solve a target problem (Christensen & Schunn, 2007; Goel, 1997). Near-field analogy appears when the target and the analogical source are from the same or very similar problem domain, while far-field analogy appears when they are from different problem domains. Despite the effect of analogies on design performance has been studied by many scholars and several literature sources share the results of carefully conducted experiments (e.g. Chan et al., 2011), contradictory conclusions emerged when comparing the outcomes of stimuli characterized by different analogical distances. It is therefore necessary to extend such studies to a wider set of tests to understand the impact of factors not yet controlled by the researchers. Among these, the authors intend to explore the impact of specific training on the capability of designers, either novices or experts: the recognition of analogies and the generation of idea through the transposition of some elements into the desired target is not necessarily intuitive and can be improved by practice.

With this purpose, a sample of novice mechanical designers, subdivided in small design teams, participated in an experimental activity structured as described in the previous section. 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 subjects also composed the groups exposed to near-field analogies, middle-field analogies and far-field sources of analogy. Each group counted 7 design teams 3 members each to recreate a typical collaborative design session.

The first round of the test consisted of a 10-minute brainstorming activity with no stimuli. In the second round, two different stimuli were introduced every 10 minutes for a total amount of 40 minutes (8 stimuli per team).

The design task was the proposal of ideas for the next generation of vacuum cleaning robots. The introductory presentation discussed some of the most common problems of these device as a preliminary design brief: quality and efficiency of cleaning are low, especially for corners and edges of the room; the collection device is hard to clean; the robot easily gets stuck and

the wheels are also easily twined by strings, cables, etc.; other problems, such as high noise, insufficient energy, etc. During the presentation of the case study, students were encouraged to generate also ideas beyond the scope of the proposed problems. The design stimuli, presented to the subjects as a combination of text and picture, were devices or products supposedly suitable as stimulus to address the problems of cleaning, moving, saving energy and decreasing noise.

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.

The results of the Kruskall-Wallis test on the output of the first round of the experiment confirmed that there are no significant differences between the groups in terms of quantity of generated ideas ($p=0.904>0.05$), novelty of the generated ideas computed as the average score of novelty for each of the groups ($p=0.228>0.05$), average quality of ideas among groups ($p=0.692>0.05$) and related variety of what they ideated ($p=0.838>0.05$). In other terms, it can be stated that groups have been properly created through randomization and therefore the emerging differences between groups in the second phase, if any, depend on the effect of analogical stimuli the groups are exposed to.

The analysis of the ideas generated in the second round of the experiment through 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 in this test had a statistically evident influence on the design outcome. On the contrary, the results show that the differences among quantity of ideas among the groups did not (statistically) depend on the stimuli ($p=0.116>0.05$). The authors interpret the latter as due to the small amount of time assigned to the subjects for the generation of ideas after the introduction of each pair of stimuli.

Figure 2 shows a comparison of the performance metrics related to novelty, variety and quality achieved by the four groups of treatment.

Without focusing on details that are beyond the goal of the present paper, the following essential results emerge from this experiment:

- Average novelty index displays that the ideas inspired by near-field analogies are significantly more novel than the ones generated with different treatments, including the Control Group. In addition, 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.
- Similarly, the quality of ideas inspired by near-field analogies is better than the ones inspired by the far-field analogy and control conditions. 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.

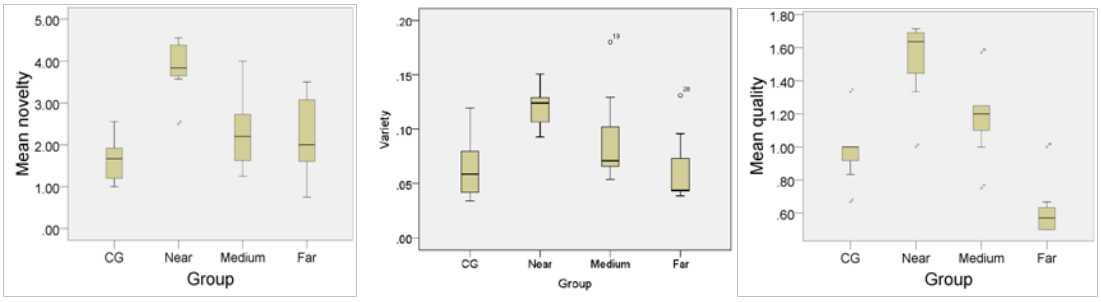

Figure 2: Comparison of different analogical distance stimuli (metrics from Shah et al., 2003)

- 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 shows that near-field analogies play the most beneficial role in the ideation at least when proposed to designers not specifically trained on this practice.

Testing information extracted from patents as creative stimuli

As a second exemplary application of the same experimental framework, the authors propose hereafter an experiment with significantly different conditions and specific objectives with respect to the study described in the previous section: practitioners instead of students as subjects; information extracted from patents as design stimuli; analysis of the impact of different representations on the ideation performance of designers.

The rationale of this study (Authors, 2017) is that among the creative stimuli proposed in literature, patents represent a suitable option for R&D engineers, since they are supposed to be acquainted with this kind of documents, and they are not expected to be as sceptic as for other documents of non-technical origin. Patents as stimuli, even if recalled from the same domain, can be meaningful for these target users since also incremental innovations might allow the exploitation of a fruitful business for the patenting subjects.

To study the effect of different patent content representations as creative stimuli in ideation processes, the investigation requires comparing the outcomes of alternative approaches using patents. Since analogies have witnessed to be effective for increasing the novelty and the quantity of ideas (Chan et al., 2011), and patents as sources of analogies demonstrated effectiveness in idea generation (Fu et al, 2015), their simplest representation (Patent Full-Text - PT) has been chosen as one of the stimuli for the comparison of treatments.

Two more representation forms have been added in the experiment:

- Problem-Solution Matrix (PS) Map (Suzuki, 2011);
- An original diagram (Figure 3, Parvin et al., 2017) based on the concept of TRIZ contradiction, since compared with alternatives such as fishbone and fault-tree diagrams, TRIZ contradictions focus on problems and solution using design parameters, corresponding respectively to design variables (what designers can change) and design requirements (Becattini & Cascini, 2013).

Overall, four different treatments were adopted for this experiment: Brainstorming (Control Group - BS Group); the Problem-Solution Matrix Map (PS group); the TRIZ Contradiction map (PS+TC group), and Patent Full-Text (PT group). The last three sets of stimuli were built out of the same patent corpus.

The experiment was conducted in Iran in collaboration with a training and consulting institute: fifty-six (56) R&D engineers (45 M; 11 F), with an average working experience of 9,2 years (Std. deviation 1.5) and different backgrounds (mechanical/industrial, chemical, electronics, computer science…), were involved and randomly subdivided into 4 groups, one per each of the abovementioned treatments. The 14 people in each group were randomly organized into 7 teams of 2 engineers each. This enables gathering a sufficient amount of experimental data for each treatment (7 data points) and have real-like design conditions for the team (2 people sharing and generating ideas on the design task).

The experiment was organized into two rounds – 30 minutes each consistent with the findings of Howard et al. (2010) on the saturation of generativity. During the first round, the 28 teams of R&D engineers were asked to generate inventive ideas they consider worth of patenting. The second round started after a 15 minute break and the 4 groups received a different treatment each.

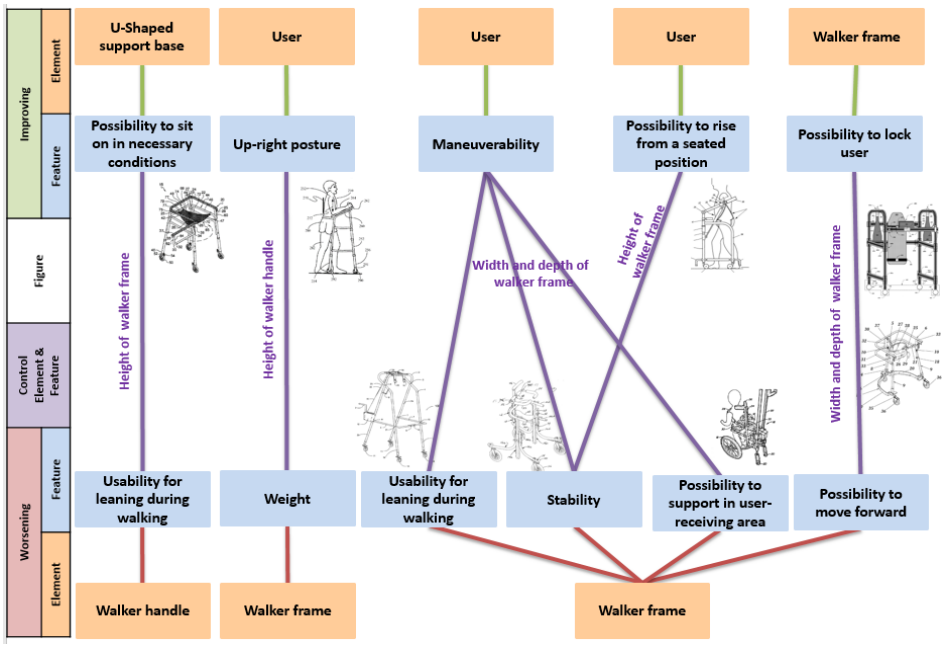

Figure 3: Exemplary TRIZ contradiction map used as patent information-based creative stimuli

To avoid biases due to misinterpretations, the design task regarded the generation of inventive ideas for a walker (or walking frame), a device for disabled or elderly people who need additional support to maintain balance or stability while walking. Its intrinsic simplicity released the subjects from having specific competencies to understand its current functions and related working principles. The data collected in the experiment were analysed with the same metrics described in the previous example and the same logic (for more details, see Parvin et al., 2017). Overall, the following conclusions were drawn:

- patent-based design tools trigger better performances than brainstorming;
- there is no statistically significant difference between the use of the patent map per se and with the contradiction map for none of the three metrics, despite the results suggest potential advantages regarding enhanced novelty;
- the combined use of the problem-solution and contradiction maps increases, with statistical significance, the quantity of ideas, compared to full-text patent stimuli and simple brainstorming;
- the combined use of the problem-solution patent map and contradiction map increases, with statistical significance, the novelty of ideas, compared to simple brainstorming;
- the use of the problem-solution patent map per se increases, with statistical significance, the variety of ideas, compared to full-text stimuli and simple brainstorming.

Conclusions

The paper proposes a structured experimental approach to test creativity-enhancing techniques in terms of both impact on design performance and usability, as well as to monitor the learning process of target group.

The proposed experimental set-up proved to be applicable in several contexts, to compare the impact of different creativity enhancing techniques. So far, it allowed to compare the effects of: creative stimuli (as examples taken from fields at different distance from the target application) on analogical thinking; information extracted from patents and represented in different forms as a trigger for idea generation; the introduction of external knowledge from biology represented with different functional/causal modelling techniques (not described in this paper due to space limitations).

Among the results, it is worth mentioning that sometimes the experiments brought some unexpected results with respect to other literature studies, which can be justified by reflecting on the specific cultural or educational context where the test was conducted. Those reflections might have significant implications in the definition of educational approaches.

Despite the tests carried out so far have mostly explored the applicability of the experimental set up varying different creative techniques and operational environments, the structured definition of the experiment enables the efficient replication of a test by different researchers with controlled changes on the experimental conditions. The authors believe that efforts in this direction will lead to a higher reusability of research achievements and, ultimately, to a more efficient growth of the scientific community working on design creativity.

References

Adair JG, Sharpe D, & Huynh CL (1990) The placebo control group: an analysis of its effectiveness in educational research. *Journal of Experimental Education* 59(1): 67-86.

Bar-Cohen, Y. (2006). Biomimetics-using nature to inspire human innovation. *Bioinspiration & Biomimetics*, 1(1): 1-12.

Becattini N., Borgianni Y., Cascini G., & Rotini F. (2012) Model and Algorithm for Computer-Aided Inventive Problem Solving. *Computer-Aided Design* 44: 961-986.

Becattini, N., & Cascini, G. (2013). Mapping Causal Relationships and Conflicts among Design Parameters and System Requirements. *Computer-Aided Design and Applications*, 10(4): 643-662.

Belski I., Skiadopoulos A., Aranda-Mena G., Cascini G., & Russo D. (2016) Idea Generation with Substance-Field Analysis: the Influence of Prior Knowledge and Practical Experience. Paper presented at the TRIZ Future Conference 2016, 24-27 October 2016, ,Wroclaw, Poland.

Bonser, R. H. C., & Vincent, J. F. V. (2007) Technology trajectories, innovation, and the growth of biomimetics. *Proceedings of the Institution of Mechanical Engineers, Part C: Journal of Mechanical Engineering Science* 221(10): 1177-1180.

Chan, J., K. Fu, C. Schunn, J. Cagan, Wood K., & Kotovsky K. (2011) On the benefits and pitfalls of analogies for innovative design: Ideation performance based on analogical distance, commonness, and modality of examples. *Journal of Mechanical Design* 133(8): 081004 1-11.

Christensen, B. T., & 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.

Faure C. (2004) Beyond brainstorming: effects of different group procedures on selection of ideas and satisfaction with the process. Journal of Creative Behavior 38(1):13-34.

Fu, K., Murphy, J., Yang, M., Otto, K., Jensen, D., & Wood, K. (2015). Design-by-analogy: experimental evaluation of a functional analogy search methodology for concept generation improvement. *Research in Engineering Design* 26(1), 77-95.

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

Gonçalves M, Cardoso C. & Badke-Schaub P. (2016) Inspiration choices that matter: the selection of external stimuli during ideation. *Design Science*, 2(e10).

Howard, T. J., Dekoninck, E. A., & Culley, S. J. (2010). The use of creative stimuli at early stages of industrial product innovation. *Research in Engineering Design*, 21(4): 263-274.

Howard, T. J., Culley, S. J. & Dekoninck, E. A. (2011). Reuse of ideas and concepts for creative stimuli in engineering design. *Journal of Engineering Design*, 22(8), 565-581.

Parvin M., Cascini G. & Becattini N. (2017) Information extracted from patents as creative stimuli for product innovation. Presented at the 21st International Conference on Engineering Design (ICED17), Vancouver, Canada, 21-25 August 2017.

Shah, J. J., Smith, S. M., & Vargas-Hernandez, N. (2003). Metrics for measuring ideation effectiveness. *Design studies* 24(2): 111-134.

Suzuki, S. I. (2011). Introduction to patent map analysis. Japan Patent Office.