ICED13
INTERNATIONAL CONFERENCE ON ENGINEERING DESIGN, ICED13
19-22 AUGUST 2013, SUNGKYUNKWAN UNIVERSITY, SEOUL, KOREA

DOES EXPERIENCE IN DESIGN AND INNOVATION AFFECT THE PERCEPTION OF CREATIVITY?

Niccolò BECATTINI (1), Yuri BORGIANNI (2), Gaetano CASCINI (1), Federico ROTINI (2)
1: Politecnico di Milano, Italy; 2: Università di Firenze, Italy

ABSTRACT
The assessment of creativity arouses increasing interest within design community. The literature witnesses efforts to quantitatively measure creativity, although commonly considered intrinsically subjective. Recent experiences show a good degree of convergence between assessments employing more objective metrics and evaluations of creativity made by experts in design and innovation. With the overall goal of determining whether such judgments are reliable and repeatable, the present paper analyzes creativity assessments of commercial products performed by skilled and novice designers in order to highlight further differences due to accumulated experience. The investigation is carried out by means of a suitable questionnaire asking to evaluate the creativity of 10 market successes and 10 commercial flops. The experiment tests also whether commercial results can strongly influence the perception of creativity. The outcomes reveal that experience is supposed to play a not negligible role in evaluating creativity, while the question about the impact of market success requires further investigation.

Keywords: creativity, innovation, experts' assessment, novices' assessment

Contact:
Dr. Ing. Yuri Borgianni
Università di Firenze
Dipartimento di Ingegneria Industriale
Florence
50139
Italy
yuri.borgianni@unifi.it

1 INTRODUCTION

It is acknowledged in the literature that the capability to generate novel and original ideas is not influenced just by individual talent, but also by systematic procedures and thinking methods (Chulvi et al., 2012; Gero et al., 2012). Shared and repeatable criteria employed to measure creativity of products would strongly contribute to provide greater understanding about the mechanisms that allow designing innovative and successful products. Additionally, the end of a design task would benefit of the reliable assessment of the degree of ideas creativity, so as to be able to select the most proper design alternative. In this sense, objective metrics of creativity would also contribute to better support decisions in the initial stages of design.

The assessment of creativity for the recalled purpose is a largely debated theme within engineering design community, whose discussion is mainly focused on the following aspects:

- the dimensions to be considered for the estimation of creativity, or in other terms, the factors that participate to determine the extent of creativity (as better explained in Section 2);
- the procedures to be followed (e.g. Verhagen et al., 2011) and the metrics to be employed (e.g. Oman et al., 2013) for the measurement;
- the reliability and repeatability of people's evaluation of creativity, which is often deemed to be a very subjective perception (Caroff and Besançon, 2008).

With reference to the last issue, the literature provides sufficient evidences about the differences among designers in facing design tasks, hence leading to very diversified outputs. Innovativeness is considerably affected also by education curricula, which do not necessarily enforce design creativity (Genco et al., 2012). In addition, insightful studies reveal that experience plays a key role in shaping individuals' aptitudes to design processes, resulting in a major attention paid to the consideration of initial requirements (Cross, 2004; Atman et al., 2007). It can be inferred that a different emphasis on problem scoping vs. ideas generation should lead to varying criteria to select alternative design results in terms of effectiveness of the solution, originality, suitability to the field of use and so on. In other terms, the perception of creativity could be extensively affected by the supposed priorities (at the individual level) in performing design tasks and, in the last instance, by the degree of experience in design of the subjects requested to express creativity evaluations. In this sense, Casakin and Kreitler (2008 provide preliminary indications by observing discrepancies among creativity assessments expressed by teachers and students of design courses in the field of architecture.

It follows that the remarked differences among the various players of the design community suggest to cautiously interpret the results of individual assessments of creativity, as well as their use as a litmus test for verifying any hypothesis about measurement procedures or metrics. Furthermore, it has to be highlighted that the literature documents experiments aimed at measuring creativity of both ideas emerged during the design process (e.g. Shah et al., 2000) and products or services, whose commercial destiny is known (e.g. Borgianni et al., 2012). Little or no information is available about the potential influence of the awareness of designers about the success of design outputs in evaluating creativity.

In such framework, the paper illustrates an experiment aimed at bridging various problematic aspects concerning the assessment of creativity and its exploitation within innovation initiatives. The performed test attempts to answer the following research questions:

- do expert and novice designers disagree also about the creativity of well-known innovative products?
- does the situation differ considering separately acknowledged successful innovations and market failures?

The answers to the above queries provide preliminary indications about the convergence of designers on the degree of products creativity and, consequently, on the reliability of individual assessments. The state of the art, presented in Section 2, better elucidates the importance of undertaking research programs devoted to a major understanding of the treated subjects. Section 3 presents a preliminary experiment designed to extract evaluations of experts and freshmen with regards to the creativity of commercial products. The emerging outcomes are presented in Section 4, and discussed in Section 5, that includes planned future activities to overcome the current limitations of the experiment.

2 RELATED ART

The evaluation of several aspects of creativity within engineering design is a relatively recent branch

of research, influenced by seminal works carried out mainly in the field of cognitive psychology (Thompson and Lordan, 1999; Howard et al., 2008). It follows that initial studies have addressed the evaluation of designers' creativity, with or without the employment of specific tools and methodologies.
The work documented by Shah et al. (2000) first underlined the importance of assessing the creativity of ideas generated during a design task. Since then, many hypotheses have been formulated to reveal the factors that mainly contribute to the evaluation or assessment of products creativity. Chiu and Salustri (2010) reviewed previous experiences in academics addressed at measuring the creativity of design projects, revealing that novelty and usefulness (sometimes referred as utility or value) are the most agreed assets of creativity. A not marginal amount of contributions include however additional components of creativity, more diffusedly appropriateness and surprise. The former concerns the property of products to fit practical purposes (Runco and Charles, 1993), in terms of being correctly designed for the conventional domain they belong to. The latter refers to exceeding the boundary of expectedness (Macedo and Cardoso, 2001), by proposing objects that get people's attention by deviating from the line of products evolution dictated by seeded knowledge.
However, the two dimensions suggested by Chiu and Salustri (2010) undoubtedly represent a shared starting point for any model or criterion to assess creativity. This thought is somehow supported by the discussion reported in (Goldschmidt and Tatsa, 2005), which remarks how the precondition of effective design ideas is their goodness (thus their utility or capability to fulfill design requirements) and that they can be considered creative, when they additionally show elements of originality or newness.
Still focusing on novelty and usefulness, Sarkar and Chakrabarti (2011) report a proposal to quantitatively assess creativity of products. Its main strength stands in the employment of objective metrics, which do not require evaluations of individuals. The suggested model exploits a previously developed causal functional model to characterize the degree of novelty, i.e. SAPPhIRE (Srinivasan and Chakrabarti, 2009) and multiple criteria to estimate usefulness including the urgency of the need to be satisfied, the potential quantity of people interested in the product, the duration of the employment of the system or of the provided benefits. The main limitation of the proposal stands in a verification of the illustrated metrics by comparing creativity measures of known products with evaluations performed by experts in design, hence resorting to subjective judgments.
On the other hand, creativity assessments entrusted to individuals are the most diffused in the literature (Oman et al., 2013). Despite the noticeable amount of contributions treating experiments of creativity estimations, carried out by design experts, no standard methodology for performing the task and evaluating the results has been established. However, Horn and Salvendy (2006) remark how the Consensual Assessment Technique (CAT) and the Creative Product Semantic Scale (CPSS) are the most common references for measuring product creativity. The former (Amabile, 1982) assumes that creativity exists only if a sample of experts agrees on its presence, which is tested through a 5-point Likert scale. The latter (Besemer and O'Quin, 1986) explores a plurality of creativity components by asking individuals their perception with respect to several semantic pairs (e.g. unknown vs. familiar within the judgment of novelty).

2.1 Open issues investigated in the research

As an evidence, also when employing objective metrics, there are no references to compare creativity assessments with, but individual estimations. If subjective evaluations represent a starting point for identifying criteria to measure design creativity, convergence of judgments, at least among experts, should be ensured to guarantee sufficient reliability of any study in the field. Whereas satisfying agreement among individuals owning expertise in design and innovation would be met, as in some experiences documented by Horn and Salvendy (2006), a further issue of investigation regards the tendency of less skilled subjects in expressing similar creativity evaluations.
Moreover, if both usefulness (in terms of goodness of the design task) and novelty represent fundamental ingredients for creativity, it might be inferred that original, but unsuccessful, products cannot be regarded of being creative. However, according to authors' vision, such statement has not been sufficiently proven.
The scope of the research is to provide a major comprehension of the phenomena that lead experts and novice designers to evaluate products as creative. The understanding of such mechanisms can motivate or reject the employment of subjective evaluations (provided by narrow or large arenas of designers)

as a reference for assessments extracted by more objective criteria. According to the above general goal, the methodological objective of the paper is the fine-tuning of an experiment devoted to highlight the differences in creativity evaluations within separate groups of experts in design and engineering students, as well between the two samples. Particular attention is dedicated to underline potential dissimilarities when successful products and acknowledged market failures are evaluated apart.

3 DESIGN OF THE TEST

The experiment is based on the administration of a specific questionnaire to a sample constituted by individuals belonging to different domains of product design and innovation and having different skills. Hereinafter, a detailed description of the performed test, as well as of the criteria used to assess the results, is presented in order to clarify the followed approach and to provide other scholars, interested in deepening the treated subject, a framework to replicate such a kind of experiment.

3.1 Structure of the questionnaires

The administered questionnaire is anonymous, so as to allow respondents to answer freely and to avoid biases in the evaluation of the results. It consists of two main sections.

The first section requires to provide some personal details such as age, gender and profession/job. These data allow classifying the collected answers according to experience and competencies of the participants about product design and innovation.

The second part of the questionnaire asks to perform an assessment of the creativity level referred to a set of 20 products, as summarized in Table 1. The set deliberately includes 10 products that have observed a widespread diffusion in the market, while the residual of the sample is constituted by goods unsuccessfully commercialized. Nevertheless, the questionnaire does not report information about which items resulted successful (those identified by the numbers 2, 6, 7, 9, 10, 13, 14, 15, 18, 19 in Table 1) or failing; indeed, the products are listed according to alphabetical order. The individuation of artifacts has been performed by extensively exploiting sources devoted to provide major understanding about the reasons behind thriving (such as Kim and Mauborgne, 2005) or failing (e.g. Haig, 2010) in the market. The success or, conversely, the commercial flop of the listed products is however witnessed by a plurality of literature sources. The authors decided to differentiate the considered goods in terms of technical domains and delivered functionalities, in order to avoid any potential distortion of the creativity assessment process ascribable to an excessive focusing of the experiment on specific industries or technologies. Moreover, besides the commercial name, each product reports the year of launch in the market, a picture and a brief description (maximum 15 words) providing essential information about its main features in a neutral form, without the use of expressions potentially influencing the judgment of assessors. Such a way of presentation is supposed to facilitate the contextualization of the products within the historical period of their launch. The objective is making the task more independent from the age of the participants and evaluating the extent of creativity by comparing the treated products with older artifacts. To the purpose, the respondents were free to gather any information about the products and technologies available in the reference periods indicated in the questionnaires, previously administered by email.

3.1.1 Building of a tailored creativity assessment tool

After a short introduction on the case studies, through the essential descriptions of Table 1, the respondents were requested to evaluate the degree of creativity for each product. The most diffused techniques were considered unsuitable for the specific purpose of the work. The CAT, besides poorly exploited within the design field (Jeffries, 2012) and noticeably time-consuming, cannot be used because of the required involvement of unskilled subjects for creativity evaluation purposes. On the other hand, the experiences witnessing the use of CPSS are restricted to the judgment of few products (Horn and Salvendy, 2006) and show limited convergence among assessors. Nevertheless, the authors decided to employ a Likert-type scale, which represents a reference instrument for the wide majority of activities aiming at measuring creativity in the design context (Oman et al., 2013).

Table 1.Products submitted to the creativity assessment through the questionnaire.

#	Product	Picture	Year	Description
1	Amphicar		1961	The first car-ship for civil use. It may be driven as car or ship in the same way
2	Apple iPod		2001	MP3 reader capable to host a great volume of data, easy to use and with unmistakable aesthetic features
3	Apple Lisa		1983	First personal computer with a graphic interface and a mouse.
4	Apple Newton		1993	Personal Digital Assistant and Tablet with specific software, fax and email applications
5	BMW C1		2000	Scooter with anti-crash chassis and safety belts
6	Canon Copiers		1973	Small copiers for desk
7	Croc's shoes		2002	Slipper and plastic shoes, cheaper, with an unmistakable design
8	Dodge La Femme		1955	Car with accessories and colors dedicated to women
9	Geox		1995	Shoes with perspiration sole
10	Nintendo Wii		2006	Console for videogame that allow to play with a joy-pad, owning improved interaction
11	Nokia N-Gage		2003	Mobile phone and portable console that allows playing with other people through Bluetooth connection
12	Pepsi Crystal		1992	Transparent Cola aimed at transmitting a sense of purity of the drink
13	Pfizer Viagra		1998	Drug for erectile dysfunction
14	Polo Ralph Lauren		1967	Elegant and classical T-shirt made of high quality material.
15	Red Bull		≈ 1995	Energy drink
16	Reynolds' Smokeless Cigarettes		1988	Smokeless cigarettes, employing cartridges made of tobacco that are heated but not burnt
17	Sony Betamax		1975	Video recorder with magnetic tape for domestic use
18	Sony Walkman		1979	Portable stereo with earphones for playback through magnetic tape
19	Swatch		1983	Cheaper watch with interchangeable band.
20	Telecom Italia Fido		1998	Mobile phone that can be used within the city, that has an improved range with respect to a domestic cordless

In order to avoid the so called “neutral point” in scales, typical of uneven quantity of alternative answers, the participants were asked to choose among four options, as follows: “Definitely Not”, “More Not than Yes”, “More Yes than Not”, “Definitely Yes”. The respondents were explicitly invited to make the assessment by following an intuitive, free and personal process, thus without using specific criteria, but considering the degree of creativity according to the context in which the products appeared for the first time.

3.2 Sample of participants

The set of participants, constituted by a sample of convenience, has been built by joining two different groups of individuals. The former comprises experts in the field of design and innovation, while the latter is constituted by novice designers.

More in detail, the group of experts is constituted by 43 participants belonging to the following categories: professors of machine design coming from different Italian Universities, senior designers having a marked sensitivity towards product and process innovation, professionals dealing with innovation processes in industry and technology transfer. The group of unskilled subjects comprises 21 volunteers coming from the course of Product Design and Development that is held during the first year of the MS program in Mechanical Engineering of Florence University.

Notwithstanding the marked difference in the size of the two groups, the samples present a supposed sufficient number of individuals, besides being homogeneous and clearly defined.

3.3 Criteria to assess the results of the pool of respondents

The judgments expressed by the participants about the degree of creativity of each assessed product are converted into scores to allow their handling for statistical analyses. Such operation is required to elaborate representative values of creativity for both the sub-samples of the whole set of respondents to be compared against. The conversion is performed by applying the metrics shown in Table 2, which have been proposed in a detailed report for evaluation purposes of the teaching in Italian Universities (Chiandotto and Gola, 2000), representing a reference point for this task (Rampichini et al., 2004).

Table 2. Metrics for the conversion into scores of the judgments about the creativity of the products collected through the questionnaire.

Judgment	Score
Definitely Not	2
More Not than Yes	5
More Yes than Not	7
Definitely Yes	10

3.4 Analysis of the outcomes

The emerging data are thus analyzed in order to fulfill the research questions. In a first instance, the objective is to investigate the differences in the creativity judgments between the two samples of participants. The authors propose to perform such examination by:

1. calculating for each product the mean and the standard deviation of creativity evaluation in charge of experts and freshmen;
2. comparing the means of both the subsets with a suitable test: two-tailed t-test is proposed to the purpose, because the variances of the distributions are not known a priori and it is not known whether the two sub-samples effectively belong to two different populations; the test swivels on the following hypotheses:
    a) *H0 (null hypothesis): the means of the populations are identical, i.e. skilled and novice designers provide the same creativity evaluations, hence experience provide no influence in judging products creativity;*
    b) *H1(alternative hypothesis): the means of the populations differ, i.e. skilled and novice designers provide different creativity evaluations, hence experience plays a role in judging products creativity*
3. evaluating for each product, the significance level of the test through p-values, then establishing for each case study whether experience affects the estimation of creativity.

In a second instance, the goal is to investigate the potential influence of the commercial success/failure of artifacts in creativity evaluation. It is proposed to achieve the objective by:

1. building the contingency table putting into relationship the success of the product and the agreement between experts and freshmen with respect to judgments of creativity;
2. computing the probability of the independence between the samples by a χ^2 test, hence estimating to which extent experience determines differences in creativity assessments.

4 RESULTS OF THE EXPERIMENT

For the sake of brevity, the whole set of answers to the questionnaires is not included in the paper. The assessments of creativity are indeed replaced by statistical descriptors (mean and standard deviation, as illustrated in Table 3), which are obtained through the metrics shown in Table 2. According to the above description of the experiment, distinct values are reported for the groups of expert and novice designers; the actual sample size of the respondents is added in the Table, since some designers did not provide creativity evaluations for all the investigated products. In the last column of the Table, the p-values are reported for the hypothesis test described in Section 3.4. Tests leading to the acceptance of the test (p-value <=0,05, as a common rule of thumb), i.e. such that experience has not affected the estimation of creativity, are marked with an asterisk in Table 3.

Table 3. Assessments of creativity performed by experienced designers and engineering students. Descriptive statistics except for the last column, collecting the p-value of the Hypothesis test mentioned in Section 3.4.

	Experts in design and innovation			Freshmen in design and innovation			p-value
Product	**Mean**	**Std. deviation**	**Sample size**	**Mean**	**Std. deviation**	**Sample size**	
Amphicar	6,64	2,61	42	6,67	2,42	21	0,044*
Apple iPod	8,19	1,91	43	7,62	1,88	21	0,872
Apple Lisa	9,26	1,48	42	9,19	1,54	21	0,172
Apple Newton	7,81	1,93	43	7,43	2,04	21	0,694
BMW C1	6,70	1,74	43	6,86	2,01	21	0,341
Canon Copiers	7,21	2,21	43	7,19	2,36	21	0,037*
Croc's shoes	5,79	2,72	43	4,23	2,30	21	0,999
Dodge La Femme	3,79	2,14	43	4,05	2,46	21	0,478
Geox	7,16	2,67	43	6,38	2,16	21	0,931
Nintendo Wii	8,53	1,93	43	8,00	1,90	21	0,847
Nokia N-Gage	5,37	1,86	43	5,38	1,91	21	0,019*
Pepsi Crystal	3,98	1,99	43	4,24	1,95	21	0,512
Pfizer Viagra	6,05	2,79	43	7,48	2,18	21	0,998
Polo Ralph Lauren	3,74	2,26	43	4,71	2,31	21	0,981
Red Bull	5,95	2,53	43	6,05	2,36	21	0,177
Reynolds' Smokeless Cigarettes	6,26	2,13	43	6,19	2,50	21	0,129
Sony Betamax	8,47	1,93	43	8,00	1,90	21	0,788
Sony Walkman	9,18	1,44	43	8,62	1,69	21	0,911
Swatch	7,95	2,31	43	6,71	2,55	21	0,966
Telecom Italia Fido	5,53	2,33	43	6,57	2,99	21	0,982

As clear from the Table, the null hypothesis is rejected in the great majority of the cases: only Amphicar, Canon Copiers and Nokia N-Gage represent exceptions. The emerging outcomes guide to infer that experience in design influences the assessment of creativity for many marketed products. It is interesting to notice that not necessarily freshmen assigned a higher creativity score to the proposed products: indeed experts ranked ten products as more creative, while seven received a higher score by the engineering students.

As clarified before, the second research question invites to verify whether the situation differs by separately considering market successes and failures. It is straightforward to observe that in both samples of case studies a marginal part of examples does not lead to the rejection of the hypothesis

supposing the missing influence of experience in judging design creativity. Indeed, the formulated *H0* is accepted for 1 success (out of 10) and 2 market flops (still out of 10, thanks to the way the experiment has been planned). The next task of the investigation requires thus to verify whether the above different behaviors result statistically significant. By considering a global number of 3 cases for which a convergence in evaluating design creativity is met, an equal distribution of these examples between successes and failures is represented in Table 4, while Table 5 reports the real observed outcomes.

Table 4. Expected outcomes of the test, by imposing a global number of 3 cases meeting the null hypothesis

	Successful products	Unsuccessful products	*Total*
Products having a significant difference in creativity evaluation	1,5	1,5	3
Products not having a significant difference in creativity evaluation	8,5	8,5	17
Total	10	10	20

Table 5.Observed outcomes of the test

	Successful products	Unsuccessful products	*Total*
Products having a significant difference in creativity evaluation	1	2	3
Products not having a significant difference in creativity evaluation	9	8	17
Total	10	10	20

Under the hypothesis that the two above distributions are non-correlated, the resulting probability of such supposition assumes the value of 0,531 according to the χ^2 test. Hence, it might be inferred that it is quite probable that the incongruence in creativity evaluations between experts and freshmen is not influenced by the market success of the products. Undoubtedly, such claim cannot be anyway assessed with great confidence.

5 DISCUSSIONS AND CONCLUSIONS

The paper illustrates an experiment about creativity assessment through questionnaires (to which 64 people participated), revealing how experience in the field of design and innovation plays a not negligible role in performing such kind of evaluations. Therefore, the outcomes of the test support, within product design, what has been already remarked by Casakin and Kreitler (2008) in the field of architecture. Just in 3 cases out of 20 it is possible to state with sufficient statistical evidence that creativity judgments of expert and less skilled designers converge. Hence, if evaluations provided by skilled designers could be considered as a reference for measuring creativity, the results of the experiment discourage the employment of inexperienced subjects for the same purposes.

Given the proven discrepancy between samples of skilled innovators and students, it should be better researched if evaluations provided by experts can represent a reliable benchmark for creativity assessment. Data about standard deviations reported in Table 3 can represent a starting point for stimulating a discussion in this sense. The extension of the sample of respondents is however required to this aim in order to determine the effective variability of experts' measures of creativity.

Also due to the very narrow quantity of akin creativity evaluations between samples of expert and novice designers, the χ^2 test carried out to verify the twisting effect of market success vs. failure provided poor information. Thus, in order to answer the second research question defined in the Introduction, the experiment should be reorganized. The redesign of the test could favorably include considerations about the knowledge of the designer with respect to the investigated products, as well as the awareness of their commercial results. Such measure would allow considering whether information about market success could represent a bias in creativity evaluations and correctly evaluating the effect played by the age of the respondents and markedly the supposed greater awareness of experts with respect to old products. At the current state of the research, it seems that this phenomenon does not take place, by observing that an unsuccessful case (i.e. Apple Lisa) is the product attributed of the maximum extent of creativity by both experts and students.

Among the planned future activities, beyond the administration of the test to a major quantity of designers and the intention to include questions about the knowledge of the surveyed products, the authors intend to evaluate the extent of a certain amount of factors in determining creativity judgments. The demographic information already provided through the experiment and a further segmentation of experts' sample (already distinguished among professors, senior designers and professionals in the field of innovation) can represent the basis fora set of explanatory factors to be analyzed. The authors commit to achieve information about the impact played by any parameter pertaining both respondents and the investigated product (e.g. its market success or failure) through statistical regressions. Such techniques are expected to determine whether the investigated factors concur to increase or diminish creativity evaluations and to which extent the same parameters result statistically significant. By employing ordinal regressions, the problem could be avoided of transforming qualitative assessments into quantitative measures, which relentlessly introduces a bias into the analysis of the results.

REFERENCES

Amabile, T.M. (1982) Social psychology of creativity: A consensual assessment technique. *Journal of Personality & Social Psychology*, Vol. 43, No. 5, pp. 997-1013.

Atman, C.J., Adams, R.S., Cardella, M.E., Turns, J., Mosborg, S., and Saleem, J. (2007) Engineering Design Processes: A Comparison of Students and Expert Practitioners. *Journal of Engineering Education*, Vol. 96, No. 4., pp. 359-379.

Besemer, S. P., and O' Quin, K. (1986) Analyzing creative products: Refinement and test of a judging instrument. *Journal of Creative Behavior*, Vol. 20, No. 2, pp. 115-126.

Borgianni, Y., Cascini, G., and Rotini, F. (2012) A Proposal of Metrics to Assess the Creativity of Designed Services. In Duffy, A., Nagai, Y., and Taura, T. (eds.) *Design Creativity 2012*, Glasgow, University of Strathclyde Engineering/The Design Society.

Casakin, H., and Kreitler, S. (2005) The nature of creativity in design. In Gero, J. S., and Bonnardel, N. (eds.) Studying Designers'05, Aix-en-Provence, Key Centre of Design Computing and Cognition.

Caroff, X., and Besançon, M. (2008) Variability of creativity judgments. *Learning and Individual Differences*, Vol. 18, No. 4, pp. 367-371.

Chiandotto, B., and Gola, M. (2000) Questionario di base da utilizzare per l'attuazione di un programma per la valutazione della didattica da parte degli studenti (Basic questionnaire to be employed for putting into practice a program for teaching evaluation by the students), final report to the Italian Ministry for University and Scientific and Technological Research, in Italian.

Chiu, I, and Salustri, F. A. (2010) Evaluating Design Project Creativity in Engineering Design Courses. *Proceedings of the 1st CEEA Conference*, Kingston, Canadian Engineering Education Association.

Chulvi, V., Mulet, E., Chakrabarti, A., Lòpez-Mesa, B., and Gonzàlez-Cruz, C. (2012) Comparison of the degree of creativity in the design outcomes using different design methods. *Journal of Engineering Design*, Vol. 23, No. 4, pp. 241-269.

Cross, N.(2004) Expertise in design: an overview. *Design Studies*, Vol. 25, No. 5., pp. 427-441.

Genco, N., Hölttä-Otto, K., Seepersad, C. C. (2012) An Experimental Investigation of the Innovation Capabilities of Undergraduate Engineering Students. *Journal of Engineering Education*, Vol. 101, No. 1, pp. 60-81.

Gero, J. S., Jiang, H., and Williams, B. (2012) Design Cognition Differences When Using Structured and Unstructured Concept Generation Creativity Techniques. In Duffy, A., Nagai, Y., and Taura, T. (eds.) *Design Creativity 2012*, Glasgow, University of Strathclyde Engineering/The Design Society.

Goldschmidt, G., and Tatsa, D. (2005) How good are good ideas? Correlates of design creativity. *Design Studies*, Vol. 26, No. 6, pp.593-611.

Haig, M. (2005) *Brand failures: the truth about the 100 biggest branding mistakes of all time*, London, Kogan Page.

Horn, D., and Salvendy, G. (2006) Consumer-Based Assessment of Product Creativity: A Review and Reappraisal. *Human Factors and Ergonomics in Manufacturing*, Vol. 16, No. 2, pp. 155-175.

Howard, T.J., Culley, S.J., and Dekonick, E. (2008) Describing the creative design process by the integration of engineering design and cognitive psychology literature. *Design Studies*, Vol. 29, No. 2, pp. 160-180.

Jeffries, K. K. (2012) Amabile‘s Consensual Assessment Technique: Why Has It Not Been Used more in Design Creativity Research?. In Duffy, A., Nagai, Y., and Taura, T. (eds.) *Design Creativity 2012*, Glasgow, University of Strathclyde Engineering/The Design Society.
Kim, W. C., and Mauborgne, R. (2005) *Blue ocean strategy: how to create uncontested market space and make competition irrelevant*, Boston, Harvard Business School Press.
Macedo, L., and Cardoso, A. (2001) Using Surprise to Create Products that Get the Attention of Other Agents. In Cañamero, L. D. (ed.) *Emotional and Intelligent II: The Tangled Knot of Social Cognition, Technical Reports of the AAAI 2001 Fall Symposium*, North Falmouth, MA, AAAI Press.
Oman, S. K., Tumer, I. Y., Wood, K., and Seepersad, C. (2013) A comparison of creativity and innovation metrics and sample validation through in-class design projects. *Research in Engineering Design*, Vol. 24, No. 1, pp. 65-92.
Rampichini, C., Grilli, L., and Petrucci, A. (2004) Analysis of university course evaluations: from descriptive measures to multilevel models. *Statistical Methods and Applications*, Vol. 13, No. 3, pp. 357-373.
Runco, M. A., and Charles, R. E. (1993) Judgments of originality and appropriateness as predictors of creativity. *Personality and Individual Differences*, Vol. 15, No. 5, pp. 573–546.
Sarkar, P., and Chakrabarti, A. (2011) Assessing design creativity. *Design Studies*, Vol. 32, No. 4, pp. 348-383.
Shah, J. J., Kulkarni, S.V., and Vargas-Hernandez, N. (2000) Evaluation of idea generation methods for conceptual design: Effectiveness metrics and design of experiments. *Journal of Mechanical Design*, Vol. 122, No. 4, pp. 377–384.
Srinivasan, V., and Chakrabarti, A. (2009) SAPPhIRE - An approach to analysis and synthesis. In Norell Bergendahl, M., Grimheden, M., Leifer, M. L., Skogstad, P. And Lindemann, U. (eds.), *Proceedings of the 17th International Conference on Engineering Design (ICED'09), Vol. 2*, Stanford, The Design Society.
Thompson, G., and Lordan, M. (1999) Review of creativity principles applied to engineering design. *Proceedings of the Institution of Mechanical Engineers, Part E: Journal of Process Mechanical Engineering*, Vol. 213, No. 1, pp. 17-31.
Verhaegen, P.-A., Peeters, J., Vandevenne, D., Dewulf S., and Duflou, J. (2011) Effectiveness of the PAnDA ideation tool. *Procedia Engineering*, Vol. 9, pp. 63-76.