

Testing ideation performance on a large set of designers: effects of analogical distance

This paper reports an experiment that aims at providing new evidence to reinforce contradictory conclusions, as available in engineering design literature, on the effect of analogical distance on ideation performance of designers. The experiment involved 84 graduate students in Mechanical Engineering, with typical competencies on engineering design, but without any specific skill on analogy-based idea generation. The subjects were exposed to three different treatments (near-, medium- far-field analogies) whose design performance has been compared to a control group. The observation confirms that designers not specifically trained in designing-by-analogy get benefit from sources of inspiration that share the same (sub-)functions and context of the target system, as their idea generation process lead to ideas having higher novelty and quality compared to the outcomes of subjects exposed to more distant analogies. The observed results also show that exploration of the design space gets positively affected.

Keywords: analogical reasoning, idea generation, creative process, creativity evaluation, inspiration

1. Introduction and background

Within the research threads of cognitive science, reasoning by analogy captured the attention of psychologists from the early 70s. Several scholars recognized its importance in cognition as it emerges and occurs quite naturally during thinking (e.g. Ross, 1970; Sternberg, 1977; Gick and Holyoak, 1980). Gentner (1983), for instance, described it as a process of knowledge mapping between two domains or situations, through the analysis of similarity among attributes, relations or representations that should match with each other.

More recently, other scholars elaborated on the importance that analogy plays in design activities, such as problem interpretation, problem-solving and idea generation (among the others: Dahl and Moreau 2002; Kao 2014; Ward 1998), while Gentner and

Jeziorski (1989) also highlighted that analogy plays a relevant role in science in general, from teaching students to scientific discovery. Engineering design literature tackles analogical reasoning from various perspectives. Grounding on the improved understanding of what analogy is and how analogical reasoning works, recent research trends focus on how to leverage it and make people find inspiration to generate new ideas and concepts. Creative stimuli that trigger reasoning by analogy have been progressively studied in order to clarify their effects on the design process outputs, according to an increasing number of conditions, such as the definition of analogical distance, the kind of design task or the experience of the involved designers (For instance in Ahmed and Christensen, 2009; Chan, Dow, and Schunn, 2015; Chan et al., 2011; Linsey, Wood, and Markman 2008; Tseng et al. 2008).

1.1. Analogy-based design

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, typically, for helping the designers to get inspired to solve the target problems and, more in general, to improve the overall efficiency of the design process (Christensen and Schunn 2007; Fu et al. 2013; Goel 1997).

The investigation on the role of analogies in design takes into account different factors that might affect designers' performance. Different sources of information have proved to be effective to stimulate creativity by leveraging analogical thinking. Several studies considered the effectiveness of technical sources of inspiration by using contents available in patents (e.g. Fu et al, 2015). In stark contrast, or as a complementary viewpoint on sources of inspiration, the effectiveness of contents from the biological

domain also proved to provide an effective contribution to the exploration of the design space and idea generation (Wilson et al, 2010). The organization of such knowledge is also crucial, as it should allow to both retrieve relevant contents and display these contents to designers in order to improve their idea generation effectiveness (Author, 2015; Authors, 2017). On a similar wavelength, different authors carried out experiments with different modalities of administration for stimuli, showing that mixed textual and graphical contents are more effective to reduce design fixation than purely textual or graphical ones (Goncalves et al., 2016). Chan et al. (2011), as well, explored how different modalities of analogical stimuli influence the design outcomes. They showed evidence that different modalities might affect generativity, but in their results there is no effect on the novelty and quality of solution concepts.

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). From a more abstract perspective, Ward (1998) and Bearman, Ball and Ormerod (2002) discussed the role analogy performs in engineering design and analysed the relationships between analogical distance and functions (from a target domain or situation) as this has been considered a key factor to effectiveness in design.

Beyond the specific differences and nuances that different scholars put under their lenses, most of the abovementioned ones provided just limited experimental data that allowed for a preliminary estimation of these effects on design performances.

1.2. Analogical Distance

As briefly mentioned above, a relevant research thread focuses on analogical distance, which is a measurement of the degree of structural (relational or functional) similarity and surface (attributonal) similarity. To this regard, Dunbar (1997) provides a preliminary classification of analogical distance from a biological perspective. He

identified three degrees of distance: 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 the 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.

Overall, in engineering design literature, the analogical distance of two systems has been considered from two different perspectives. First, as the proximity of the two systems in terms of purpose/function and, second, as the affinity of their product category and industry (see, for instance, Chan et al. 2011; Fu et al. 2013; Kalogerakis, Lüthje, and Herstatt 2010).

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). Kennedy et al (2018) also showed that far-field analogies are beneficial on ideation effectiveness for novelty, relevance and effectiveness of solutions. This study also shows that novelty improves with stimuli from the biological domain, while relevance and effectiveness are positively correlated with far-field analogies in the technological domain. 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). Srinivasan et al. (2018) recognized that near-field analogies are used more frequently, but they also noticed a positive correlation between distance and novelty (increasing distance, increasing novelty) and a negative correlation for quality (increasing distance, decreasing quality). A previous study by the same group of authors showed that there was no correlation between analogical distance (domain proximity) and novelty and quality of ideas (Venkatamaran et al, 2017). A further issue that might explain the contradictory outcomes of apparently similar studies in this field is the impact of specific training on the skill of the 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, as for instance noticed by Keshwani et al (2017). They noticed that not fully expected results might depend on the lack of the specific knowledge in the domain the stimulus comes from (e.g. biological) and that training, as well, is critical to improve the effectiveness of ideation.

1.3 Motivation

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, in order to provide new evidence and more definitive conclusions about what still remains controversial. Results of experiments on a large scale (in terms of participating subjects) allow populating database or repositories collecting the outcomes of experiments on design creativity, so that these results can be eventually cross compared among different experiments and, eventually, interpreted according to the background and skills differences among participants. This paper presents a first step towards the extension of these studies to a large set of subjects, which here share the same background and geographical location. 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 subjects were asked to generate ideas freely brainstorming. 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 (Figure 1).

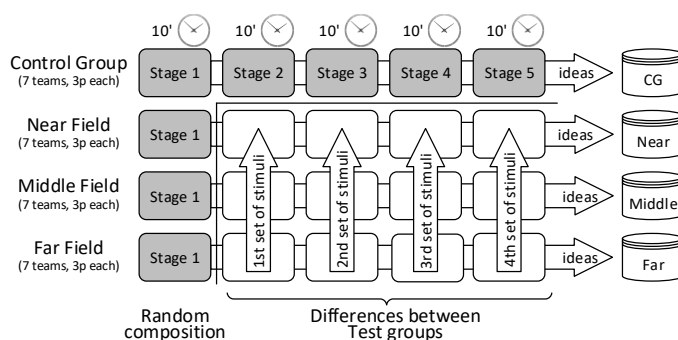


Figure1. Experiment setup

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 devices 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.

Figure 2 provides some examples of ideas as noted from the subjects participating in the experiment.

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 a 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. The stimuli used along the experiment include a description of the function, of the working principles and/or the context of use, together with a web-retrieved picture of the product used as a source of analogy. The whole set is available with the supplementary material of this paper. Table 2 shows an example of stimulus for each of the considered analogical distances.

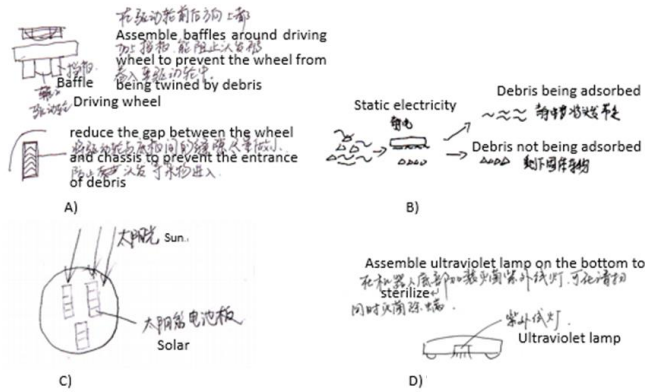


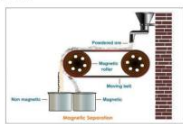


Figure 2. Four examples of ideas generated and noted by the subjects as a combination of text and sketches.

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. Examples of stimuli from different analogical distances

Type	Example	Content of the stimulus
Near-field stimuli		<p><i>Text:</i> Dust particles are adsorbed through electrostatic and fall into the collection plates after losing electricity.</p> <p><i>Function:</i> clean dust</p> <p><i>Context:</i> domestic use (same context of the vacuum cleaner)</p> <p><i>Source:</i> http://i05.c.aliimg.com/img/ibank/2013/021/416/911614120_93291792.jpg</p>
Medium-field stimuli		<p><i>Text:</i> The stress that bondless windscreen wipers support is distributed. And they fit with glass perfectly, reduce the damage and eliminate the vibration caused by wind.</p> <p><i>Function:</i> clean dust</p> <p><i>Context:</i> outdoor use/road infrastructure</p> <p><i>Source:</i> http://i3.sinaimg.cn/qc/photo1/09newmtzlt/U121P33T148D334546F2100DT20090303212900.jpg</p>
Far-field stimuli		<p><i>Text:</i> Being applied in mining iron, garbage classification contexts, magnetic separation can extract magnetic material from a mixture through a magnetic field.</p> <p><i>Function:</i> separate solids (different but relevant function)</p> <p><i>Context:</i> industrial use</p> <p><i>Source:</i> http://images.tutorvista.com/content/metals-non-metals/magnetic-separation-process.jpeg</p>

The ideation performance of each team has been measured according to four metrics: Quantity, Variety, Novelty and Quality. These 4 metrics allow comparing different design approaches (Shah et al, 2003). Figure 3 presents the scores for each metrics.

Novelty measurement presents a variety of approaches: beyond Shah et al. (2003), the literature also presents some proposals of refinement which consider different degrees of novelty depending on differences at functional and structural level, obviously considering also changes for physical and working principles – similar to Shah's variety (e.g. Peeters et al, 2010). With a similar degree of granularity for novelty assessment, Sarkar and Chakrabarti (2011) proposed a different modelling reference framework to characterize novelty with respect to the previous generation of a technology. In this work, differently from other metrics, Novelty has been measured according to the approach proposed by Sarkar and Chakrabarti (2011), as it provides a more repeatable approach based on structured comparisons with previous versions of the solution. The assessment of ideas according to the metrics has 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 generating 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 3 yields the value for Variety, where k means the number of functions, f_k means the weight of function k , T_k is the number of ideas related to function k in team i , and S_k is the overall number of different ideas related to function k in all the teams of that group. This choice to limit the investigation on variety to the amount of functions and to account for the related sets of ideas is due to the detail level of description of ideas generated during the experiment (not all of them described solutions with details on the embodiment and on the details). However, the number of ideas is sufficiently representative of different working/physical principles, which correspond to the branches to count as proposed by Shah et al. (2003) to compute variety (weights for branches - e.g. 10, 6, 3, 1 - at different levels are, therefore, meaningless). From this perspective, the only meaningful parameter to determine is the weight for functions, as the chosen metrics allows for a clear description of exploration through variety, preventing the risks highlighted by Nelson et al. (2009). It appears also not particularly relevant to this measurement. Table 3 shows the list of functions and the normalized weights.

Table 3. Functions and weights

<u>Function</u>	<u>Weight</u>	<u>Function</u>	<u>Weight</u>
<u>Clear debris</u>	<u>0.10</u>	<u>Enhance capacity of the battery</u>	<u>0.09</u>
<u>Fill debris</u>	<u>0.10</u>	<u>Transform available energy into electricity</u>	<u>0.09</u>
<u>Reduce noise</u>	<u>0.08</u>	<u>Protect brush from being twined</u>	<u>0.11</u>
<u>Emit music</u>	<u>0.04</u>	<u>Protect robot from being stuck</u>	<u>0.11</u>

<u>Adjust quality of air</u>	<u>0.05</u>	<u>Protect wheel from being twined</u>	<u>0.11</u>
<u>Reduce energy consumptions</u>	<u>0.09</u>	<u>Clear glass</u>	<u>0.03</u>

Both Quantity and Variety are computed considering the whole set of ideas generated by the teams.

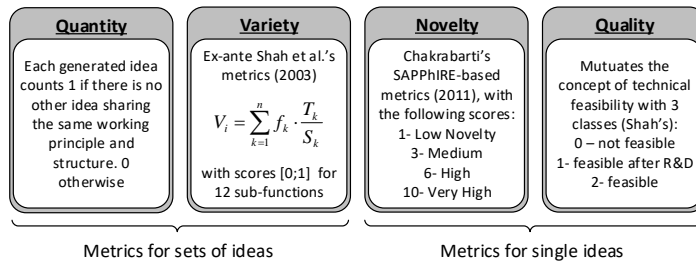


Figure 3. Metrics for the evaluation of ideation performance

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:

$$\text{Average}_{\text{Novelty (Quality)}} = \frac{\sum_1^{\text{Quantity}} \text{Novelty (Quality) score}_i}{\text{Quantity}} \quad (1)$$

- 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:

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

- 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.

$$\text{Max}_{\text{Novelty (Quality)}} = \sum_1^N \text{Novelty (Quality) score}_i$$

$$\text{Novelty (Quality) score}_i \in \text{TopRated} = \{\text{ideas with Top - Rated scores}\}^{(3)}$$

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 it was taken into account just the idea which received the highest score in the whole set. 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 estimating 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 2, the group composition was determined by chance, thus following a randomization process. Table 3 summarizes the outcomes recorded after the first stage of designing, as for Figure 1. Groups in Table 3 are organized as for Figure 1 as well.

As in Stage 1 there is no treatment, whatever the group is, Table 4 describes groups with capital letters (A, B, C and D) which correspond to, respectively, the treatments they will be exposed to, from Stage 2 to 5 (Control Group, Near-field analogies, Medium-field analogies, Far-field analogies).

Table 4. Outcomes of idea generation after Stage 1

		Team 1	Team 2	Team 3	Team 4	Team 5	Team 6	Team 7
Group A (Control Group from stage 2)	Quantity	1	1	2	1	1	1	2
	Avg novelty	3.00	3.00	1.00	1.00	6.00	1.00	2.00
	Avg quality	1.00	1.00	1.00	2.00	2.00	2.00	0.50
	Variety	0.09	0.01	0.03	0.05	0.01	0.02	0.02
Group B (Far-field Group from stage 2)	Quantity	2	1	2	1	2	1	1
	Avg novelty	1.00	1.00	0.50	3.00	2.00	1.00	1.00
	Avg quality	1.00	2.00	2.00	2.00	1.00	2.00	2.00
	Variety	0.02	0.01	0.05	0.04	0.05	0.01	0.02
Group C (Medium-field group from stage 2)	Quantity	1	2	1	1	3	1	2
	Avg novelty	1.00	1.50	3.00	1.00	0.67	3.00	1.00
	Avg quality	2.00	2.00	2.00	0.00	0.00	2.00	1.00
	Variety	0.01	0.02	0.01	0.01	0.06	0.01	0.09
Group D	Quantity	1	2	2	1	1	2	1
	Avg novelty	3.00	1.00	1.00	0.00	1.00	0.50	1.00
	Avg quality	1.00	2.00	1.50	2.00	2.00	0.00	1.00

(Far-field group from stage 2)	Variety	0.02	0.02	0.06	0.02	0.02	0.03	0.03
--------------------------------	---------	------	------	------	------	------	------	------

The results of the Kruskal-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, despite the results on variety do not have particular meaning in case of a team generating one single idea during Stage 1. According to the recorded results, 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, for what concerns Stages 2-5, show that after the first stage of idea generation all the groups started being more productive. Table 5 presents the quantity of ideas generated by the different teams and according to the treatment they were exposed to. These figures show that the differences in terms of quantity of ideas among the groups do not (statistically) depend on the stimuli ($p=0.116>0.05$ with Kruskal-Wallis test).

On the contrary, the recorded results show that differences among groups are statistically 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.

Table 5. Distribution of quantity among each team

	Control Group	Near Group	Medium Group	Far Group
Team 1	4	6	5	5
Team 2	4	6	4	7
Team 3	9	6	8	4
Team 4	5	7	5	4
Team 5	4	9	6	4
Team 6	6	7	7	4
Team 7	6	11	4	9
Average	5.43	7.43	5.57	5.29
Std. Dev	1.81	1.90	1.51	1.98

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 4 (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 4 (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 it 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 3 provides the values for each of the 7 teams in each group. The Mann-Whitney test displays that, as for the results about Quality, significant differences of Variety occurred just between the Near group and, respectively, the Control Group and the Far-Group. The distributions of the treated groups (Figure 4, 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 what emerged in terms of Quality of ideas, the performance of Variety 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.

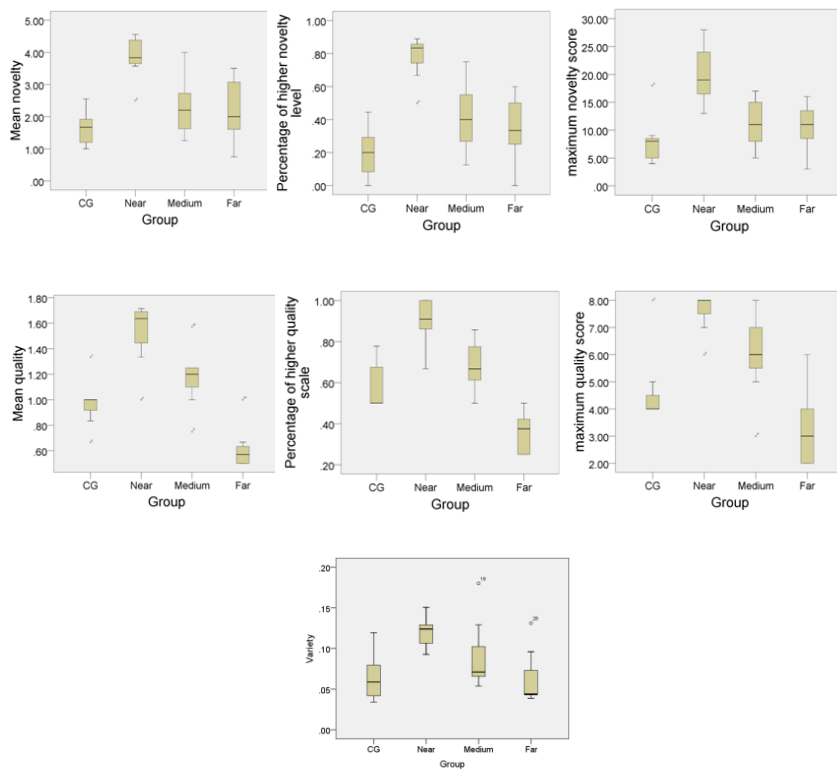


Figure 4. Boxplots. First row: Novelty. Second row: Quality. Third row: Variety.

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 as shown in Figure 5. This effect might depend on the depletion of ideas or on the tiredness of participants.

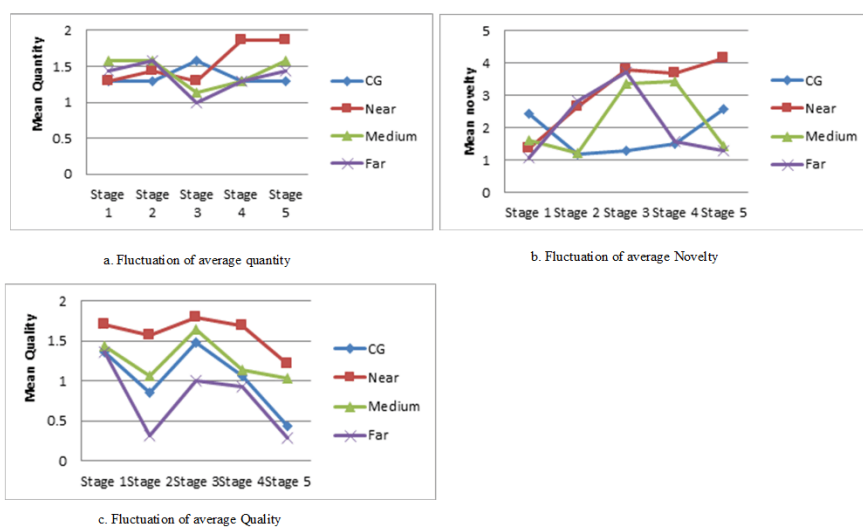


Figure 5. Stage-based diagrams with recorded values for the metrics of average quantity, novelty and quality.

More in detail, the diagram for the quantity of ideas (Fig 5-a) shows that generativity does not sharply vary from Stage 1 to Stage 5, whatever the treatment is. The only exception regards the near-field stimuli, which work better than others, especially in the latest stages.

Figure 5-b indicates that the average novelty of ideas generated in Near group has been rated with a higher score that goes on increasing also in the last stages. On the contrary, novelty in Medium and Far group begins to decrease. Unexpectedly, novelty in Control group increases sharply in stage 5, where probably participants' inspiration finally sparkles after a long-term low performance.

The fluctuations of average quality of ideas is displayed in Figure 5-c. The diagram still shows that near-field analogies provide better results than the others. Overall, the analysis of the dynamics of ideation effectiveness highlight that quality, more than other metrics, shows an emerging rhythm. In such a case, the treatments appear to affect the outcomes of quality of ideas by making the overall trend more or less intensively marked.

Considering the exploratory nature of this kind of time-dependent analysis, the authors believe that further investigation is required to clarify this issue.

With reference to the results presented in Section 3, all the chosen criteria for novelty (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 4 reveals 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. Overall it is possible to state that the results show a negative correlation for analogical distance and novelty of the ideas generated using such stimuli.

This is contradictory with respect to what described by Srinivasan et al (2018) after their experiment involving 105 subjects, as they observed a positive correlation for novelty and analogical distance: higher distance from the problem domain helps generate solutions of higher novelty.

With regard to quality, although there is slight discrepancy among average quality, maximum quality and the percentage of ideas having higher quality, the results of the second row of Figure 4 show a negative correlation between analogical distance and the quality of ideas: Stimuli of closer analogical distance (Near-field) generally

trigger the generation of ideas which are more effective and more capable of matching the design specification. This is consistent with the findings of Chan, Dow, and Schunn (2015) and Srinivasan et al (2018): also in this experiment the quality of generated ideas, as inspired by near-field analogies is better than the ones inspired by the far-field analogy and control conditions. As a possible explanation, this might be due to the nature of distant analogies, which may require designers to retrieve, map and transfer a hardly manageable amount of information for cognition and this appears as more difficult without dedicated training. Nevertheless, other recent studies produced opposed results about quality of ideas, e.g. Kennedy et al (2018) noticed that far-field analogies (within the technological domain) are capable of triggering ideas of high quality (relevant and effective). It is also worth noticing that both the findings of this paper and Srinivasan et al (2018) emerged from experimental activities with a large number of subjects (respectively 84 and 105). Therefore, the extension of this kind of investigation to larger set of subjects, mentioned in section 1.3 as a potential strategy to clarify the effect of analogies on ideation performance, is not, in itself, capable of fully disambiguate the contradictory results emerged in literature. These contradictory results highlight that some variables, which are relevant to describe the whole phenomenon, might be obscure yet. Differences in the experimental protocol concern design team size, the design task as well as the nature of stimuli. Moreover, both these studies have been carried out with the majority of subjects from Asia (respectively, China and Singapore), that have a slightly different background (Engineering students). Obviously students' individual profile might also impact the results, as the two universities where the experiments have been carried out follow different policies for students' access. Both the background and cultural factors, together with others yet to consider, might also be responsible of the misalignment of results across different studies. Then, in

order to clarify the effect of stimuli for design-by-analogy on ideation performance, scholars across the world should define and adopt a more uniform testing protocol which enables cross comparisons to be performed so that results remain meaningful and the currently “hidden” dimensions influencing ideation performance might emerge.

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 in this practice.

5. Conclusions

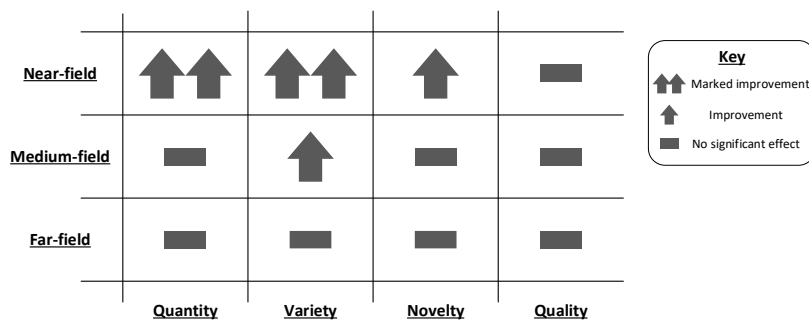
This paper explores the effects of the distance between analogical stimuli and the design target (using near-field, medium-field and far-field analogies) on the idea generation process. The experiment here presented aims at providing further evidence to clarify the controversial results emerged in the 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 (Figure 6) of this experiment show that:

- **Novelty:** Near-field analogies (stimuli sharing the same function and the same context of the design target) are the most effective to generate novel ideas. The effect of medium- (stimuli that just share the function with the design target) 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 (leveraged with stimuli that do not share function nor context with the design target) 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. Stimuli which share function and context with the design target provides the best results for generating ideas which are different from each other.

- There is no statistical significance that the creative stimulation by analogy produces an effect on the quantity of generated ideas.



Codice campo modificato

Figure 6: Summary of the effects of creative stimuli by distance of the source of analogy, with reference to the metrics of design creativity.

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 does 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 need 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 an introduction, however, should not narrow the exploration of the design space by

excessively constraining it. In fact, the set of problems presented before the experiments for the current robot for vacuum cleaning 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 performance. 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 background and previous studies, cultural differences, talent of the subjects as well as the target application field. From this perspective, the authors plan to extend the experiment to a broader sample of designers with a controlled profile. The results presented in the current study can be also used as a potential dataset to compare new results with and address some of the above-mentioned issues.

6. References

- Ahmed, S., and B. T. Christensen. 2009. "An in situ study of analogical reasoning in novice and experienced design engineers." *Journal of Mechanical Design* 131(11): 111004. doi: 10.1115/1.3184693.
- Bearman, C. R., L. J. Ball, and T. C. Ormerod. 2002. "An exploration of real-world analogical problem solving in novices." *Proceedings of the 24th Annual Conference of the Cognitive Science Society, Fairfax, USA, August 7-10.*
- 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.
- Dahl, D. W., and P. Moreau. 2002. "The influence and value of analogical thinking during new product ideation." *Journal of Marketing Research* 39(1): 47-60. doi: 10.1509/jmkr.39.1.47.18930.
- 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.
- Fu, K., Murphy, J., Yang, M., Otto, K., Jensen, D. and Wood, K. (2015), Design-by-analogy: experimental evaluation of a functional analogy search methodology for concept generation improvement, *Research in Engineering Design*, 26(1), 77-95.
- Gentner, D. 1983. "Structure-Mapping: A Theoretical Framework for Analogy." *Cognitive Science* 7(2): 155-170. doi: 10.1016/S0364-0213(83)80009-3

- Gentner D. and Jeziorski, M. (1989). 11. Historical shifts in the use of analogy in Science. *Psychology of science: contributions to metascience*, 296.
- 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.
- Gick, M. L., and K. J. Holyoak. 1980. "Analogical problem solving." *Cognitive Psychology* 12(3): 306-355. doi: 10.1016/0010-0285(80)90013-4.
- 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.
- Gonçalves M, Cardoso C. and Badke-Schaub P. (2016), Inspiration choices that matter: the selection of external stimuli during ideation. *Design Science*, vol.2, e10 DOI: 10.1017/dsj.2016.10
- Kalogerakis, K., C. Lüthje, and C. Herstatt. 2010. "Developing innovations based on analogies: experience from design and engineering consultants." *Journal of Product Innovation Management* 27(3): 418-436. doi: 10.1111/j.1540-5885.2010.00725.x.
- Kao, C. Y. 2014. "Exploring the relationships between analogical, analytical, and creative thinking." *Thinking Skills and Creativity* 13: 80-88. doi: 10.1016/j.tsc.2014.03.006.
- Kennedy, E.B., Miller, D.J. and Niewiarowski, P.H., 2018. Industrial and Biological Analogies Used Creatively by Business Professionals. *Creativity Research Journal*, 30(1), pp.54-66.
- Keshwani, S., Lenau, T.A., Ahmed-Kristensen, S. and Chakrabarti, A. 2017. Comparing novelty of designs from biological-inspiration with those from brainstorming. *Journal of Engineering Design*, 28(10-12), pp.654-680.
- Linsey, J. S., K. L. Wood, and A. B. Markman. 2008. "Increasing innovation: presentation and evaluation of the wordtree design-by-analogy method." *ASME 2008 International Design Engineering Technical Conferences and Computers and Information in Engineering Conference*, New York, USA, August 3-6. doi: 10.1115/DETC2008-49317.
- [Nelson, B.A., Wilson, J.O., Rosen, D. and Yen, J., 2009. Refined metrics for measuring ideation effectiveness. *Design Studies*, 30\(6\), pp.737-743.](#)

- 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.
- Parvin, M., Cascini, G., & Becattini, N. (2017). Information extracted from patents as creative stimuli for product innovation. In *International Conference on Engineering Design 2017* (pp. 297-306).
- [Peeters, J., Verhaegen, P.A., Vandevenne, D. and Dufloy, J.R. \(2010\). Refined metrics for measuring novelty in ideation. IDMM Virtual Concept Research in Interaction Design, Oct, pp.20-22.](#)
- Rice, J. A. 2006. *Mathematical statistics and data analysis*. Belmont: Thomson Higher Education.
- Rosa, F., Cascini, G., & Baldussu, A. (2015). UNO-BID: unified ontology for causal-function modeling in biologically inspired design. *International Journal of Design Creativity and Innovation*, 3(3-4), 177-210.
- Ross, J. F. 1970. "Analogy and the resolution of some cognitive problems." *The Journal of Philosophy* 67(20): 725-746. doi: 10.2307/2024008.
- 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
- Srinivasan, V., Song, B., Luo, J., Subburaj, K., Elara, M.R., Blessing, L. and Wood, K., 2018. Does Analogical Distance Affect Performance of Ideation? *Journal of Mechanical Design*, 140(7), p.071101.
- Sternberg, R. J. 1977. "Component processes in analogical reasoning." *Psychological Review*, 84(4): 353-378. doi: 10.1037/0033-295X.84.4.353.
- Tseng, I., J. Moss, J. Cagan, and K. Kotovsky. 2008. "The role of timing and analogical similarity in the stimulation of idea generation in design." *Design Studies* 29(3): 203-221. doi: 10.1016/j.destud.2008.01.003.
- Venkataraman, S., Song, B., Luo, J., Subburaj, K., Elara, M.R., Blessing, L. and Wood, K., 2017. Investigating effects of stimuli on ideation outcomes. In *DS 87-8 Proceedings of the 21st International Conference on Engineering Design (ICED*

17) Vol 8: Human Behaviour in Design, Vancouver, Canada, 21-25.08. 2017
(pp. 309-318).

Verhaegen, P.A., Vandevenne, D., Peeters, J. and Duflou, J.R. (2013). Refinements to the variety metric for idea evaluation. Design Studies, 34(2), pp.243-263.

Ward, T. B. 1998. "Analogical distance and purpose in creative thought: Mental leaps versus mental hops." In Advances in analogy research: Integration of theory and data from the cognitive, computational, and neural sciences, edited by K. J. Holyoak, D. Gentner, and B. N. Kokinov, 221-230. Sofia: New Bulgarian University.

Wilson, J. O., D. Rosen, B. A. Nelson, and J. Yen. 2010. "The effects of biological examples in idea generation." Design Studies 31(2): 169-186. doi: 10.1016/j.destud.2009.10.003.

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.

Table 1. Summary of selection guidance for analogical stimuli

Table 2. Stimuli from different analogical distance

Table 3. Outcomes of idea generation after Stage 1

Table 4. Distribution of quantity among each team

Figure 1. Experiment setup

Figure 2. Four examples of ideas generated and noted by the subjects as combination of text and sketches.

Figure 3. Metrics for the evaluation of ideation performance

Figure 4. Boxplots. First row: Novelty. Second row: Quality. Third row: Variety

Figure 5. Stage-based diagrams with recorded values for the metrics of average quantity, novelty and quality.

[Figure 6: Summary of the effects of creative stimuli by distance of the source of analogy, with reference to the metrics of design creativity.](#)