Brainstorming versus creative design reasoning

Akın Kazakçı, Thomas Gillier, Gerald Piat, Armand Hatchuel

Brainstorming versus creative design reasoning: a theory-driven experimental investigation of novelty, feasibility and value of ideas

Akin Osman KAZAKCI
Ecole Nationale Supérieure des Mines de Paris, Paris, France
Thomas GILLIER*
Grenoble Ecole de Management, Grenoble, France

Gerald PIAT
R&D Division, Electricité de France (EDF), Paris, France
Armand HATCHUEL
Ecole Nationale Supérieure des Mines de Paris, Paris, France

In industrial settings, brainstorming is seen as an effective technique for creativity in innovation processes. However, bulk of research on brainstorming is based on an oversimplified view of the creativity process. Participants are seen as idea generators and the process aims at maximizing the quantity of ideas produced, and the evaluation occurs post-process based on some originality and feasibility criteria. Design theories can help enrich this simplistic process model. The present study reports an experimental investigation of creativity process within the context of real-life design ideation task. Results lead to the rejection of the classical 'quantity breeds quality' hypothesis. Rather, we observe that successful groups are the ones who produce a few original propositions that hold great value for users while looking for ways to make those propositions feasible.

1 Research problem: creativity beyond idea generation

Considering current economical and social challenges at the global scale, innovation is widely considered as a vital component of today's industry.

It is generally assumed that creativity is a main component in innovation processes. Based on these premises, an astounding amount of research has been undertaken on creativity techniques and processes since 1950's [1-8]. A technique that has received particular attention and acknowledgement is brainstorming [3]. Brainstorming is considered as an effective technique by professionals in overcoming particularly hard tasks that require creative insights [9]. There is a large room for debates on this issue considering the conflicting results from research [1, 2]

Since the pioneering work of Osborn [3], a significant amount of work has been undertaken to decipher this creative process and to provide *control parameters* and various *extensions* [4-6]. Our main claim of about the existing research work is that they are working with an oversimplified (and implicit) process model for creativity process. Not only they lack rich definitions of creative behavior, thus reducing the phenomenon mainly to a simple idea generation phase, they also implicitly assume that the evaluation process can be dissected from the generation phase. Under such hypotheses, naturally, brainstorming research tends to evaluate output of a creative process by the quantity of ideas produced, making the assumption that quantity breed quality (Osborn, 1963).

Compared to brainstorming research, design research has focused in understanding the *dynamics* of creativity within design processes [10-14]. Arguably, design processes are most significant creative processes driving economic and social innovations [15]. As design research progresses towards maturity, several formal models and theories has been produced describing design reasoning leading to creativity [16-24]. The paper defends the thesis that such models would enable better-designed experiments to understand creative processes and would allow new predictions to be tested.

To demonstrate this concept, the current work reports an on-going exploratory investigation of a real-world design experiment. Ten teams of three people have been given an innovative design task (the design of an Antarctica-like museum) within the context of an industrial setting. Their reasoning processes have been analyzed using verbal protocol analysis. Two hypotheses, regarding the quantity of properties produced and the quantity of novel properties produced have been tested – both of which have been rejected, as predicted based on a particular model of design, namely, C-K theory. After reviewing brainstorming research in section 2, we introduce a creative design process description based on C-K design theory in section 3. Section 4 discusses methodology and data collection. Section 5 presents results, before concluding in section 6.

2 Brainstorming research

2.1 Overview

Focus of brainstorming research has been to increase number of ideas generated by groups [25, 26]. Initial research wanted to show that group creativity is better then same number of individuals (called nominal groups) working separately. Experiments have shown that nominal groups are much productive. A second episode of research on brainstorming tried thus to identify causes for this phenomenon. A number of factors such as the production loss, blocking, free riding, fear of evaluation have been identified [25]. The natural tendency followed was to rectify this situation by eliminating these factors. Numerous techniques such brainwriting and anonymous evaluation has been proposed.

2.2 Evaluation of a creativity in brainstorming literature

Brainstorming research has accepted, often unquestioningly, a major hypothesis initially introduced by Osborn himself:

> (QBQ) Quantity breed quality hypothesis: "It is almost axiomatic that quantity breeds quality in ideation. Logic and mathematics are on the side of the truth that the more ideas we produce, the more likely we are to think up some that are good. "(Osborn 1963, p. 131).

In the literature, this hypothesis has manifested itself as an evaluation process that is based essentially on the *number* of ideas. Literature often discusses criteria introduced early on [27, 28] such as fluency, originality, and flexibility. Fluency is a quantity of the number of non-redundant ideas generated during the process Originality represents the uncommonness of an idea, given a problem, sometimes measured as the relative rarity of an idea given the pool of ideas produced by several participants for the same task [28, 29]. Flexibility measures the ability to produce ideas that belong to uncommon categories of solutions given to the task at hand [27, 28]. Research indicates that these dimensions can be co-related but it is not always the case. Some studies indeed report that fluency and originality are co-related (i.e. that QBQ is valid [1] but this is not always the case [30] reports that previous work where both flexibility (number of categories used) and within-category fluency were present, no systematic correlation between the two was found [31] indicating these measures might be independent. Although the literature consistently argues that total production is correlated with high-quality output

2.3 Feasibility vs. originality

Despite the previously mentioned criteria, in experiments, the production of an ideation process is usually measured based on a dichotomy of originality and feasibility. Rietschel *et al*. [25] state that "*individuals or groups that generate the most ideas, also generate the highest number of good ideas (with good ideas usually defined as ideas that score high on both originality and feasibility)*. As Rietschel and colleagues [32] puts it there is a general agreement among brainstorming researchers that quality in creativity tasks is some combination of *originality* and feasibility. Measuring idea quality by having external judges rating the originality and the feasibility of the generated ideas, it is claimed that QBQ has been verified [1, 32].

2.4 Critics of idea evaluation in brainstorming

Contrary to mainstream brainstorming research, concerns have been raised in various research projects regarding the QBQ hypothesis and the evaluation process used in ideation tasks. After a comprehensive review of brainstorming research, [26] concludes that evidence on QBQ is not conclusive, or even conflicting.

Quantity is not the issue in real-life innovation processes
Williams and Sternberg [33] instructed teams of participants to produce a best idea ratter than as many ideas as possible. Contrary to usual work comparing nominal groups and interactive groups, they found out that teams were more successful in generating an overall superior idea than individuals. Their instruction clearly violated the QBQ hypothesis, while approaching a more realistic setting. As we shall argue later on, the objective of finding a best idea enforces the evaluation of ideas while they are being generated, rather than a pure generative process, the activity becomes a *reasoned* process. Rowatt et al. [34] demonstrated with a series of experiments that people have indeed a preference for *quality over quantity*. They interpreted their finding as a reason for revising brainstorming instructions to downplay the importance of quantity and emphasize the importance of quality. Paulus [35] interpret this preference as fear for novelty and judging somewhat unfortunate that people tend to focus more on usefulness and validity.

Phased separation between idea generation and evaluation

Rietschel *et al*. [25] acknowledge that idea generation is only a part of the creative process and not a goal in itself. However, their diagnostic for situating the dynamics of idea generation within the broader process share the same flaw many other brainstorming researchers: they make an implicit assumption that idea generation and evaluation are two separate processes: *A question largely unaddressed by brainstorming research thus is how exactly the production of ideas contributes to creative solutions or innovations after the idea generation stage*. [25]

Paulus [35] argues that a lot of potential that is built up during the idea generation phase is wasted because people do not know how to evaluate ideas. Citing Rietzschel and colleagues [36], he states that idea generation is still an elusive issue for brainstorming research. *Typically it is suggested that generation and selection should be two separate phases, but evidence thus far is not clear on this issue. Alternatively, it might be best to mix short idea generation sessions with evaluation sessions. This will be a puzzle for future research to resolve*. As we shall see, this is far from being a puzzle but an elementary property of creative processes in design research.

3 Design theory and models: rich descriptions of creativity

Contrary to brainstorming research having strong relationship with experimental psychology literature, design research combined several methodological approaches to understand complex and real-life creativity situations that are design processes. One of the methodologies that have been used is theoretical modeling ([16-24]. In the present work, we are going to consider a particular theory of design, namely the C-K theory [20]– and contrast it with the brainstorming underlying process model.

3.1 An overview of value, feasibility and originality in C-K theory

C-K theory describes design based on the interaction between two spaces. In the knowledge space, propositions about the known world exist. They are either true or false. In the concept space, there exist definitions for classes of objects. The theory claims that, in innovative design processes, those definitions are undecidable: it cannot be stated that corresponding objects can or cannot exist - until the end of design. In C-K theory creative propositions (and thus, originality) stems from a particular type of operation a designer applies in order to elaborate an undecidable propositions. This kind of operation, called expansive partitions or conceptual expansions, adds to a concept an unusual or unknown property in order to build new and unprecedented object definitions (e.g. an Antarctica-like mobile museum). A second type of operation, called restrictive partitions, adds to a concept a usual and known property (e.g. a history museum). Restrictive partitions does not necessarily create easy-to-realize object definitions, since creative design already starts with a conceptual expansion – and the known property added by the restriction stills need to be connected with the unusual properties within the concept (e.g. an Antarctica-like mobile history museum). Although not explicitly stated by the theory, in practice it is often assumed that concepts hold value. Value is constructed progressively, extended if necessary, using partitioning (expansion or restrictions). Thus, conceptual expansions create or add new values to a type of object (e.g. an Antarctica-like mobile museum provides opportunity for children to be immersed into an Antarctica-like universe, not far from where they live).

There is no guarantee that a creative concept is going to be validated (i.e. acknowledged as feasible): concepts are validated if, during the process, knowledge warranting the existence of such an object is produced, activated or found. Expansive partitions, by definition, introduce originality into a design, but makes more difficult to validate a concept (i.e. to make it feasible).

Thus, in design, separation between the originality and the feasibility of a concept is not a problem; it is an opportunity to achieve both: a concept is unfeasible by nature (it contains creative expansions that makes it unfeasible). It is by the process of design that those concepts are made to exist. Thus, unfeasible yet original ideas cannot be discarded as invaluable or of poor quality. They allow the exploration of both value and feasibility.

3.2 Theory-driven predictions about brainstorming hypotheses

Based on the dynamics described by C-K theory, it is possible to produce and predict the outcome for a myriad of hypotheses about the dynamics of a design process and its impact on performance in terms of feasibility (F), originality (O) and value (V) of ideas. To demonstrate the approach, we shall introduce a second hypothesis. As mentioned earlier, the conceptual separation between idea generation and evaluation is a simplistic model that is bound to induce interpretation problems regarding the creative processes. Paulus [35] states: *People tend to focus more on usefulness and validity. There is clearly fear of novelty.* This seems to assume that originality should prime over feasibility and value of ideas. We might formulate this hypothesis as quantity of novel (unknown) properties breeds quality:

> (QNOQ) Higher the number of novel properties, higher the team performance

C-K theories and more generally, work on design research goes clearly against this hypothesis. For instance, Girorta *et al.* [37] argue that organizations prefer a single outstanding idea to several good ideas. In design processes, the objective of finding a best idea enforces the evaluation of ideas while they are being generated – but the evaluation steps do not necessarily eliminates ideas. As C-K theory points out, new *knowledge* allowing further elaboration and improvement of those ideas are generated as well. As can be expected from such process, designers tend to generate sufficient number of original ideas that embodies value but they also need to make sure the feasibility of those ideas by elaboration. Brainstorming research, seeing the creative process as the generation of a sequence of unconnected ideas, misses this crucial insight. Thus, based on C-K theory's description of design process, QNOQ should prove to be false.

4 Research design and methodology

4.1 Overview

In order to address our research problem, we realized verbal protocol analyses of several design teams' creativity sessions. Our groups were composed of three designers with experiences in R&D activities. The creativity sessions were part of an innovative real-life project, supported by a French cross industry innovation partnership: In 2012, MINATEC IDEAs Laboratory® decided to organize a series of innovative workshops involving multiple external professional designers. The sessions took place in a laboratory setting. Data were recorded and analyzed following the principles of verbal protocol analysis [38, 39].

Our research question was to explore how design teams ideate when they need to elaborate breakthrough concepts. Contrary to research that mostly used case-study methodologies for understanding group creativity at the firm level, verbal protocol analysis permits tracking the team cognitive process in a more fine-grained level. More details about our protocol are being presented elsewhere [40]

Design Computing and Cognition DCC'14. J.S. Gero (ed),
pp. xx-yy.

4.2 Data collection

4.2.1 Research protocol of the innovative sessions

A. Participants and formation of design teams

The 30 participants were all either engineering designers or industrial designers with an average of 12 years of professional experience in R&D and innovation. Participants came from various industrial sectors and held different positions at the moment of the experiments: 10 were innovation managers, 11 engineering designers, 4 industrial designers, 3 industrial buyers and 2 B-to-B marketeers. Not any participant had previous experience in the design of museum or in the creation of important public social events.

B. Presentation of the design brief

The ten design teams were assigned to elaborate a breakthrough museum concept that gives the visitor an immersive experience in an Antarctica-like world. The participants were asked to elaborate both the form and the functions of the museum, the architectural aspects and the possible museum activities. Additional requirements were the following:

- The museum aims to make people sensible to the imperative of protecting Antarctica.
- The museum is mobile – it is a touring museum that could be deployed everywhere in the world whatever the conditions.
- Practical and easy to install and transport.
- Eco-friendly as much as possible (ecological materials; energy harvesting solutions…)
- The museum size is approximately 3600 m^2.

The topic of museum is simple and easily appropriable for the subjects. Because museums are a commonplace, all participants have some experience with them. The design brief is sufficiently open-ended; it offers the opportunity to investigate how design teams think in very different ways about different domains. Also, the object to be designed, by its very definition, is outside the scope of any known instance of its category.

C. Organization of the design sessions

At the start of each session, the design task was given to the design team. They were informed that they had one hour and thirty minutes to formulate one single concept of innovative museum. The design teams were asked to summarize their final proposal on an A3 sheet of paper with sketches, user scenarios, texts and motto. All experiments were launched in the same large room; table, white papers/papers, pencils were provided. Participants did not have any access to external documents (no computer, no internet connections, no books, no phone…).

At the end of each session, the participants were given 10 minutes to present their final product concept to the organizers. In order to cover all the aspects of the design proposals and to provide reliable and comparable qualitative data between the ten designs, a semi-structured guide was used to question them about their designs. All the experiments and interviews were video and audio recorded.

4.2.2. Research protocol of the ratings of the final designs

A panel of 14 professionals assessed the ten final designs – 6 of them were experts in museum (2 directors of museum, 2 curators, 1 public programmer and 1 exhibition designer) and 8 of them were specialized in the organization of public events (1 director and 7 project managers). Judges were asked to rate the 10 final designs with 3 criteria according to a five level Likert scale:

- **novelty** of the product concept compared to the existing museums – all kinds of museum could be considered (from traditional museums to innovative ones (planetarium, 3D-relief movies…));
- **feasibility** of the product concept in terms of how it can be implemented – economical and technical feasibility were ered ;
- **value** for users; value of the product concept for potential visitors.

All judges were blind to the research. In order to increase the reliability of judges, the evaluation process was divided into 3 steps according to an adapted version of Delphi technique. In the first step, the rating criteria were presented and discussed by the judges. This aims to reduce the possible differences in the interpretation of the 3 criteria. Then, the final sketches of the 10 final design concepts were presented and each judge was asked to rate independently. Finally, the results of ratings were discussed in an unstructured form by the judges who could modify (or not) their initial ratings. Inter-rater reliability between the judges was measured. We use the inter-judge agreement reliability formula, called the Proportional Reduction Loss (PRL) of Rust and Cooil's [41]. PRL is proved to be of superior performance of several other reliability approaches for qualitative or quantitative data since it does not allow reliability to appear inflated. The inter-rater reliability was calculated for each of the criteria during both the first and second round of evaluation: *PRL-originality (2*nd*=.74;1*st*=.71);*

Design Computing and Cognition DCC'14. J.S. Gero (ed),
pp. xx-yy. © Springer 2014

PRL- feasibility (2^{nd}=.74;1^{st}=.71); PRL- value (2^{nd}=0.70;1^{st}= .71). All statistics show PRL ≥ 0.7 and so the internal consistency is largely acceptable.

4.3 Data analysis

4.3.1. Coding and analysis of the verbal protocols: process – measures metrics

Verbal protocol analysis has been used to study the conversations and the different interactions between the participants. Teams activities were recorded and transcribed. Information regarding the identification of the speaker and the specific actions and reactions of participants (drawing, handling objects, jokes, laughing, mime…) were integrated into the transcripts. In total, the ten transcripts covering 15 hours of audiovisual data were used in the present analysis. The degree of verbalization varies between teams, number of words ranges from 12.500 to 23.000.

A. Coding of design properties: classic versus novel properties

Utterances from design teams have been used to capture various design properties they dealt with. All the design properties were then either coded as "known" or "unknown" by two experts. Known properties refer to classical properties of museums (i.e. restrictive partitions). Unknown properties are properties imagined by designers that do not correspond to any traditional properties of museums. During coding, a design property was considered to be known if it could be trivially observed in existing or past museums. Several documents about various types of museums (scientific, artistic, cultural, or historical) were considered – a specific attention was given to the website of the major museums dedicated to Antarctica, the museum of Artic and Antarctic in St Petersburg (Russia). In the case where an uttered design property was not usually encountered in museum, it was coded as an unknown property. For instance, the design properties "a museum that is a zeppelin" or "museum with ice footsteps sound effects" fall in this category.

B. Inter-rater reliability measures: Identification of properties

Two of the authors completed the coding of all the data independently and without communicating. Design properties were identified by highlighting each time they appeared in the transcribed data; the two coders individually named each design property, with its timestamp and the name of the designers. For coding, the software Atlas Ti (version 6.2, www.atlasti.com) was used. Afterwards, the two coders analyze and compare their coding segments side-by-side in order to validate (or not) their identification. The Percentage Agreement (PA) between the two coders for identifying the design properties was calculated for each team. While this method does not exclude agreements occurred by chance, it is simple and appropriate for exploratory conditions. For each transcript, the PA was calculated. The naming of design properties was discussed by the two coders. All the differences in the interpretations were resolved between the two coders. The overall average PA is 0.77. Such level shows that the identification of design properties can thus be considered fully acceptable and satisfactory (PA > .7). Inter-rater reliability between the two authors was measured for the full set of design properties. The PRL for each team reach a satisfactory level (> 0,7), in overall, the PRL was 0.75.

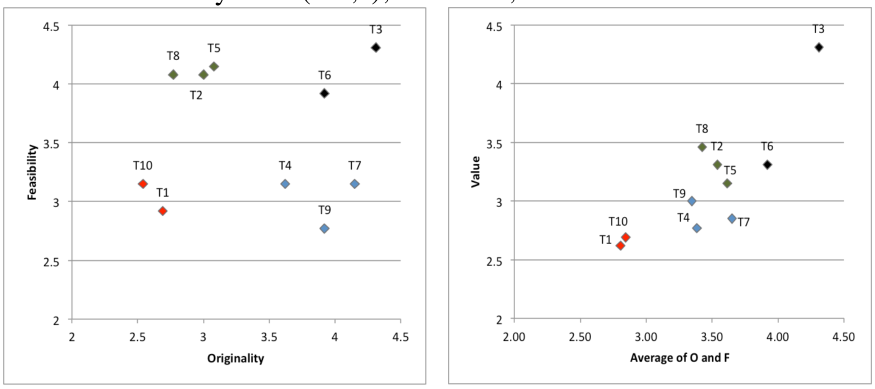

Fig1. Comparison of team performances, on the left, feasibility vs. originality criteria; on the right, value vs. average of originality and feasibility

Table 1 Team performances on O, F and V

Team	Originality (O)	Feasibility (F)	Value (V)
Team #1	2.69	2.92	2.62
Team #2	3	4.08	3.31
Team #3	4.31	4.31	4.31
Team #4	3.62	3.15	2.77
Team #5	3.08	4.15	3.15
Team #6	3.92	3.92	3.31
Team #7	4.15	3.15	2.85
Team #8	2.77	4.08	3.46
Team #9	3.92	2.77	3.00
Team #10	2.54	3.15	2.69

5 Results

5.1 An overview of team performances

Table 1 and Figure 1 give insights into the nature of team performances. Here, we can observe that there are four distinct categories of team performances. Teams #3 and #6 are the best teams overall. In particular, they have performed well both on criteria O and F. Teams #1 and #10 are the worst performing teams on both criteria compared to others. The remaining teams let appear two contrasted profiles.

Table 2 Number of properties, unknown properties, known properties and their rations per teams

Team	# of Ps	# of UPs	# of KPs	UPs/KPs Ratio	UPs/Ps Ratio
Team #1	64	31	33	0.94	0.48
Team #2	68	39	29	1.34	0.57
Team #3	60	39	21	1.86	0.65
Team #4	57	36	21	1.71	0.63
Team #5	95	75	20	3.75	0.79
Team #6	125	79	46	1.72	0.63
Team #7	88	56	32	1.75	0.64
Team #8	95	51	44	1.16	0.54
Team #9	85	59	26	2.27	0.69
Team #10	85	47	38	1.24	0.55

Teams #2, #8 and #5 are teams that have performed reasonably well (worst than the bests, better than the worsts). Their performance profiles are similar; they are all rather better on the feasibility criterion compared to originality. By contrast, teams #9, #7 and #4 (again worst than the bests, better than the worsts) did perform better on originality than on feasibility. When we consider the performances of these four categories of teams on the value criteria (Fig1b), we see that the bests and the worsts remain stable. Between the two remaining groups, the feasibility oriented group has performed slightly better with respect to value criteria than the originality oriented group. Different hypotheses that we are going to test give us a better understanding of these performances.

QBQ: Bigger the #Ps, better the team performance.

In total, the ten design teams have generated 822 design properties (S.D. = 20.61). Table 1 gives the evaluation of experts for each group and table 2 gives three metrics (number of properties #Ps, number of unknown properties #UPs and number of known properties #KPs) and. There are several interesting consequences that can be drawn from this information. First, we see that the relationship between the number of (non-reduntant) properties and team performance is questionable. The top three generating teams are team #6, #5 and #8. Team #6 generated the highest number of design properties (125 design properties), well above the average (>82,2). Its performance appears to be well on originality and value criteria (3rd in both). On feasibility criteria it's ranked 5th. Its average ranks are rather good (2nd on both O+F and O+F+V).

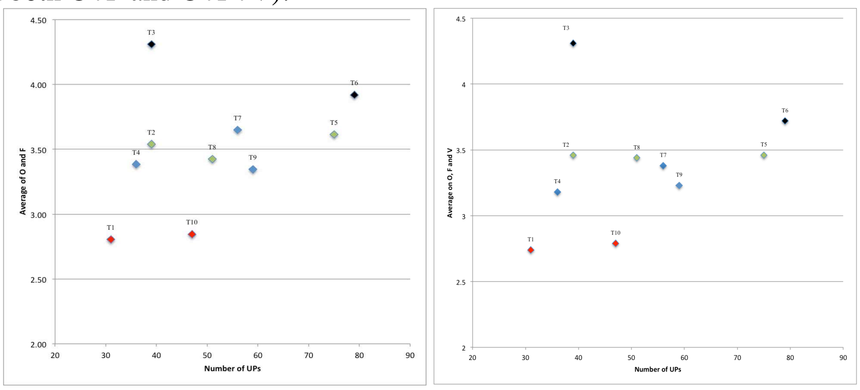

Fig2. Comparison of team performances based on #UPs, on the left, feasibility vs. originality criteria; on the right, value vs. average of originality and feasibility

This might tempt us to think that H1 is correct: the group that has generated the most ideas is a good group overall. Looking at the top 2nd and 3rd most generating teams (respectively, team #5 and #8), however, gives a clearer picture. Both have generated 95 ideas (again, well above the average). Team #5 has an unsteady performance profile (6th on O, 2nd on F and 5th on V). This team has achieved particularly well on the feasibility criterion; however, its overall rank on O+F is just above average (4th). Team #8 has also an irregular performance (8th on O, 4nd on F and 2th on V). This team has achieved particularly bad on the originality, but surprisingly, it has one of the best ranks in value criterion (2nd). However, both of its overall ranks are average at best (6th on O+F and 5th on O+F+V). Looking to the sporadic performances of these top-generating teams, QBQ cannot be confirmed clearly.

The least generative teams shed further light on QBQ. In fact, one of the two least generative teams, namely team #2, is the team with the highest performance on all criteria, thus, arguably the best team. *They have only generated 60 ideas*. This goes clearly against the QBQ. Let us remark that the opposite of the QBQ (the less the number of ideas, the better the quality) cannot be stated either. Team #1 and #4, two of the three least generative teams have arguably bad performances. Team #1, which generated 64 properties, has ranked 9th or 10th on all criteria. Team #4 has ranked 5th, 7th and 8th on O, F and V. In sum, according to our dataset, there is no relationship between the number of generated Ps and the quality of Ps. Thus, we reject QBQ.

QNOQ: Bigger the #UPs, better the team performance.
Let us consider QNOQ and the number of unknown properties (UPs) generated by the teams. The team that generated the most UPs is team #6, which is the same as the team that generated most Ps. It is indeed one of the best teams (2nd both on O+F and O+F+V)[1]. The second best group on UPs ranking is team #5, which is also a good team in overall performance (4th on O+F and 3rd on O+F+V). Going to the opposite side, the two teams that generated the least UPs (team #4 and team #1) have poor performances overall (team #1 being the least good team in the experiment, and team #4 being 7th on O+F and 8th on O+F+V). Considering these, it is tempting to suspect a positive relationship between the number of UPs generated and the team performance. However, if we consider the third most and third least generative teams, the argument breaks down. In fact, once again team #1 who is the uncontested best on performance criteria is one of the least UPs generating team (ranked 3rd). Likewise, one of the worst teams, team #9 (8th on O+F and 7th on O+F+V), has ranked 3rd according to the number of UPs generated, again going against the QNOQ.

Much as it is hard to confirm QNOQ with the available data, some variations of QNOQ deserve some attention. Let us consider:

- QNOQ': if a group's generation of UPs is below some threshold with respect to the average, it will have a bad performance.
- QNOQ": if a group's generation of UPs is above some threshold with respect to the average, it will have a good performance.

Although not conclusive from the current data, there is some evidence supporting these hypotheses. If we consider that the average UPs production is 51.20, we see that both most and least productive teams are beyond

[1] It should be noted, however, that their score on feasibility is average (5th) and thus, their being second in overall is mostly due to the compensatory nature of the averages (e.g. a bad score can be compensated by a good one, and vice versa).

the edges of standard deviation (±16.25). Team #5 and #6 have both produced above 67.45 UPs and both have good performance. Team #4 and #1 have both produced below 40 UPs (slightly above μ-σ) have performed poorly. Considering the size of our sample and the performance of team #3 (which has also produced slightly above μ-σ but has best performance), this is still inconclusive. However, there is reason to check for QNOQ' and QNOQ" in future iterations.

Another way to consider QNOQ is from the performance charts (Fig 2). Here, we see that the top-performing group (Team #3 and #6) has a significant difference in terms of number of UPs produced. Likewise, the difference between the worst performing teams is considerable (Teams # 1 and #10). This variability in terms of number of UPs is also valid for the middle performance groups. Both the feasibility and the originality groups span widely along the number of UPs axis. In the light of these observations, it cannot be stated that QNOQ is a direct predictor of team performance.

Conclusions

The full set of results from our experiment is still being processed. The major result presented in the current work is the invalidity of QBQ hypothesis of brainstorming in our experimental setup based on a design creativity task. We also demonstrated that the number of original (unknown) propositions generated by a team is not predictor of high performance. An inspiring observation is that, groups that have achieved high on feasibility (from among the middle performance groups) are also better on value than the groups that have achieved better on originality. A possible explanation is that the groups that tend to have a few original UPs and that managed to ascertain feasibility of those outperformed groups that have only searched for originality. In our data, it would seem that a more balanced UPs and KPs generation allows higher originality. All the teams that ranked close to average (Teams #3, #7, #6, #9 and #4) have high originality scores. By contrast, teams that have ranked well on feasibility criterion are either a low or high ratio of UPs/KPs. It would seem, that having high originality with a balanced unknown generation strategy is not sufficient to obtain a high value or feasibility score. In future work, we shall analyze whether successful teams identify and work on the valuable and original ideas to guarantee their feasibility. This will require going beyond quantity-based metrics by introducing and analyzing process based metrics and their relationship with output proposals.

References

1. Diehl, M. and W. Stroebe, *Productivity loss in brainstorming groups: Toward the solution of a riddle*. Journal of Personality and Social Psychology, 1987. **53**(3): p. 497.
2. Stroebe, W., M. Diehl, and G. Abakoumkin, *The illusion of group effectivity*. Personality and Social Psychology Bulletin, 1992. **18**(5): p. 643-650.
3. Osborn, A.F., *Applied Imagination, principles and procedures of creative thinking*. 1953.
4. Paulus, P.B. and V.R. Brown, *Enhancing ideational creativity in groups*. Group creativity: Innovation through collaboration, 2003: p. 110-136.
5. Paulus, P.B. and B.A. Nijstad, *Group creativity: Innovation through collaboration*2003: Oxford University Press.
6. Mannix, E.A., J.A. Goncalo, and M.A. Neale, *Creativity in groups*. Vol. 12. 2009: Emerald Group Publishing.
7. Summers, I. and D.E. White, *Creativity techniques: Toward improvement of the decision process*. Academy of management Review, 1976. **1**(2): p. 99-108.
8. De Bono, E., *The mechanism of mind*1969: Wiley Online Library.
9. Hurt, F., *Better brainstorming*. Training and Development, 1994. **48**(11): p. 57-59.
10. Goldschmidt, G., *The dialectics of sketching*. Creativity research journal, 1991. **4**(2): p. 123-143.
11. Gero, J.S. and M.L. Maher, *Modeling Creativity and Knowledge-Based Creative Design*. 1973.
12. Logan, B. and T. Smithers, *Creativity and Design as Exploration*, in *Modelling creation and knowledge-based design*, J.S.a.M. Gero, M. L., Editor 1992.
13. Goel, A.K., *Design, Analogy, and Creativity*. IEEE Expert, 1997. **12**(3): p. 62-70.
14. Dorst, K. and N. Cross, *Creativity in the design process: co-evolution of problem–solution*. Design Studies, 2001. **22**(5): p. 425-437.
15. Le Masson, P., B. Weil, and A. Hatchuel, *Strategic management of innovation and design*2010: Cambridge University Press.
16. Yoshikawa, H., ed. *General design theory and a CAD system*. Man-Machine Communication in CAD/CAM, ed. T. Sata and E. Warman1981, North Holland Publishing.
17. Takeda, H., et al., *Modeling design processes*. AI Magazine, 1990. **11**(4): p. 37-48.
18. Salustri, F.A. *A formal theory for knowledge-based product model representation*. in *2nd IFIP Knowledge Intensive CAD workshop*. 1996.
19. Hendriks, L. and A. Kazakci *Imagining Future Knowledge, Logic and Interactive Rationality, The Institute for Logic, Language and Computation, University of Amsterdam, .* 2010.
20. Hatchuel, A. and B. Weil, *C-K design theory: an advanced formulation*. Research in Engineering Design, 2009. **19**(4): p. 181-192.

21. Maimon, O. and D. Braha, *A Mathematical Theory of Design*. International Journal of General Systems, 1996. **27**(4-5): p. 275-318.
22. Braha, D. and Y. Reich, *Topological structures for modelling engineering design processes*. Research in Engineering Design, 2003. **14**: p. 185-199.
23. Gero, J.S., *Design prototypes: a knowledge representation schema for design*. AI Magazine, 1990. **11**: p. 26-36.
24. Kazakci, A., *On the imaginative constructivist nature of design: a theoretical approach*. Research in Engineering Design, April 2013, Volume 24, Issue 2, pp 127-145, 2013.
25. Rietzschel, E.F., K.W. De Dreu, and B.A. Nijstad, *What are we talking about, when we talk about creativity? Group creativity as a multifaceted, multistage phenomenon*." in *Creativity in groups*2009.
26. Reinig, B.A. and R.O. Briggs, *On the relationship between idea-quantity and idea-quality during ideation*. Group Decision and Negotiation, 2008. **17**(5): p. 403-420.
27. Torrance, E.P., *Guiding creative talent*. 1962.
28. Guilford, J., *Creativity: Yesterday, today and tomorrow*. The Journal of Creative Behavior, 1967. **1**(1): p. 3-14.
29. Amabile, T.M., *Creativity and innovation in organizations*1996: Harvard Business School.
30. Förster, J., R.S. Friedman, and N. Liberman, *Temporal construal effects on abstract and concrete thinking: consequences for insight and creative cognition*. Journal of Personality and Social Psychology, 2004. **87**(2): p. 177.
31. Nijstad, B.A., W. Stroebe, and H.F. Lodewijkx, *Production blocking and idea generation: Does blocking interfere with cognitive processes?* Journal of Experimental Social Psychology, 2003. **39**(6): p. 531-548.
32. Rietzschel, E.F., B.A. Nijstad, and W. Stroebe, *Relative accessibility of domain knowledge and creativity: The effects of knowledge activation on the quantity and originality of generated ideas*. Journal of Experimental Social Psychology, 2007. **43**(6): p. 933-946.
33. Williams, W.M. and R.J. Sternberg, *Group intelligence: why some groups are better than others*. Intelligence, 1988. **12**(4): p. 351-377.
34. Rowatt, W.C., et al., *Perceptions of brainstorming in groups: The quality over quantity hypothesis*. The Journal of Creative Behavior, 1997. **31**(2): p. 131-150.
35. *Interview: Paul Paulus on Group Creativity*. Creativity and Innovation Management, 2013. **22**(1): p. 96-99.
36. Rietzschel, E.F., B.A. Nijstad, and W. Stroebe, *The selection of creative ideas after individual idea generation: Choosing between creativity and impact*. British Journal of Psychology, 2010. **101**(1): p. 47-68.
37. Girorta, K., C. Terwiesch, and K.T. Ulrich, *Idea Generation and the Quality of the Best Idea by girorta*. Working paper, 2008.
38. Dunbar, K. and I. Blanchette, *The in vivo/in vitro approach to cognition: the case of analogy*. TRENDS in Cognitive Sciences, 2001. **5**(8): p. 334-339.
39. Ericsson, K.A. and H.A. Simon, *Protocol analysis*1985: MIT press.

40. (authors to be inserted) *The elaboration of breakthrough concepts by design teams*. Submitted, 2013. **2013**.
41. Rust, R.T. and B. Cooil, *Reliability measures for qualitative data: Theory and implications*. Journal of Marketing Research, 1994: p. 1-14.