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

EMPIRICAL EVALUATION OF FLEXIBLE DESIGN CONCEPT GENERATION PROCEDURES: A STUDY IN EMERGENCY SERVICES

Michel-Alexandre CARDIN, Howard KA-HO YUE, Yixin JIANG, Yinghan DENG, Deepak SANTHANAKRISHNAN
National University of Singapore, Singapore

ABSTRACT

This paper presents the results of an empirical study of concept generation procedures enabling flexibility in engineering systems design. Evaluation of two educational training procedures (analogies vs. explicit) and two ideation procedures (free undirected brainstorming vs. prompting) was done. The procedures aim to improve quantitative lifecycle performance, while providing users with satisfaction with the process. Controlled experiments involved ninety participants working on a design problem in emergency services. Results suggest that combining explicit training on flexibility and free undirected brainstorming was best to improve lifecycle performance, measured as average response time and net present cost of infrastructure. No statistically significant effect was measured when comparing procedures against one another based on lifecycle performance, suggesting that any procedure could be used. Analogies combined with free undirected brainstorming led to better process satisfaction for users. The results give insights on the true effects of concept generation procedures, considering quantitative performance impacts, as well as qualitative user impressions.

Keywords: collaborative design, creativity, design methods, flexibility in engineering design, real options

Contact:
Dr. Michel-Alexandre Cardin
National University of Singapore
Department of Industrial and Systems Engineering
Singapore
117576
Singapore
macardin@nus.edu.sg

1 INTRODUCTION

This paper is motivated by a need to understand the effects of procedures supporting the design of complex engineering systems for flexibility. Designing a complex system for communications, emergency services, or transportation is a challenging task (ESD, 2011). Uncertainty in demographics, environments, markets, regulations, and technology inevitably affect the performance of such systems, since they operate over long-time horizons, and require significant capital investments. Many procedures have been developed to deal with uncertainty systematically by exploiting ideas of flexibility (de Neufville and Scholtes, 2011). There has not been an equivalent body of work focused on understanding their effects on flexible design concept generation activities. Many studies describe a novel design procedure, and show how to use it via application in a single case study – e.g. (Mikaelian et al., 2012). This approach does not provide sufficient grounds for thorough procedure evaluation, and to measure any statistical effect (Frey and Dym, 2006). Every system is different and is exposed to different exogenous factors affecting the responses, thereby affecting the results.

This paper presents the results of a study evaluating concept generation procedures to help designers generate flexibility in engineering systems. Participants were involved in experiments related to the design, deployment, and operations of an emergency services system in a hypothetical city in Asia. This system was chosen because it is a critical infrastructure to any modern city, and appropriate given the background of the sample population. Two procedures were evaluated in a controlled setting to: free undirected brainstorming and prompting. Two procedures were used to train participants: analogies and conceptual training. The main and interaction effects of the procedures are reported on quantitative lifecycle performance, and qualitative user impression of satisfaction with the processes.

2 RELATED WORK

2.1 Flexibility in Engineering Systems Design

Flexibility enables a system to change in the face of uncertainty (Fricke and Schulz, 2005). It is associated to the concept of real options, providing the “right, but not the obligation, to change a project in the face of uncertainty” (Trigeorgis, 1996). Real options exist “on” a project, involving higher-level managerial decisions like abandoning, deferring until favorable market conditions, expanding/contracting/reducing capacity, deploying capacity over time, switching inputs/outputs, grow by investing in R&D, and/or mixing the above (Trigeorgis, 1996). Real options “in” a project are technical engineering and design components enabling flexibility in operations (Wang and de Neufville, 2005). While the literature on real options focuses on the economic valuation of flexibility, very little work exists on how to generate and enable such flexibility in engineering design. Many studies have shown that flexibility and adaptability can bring lifecycle performance improvements ranging between 10% and 30% compared to standard design and evaluation approaches (de Neufville and Scholtes, 2011, Engel et al., 2012). Design for flexibility, however, is a difficult process that requires assistance from the engineering design community. It can improve expected performance by affecting the distribution of possible outcomes, as opposed to optimizing the design to a deterministic point forecast. It protects from downside risks (i.e. like buying insurance), while positioning the system to capitalize on upside opportunities (i.e. like buying a stock option to capture more profits). One example in real estate is the ability to expand a building vertically, as done by Blue Cross Blue Shield in Chicago (Guma et al., 2009). The strategy was to “build small first, then expand *when needed*”. This strategy reduced exposure to losses because less capital was required upfront. It also gave access to upside opportunities under favorable market conditions because more offices could be built, and hiring personnel would help generate more profits. This strategy was enabled carefully in the engineering design in the 1990s (e.g. larger elevator shafts, stronger structure), and exercised a few years later to see the second phase completed in 2011.

2.2 Flexible Design Concept Generation and Experimental Evaluation

Flexibility generation involves 1) *generating* concepts in response to major uncertainty sources, and 2) *identifying* areas in the design to enable flexibility concretely and in future operations. Trigeorgis (1996) provided a set of generic real option strategies that can stimulate early concept generation. Suh et al. (2007) combined Change Propagation Analysis (CPA) and Design Structure Matrix (DSM) to identify areas and design variables suitable for flexibility. Mikaelian et al. (2012) suggested a systematic approach based on logical multi-domain matrix (MDM), and enterprise version of a DSM. Many studies have evaluated concept and idea generation procedures in an experimental setting. Kurtoglu et al. (2009) evaluated an online design library procedure integrating artificial intelligence principles. M Yang (2009) studied correlations between concept quantity and quality for brainstorming, morphology charts, and sketching procedures. Others have studied the effects of education and pedagogy on creative design activities. Eppinger et al. (1990) studied how an interdisciplinary classroom environment affects product design and development.

2.3 Research Gaps and Contributions

This overview suggests that there is not much work on developing and evaluating simple and intuitive concept generation procedures for flexibility. Much work on flexibility generation requires developing a DSM first, which is non-trivial and can be time-consuming. Besides the study by Cardin et al. (2012) focusing on a real estate problem in the United States, no other study has focused on this particular issue. There is a need to determine whether the effects observed for different procedures hold when used in other contexts, for different design problems, and by other sample populations. Generalizations based on qualitative arguments may not suffice. This paper presents the results of a controlled user study where similar (except for one) concept generation procedures are evaluated as in Cardin et al. (2012), although using a different design problem, and sample population.

3 DESIGN PROCEDURES

Table 1 summarizes the four concept generation procedures evaluated in this study. Most material used in experiments (e.g. lecture slides, survey, etc.) is available online as indicated in (Cardin et al., 2012). The analogy lecture and design problem description are available upon request.

Table 1. Setup for 2 x 2 Design Of Experiment (DOE)

Educational Training on Flexibility (*E*)	**Ideation Mechanism (*I*)**	
	Brainstorming (–1)	**Prompting (+1)**
Explicit (+1)	Treatment 1	Treatment 2
Analogy (–1)	Treatment 3	Treatment 4

3.1 Educational Training (*E*)

A short conceptual training program was devised to help participants generate flexibility (E = +1, referred as *explicit* training). The treatment was a short 15-20 minutes lecture on flexibility describing generic sources of uncertainty affecting performance, why flexibility can improve such performance, and why it must be considered early in the design process. The lecture also provided real-world applications of flexibility principles in the aerospace and oil industries. The other training procedure consisted of a set of examples flexible systems (E = –1, referred as *analogy*). Learning by examples is widely used in education, motivating this choice (Knoll and Horton, 2010). Also, analogies have never been studied experimentally to train designers on flexibility. Example real-world systems exhibiting flexibility properties (or a lack thereof) were presented to help participants identify features they could reproduce in the design problem during experiments.

3.2 Ideation Mechanism (*I*)

There is a wide range of procedures to support the conceptual design process (Shah et al., 2000, Spitas, 2011). Here, two procedures imposing very different levels of structure were studied. A *prompting* procedure was used to scaffold the thought process for flexibility systematically, as captured by level I = +1. Example prompts were “what are the major sources of uncertainty affecting the future performance of this system?”, “what flexible strategies would enable the system to change and adapt if the uncertainty scenarios you just discussed occur during operations?”, “how should you prepare, engineer, and design this system to enable the flexibilities just discussed?”, or “how should you manage and decide when it is appropriate to use, or exercise, the flexibilities in this system?”, supported by general real option strategies, and generic examples. Free undirected *brainstorming* (I = –1) was used as a simple and intuitive approach to stimulate creativity (Osborn, 1957). It provided a good point of comparison as it is widely used in U.S. industry and academia (Yang, 2007). It is easy to teach to participants, only involving a few principles like encouraging concept quantity, welcoming unusual ideas, avoiding criticism, and combining ideas or improving existing ones. All ideation mechanisms were supported by Group Support System (GSS) technology to minimize the impact of productivity loss (Bostrom and Nagasundaram, 1998). GSS is defined here as “socio-technical systems blending software, hardware, meeting procedures, and facilitation support to support a group engaged in intellectual collaborative work.” (de Vreede et al., 2003)

4 METHODOLOGY

4.1 Participants

Ninety participants distributed among twenty-four teams were recruited from professional masters and doctoral programs in engineering systems, design, and management at a top institution in South-East Asia. They were recruited via class and electronic email announcements for voluntary participation in experiments on flexibility in engineering design. Most participants were mature graduate students with training in engineering, science, and/or management, and with a wide range of industry experience. There were six replicates (or teams) for treatments 1, 2, and 4, and seven for treatment 3. Six control groups used brainstorming in both sessions, and without any form of training.

4.2 Design Problem

The design problem focused on the deployment of fire stations to provide timely response to fire-related emergency calls in a hypothetical city in Asia. Participants were asked to brainstorm about possible deployment and operations strategy for such emergency services. Demand for emergency services was assumed to be subject to many uncertainty sources like demographics changes, call pattern variations during day and night, construction of new population and industrial estates (SCDF, 2012, Ong et al., 2009). This uncertainty was captured via uncertain and fluctuating fire rates. Participants were given a map of the city, divided into five sectors, and average travel time estimates between each sector. They were given the current rate of fires in each sector, but told that the rates would change over time. They were warned that a factory would be built sometime in the future, with the corresponding probability of being built in each sector. This factory represented a major system disturbance although it was possible it would never be built. Characteristics of three different types of fire stations were given: basic, enhanced, and privatized. Construction and operational costs, together with the capabilities of each type was explained. Participants were asked to generate engineering systems design concepts that would minimize 1) average response time, and 2) lifecycle cost, including capital investment, and operations cost (the quantitative performance metrics). Qualitative descriptors (e.g. high, med, low) were used instead of numerical values to avoid leading players into addressing this problem as an optimization problem, rather than a concept generation exercise.

The benchmark design was presented as the typical solution when optimizing emergency services allocation subject to deterministic conditions. An enhanced facility was located in the southern sector with a basic station in the northern sector. This design was selected because the factory was most likely to be built in the South, and the North sector had the best traveling times to other sectors. No privatized station was used in the benchmark to discourage public reliance on private services.

The problem was devised to provide enough room for creativity, while being constrained enough to be tractable in the concept evaluation phase. Participants could decide what kind of fire stations to deploy, as they each had different cost characteristics and capabilities. These decisions taken in real life would naturally impact the more detailed aspects of engineering design. For example, a basic station would be easier to build, require less mechanical and operational resources, but could only respond to minor fires. For an extra cost, enhanced stations could be built to handle both minor and major fires, an example of switching flexibility. Alternatively, privatized stations, which have the same capabilities as an enhanced station, could be deployed at lower capital cost, but higher operational cost. Participants could decide when and where to deploy fire stations, allowing different flexible deployment strategies, and benefiting from the time value of money by deferring capital costs.

4.3 Experiments

4.3.1 Session Structure

At the beginning of each session, the moderator welcomed participants and described the design problem. Then followed a short training on free undirected brainstorming and how to use GSS technology. The task was assigned in session 1 to brainstorm for 25 minutes and suggest alternative design, deployment, and operations plans improving lifecycle performance compared to the benchmark design. Participants had 5 minutes to vote on design concept quality, using a 1 (low) to 10 (high) Likert scale. Quality scores were used in coding analysis to discriminate between two seemingly opposite concepts. Participants repeated the exercise in session 2 for 25 minutes under one of the treatments in Table 1 (except the control group), then followed by another 5 minutes for voting. A post-experimental debrief explained the purpose of the study. Demographics information and user impressions were collected using a validated questionnaire (Briggs et al., 2006).

4.3.2 Control Conditions

Each experiment was structured following a pretest-posttest design to enhance signal-to-noise response. All treatments were done using two rooms, so that two treatments were held simultaneously, and four could be done successively. Providing the same content in all activities (e.g. introductions, training, task definitions, lectures, etc.) controlled for information variability. Teams of three participants were assigned in each experiment to control for team size effects – although a few last-minute cancellations forced the arrangement of six teams of two, and five teams of four. The same time was allocated for each activity to control for possible effects on concept quantity and quality.

4.4 Data Collection

4.4.1 Online GSS Interface and Raw Data Description

GroupSystems' ThinkTank® online software was used as GSS technology. It is an easy-to-use interface enabling participants to type in real-time descriptions of their design solutions. The moderator posted the ideation task of improving performance compared to the benchmark design. Each participant described different solutions, which were displayed to others to stimulate creativity and engage discussions. Each participant could reply, comment, or append new ideas to a thread. Raw data consisted of written descriptions of the systems design concepts in a Word® document produced by the GSS software, quality scores for each idea/concept, and online survey results collected using LimeSurvey®. This data was analogous to the raw data obtained after interviewing engineers, managers, and/or decision-makers in case studies (Suh et al., 2007).

4.5 Computer Model

A Matlab® simulation model was developed based on standard operations research techniques (Larson and Odoni, 1981). For each scenario, the location/construction year of the factory was simulated, as well as rates of fires in all sectors. Based on the station type, location, and deployment strategies described by participants, Expected Net Present Cost (*ENPC*) and Expected Average Response Time (*EART*) were calculated as key performance indicators (KPI).

The probability that the factory was built in each sector in each year was captured by variables (p_N, p_S, p_E, p_W, p_C), with $p = 0.25$ each year that the factory would not be built. After being built, the location of the factory could not be changed. The rate of minor and major fires would increase by a factor of 2 following construction, with annual volatility +/– 10% in both minor and major rates of fire. Minor and major rates of fires were assumed to be independent random variables. For instance, the rate of minor fires $\lambda_{N1,t}$ at any time t (subscript 2 for major fires) in the North sector in shown in Equation 1:

$$\lambda_{N1,t} = \lambda_{N1,t-1}\left(1 + dW_t\right)F_{N,t} \tag{1}$$

Here a Wiener process $dW_t \sim U\,(-0.1,\ 0.1)$ was assumed, and $F_{N,t}$ adjusted the rate of fire for the presence of the factory. $F_{N,t} = 2$ when the factory was in the North at time t, and 1 otherwise. All other fire rates were calculated in a similar fashion. Flexibility strategies and management decision rules were implemented to mimic flexible decision-making in light of uncertainty realizations, using logical programming statements (e.g. `IF`, `ELSE`) depending on the nature and complexity of the decision rule. A one-year time-to-build period was assumed between decisions to construct/upgrade a station, and when it was operational. Construction costs were incurred before a fire station became operational. For instance, one common strategy was upgrading. Initially, the configuration could be one basic station in the South, and one enhanced station in the North. The decision rule could be "*if the factory is built in* *the South sector, upgrade the basic fire station to an enhanced fire station.*” Such strategy affected *EART* and *ENPC* positively because it would enable more emergency calls to be served, while deferring upgrading costs to later in the future due to discounting (see Equation 3 below).

ENPC was chosen as a measure of design’s efficiency, based on communications with an emergency services provider (SCDF, 2012), and calculated as shown in Equation 2:

$$ENPC = -ENPV = -E\left[NPV\right] = -\frac{1}{M}\sum_{m=1}^{M} NPV_m \tag{2}$$

In a simulation m, *NPV* was calculated by summing all discounted cash flows incurred over the time horizon T.[1] Since costs C_t were positive and there were no revenue streams, the negative sign below ensured that *NPV* would be negative, and thus *ENPC* always positive. The discount rate r accounted for the time-value of money (Equation 3), while C_t was calculated as the sum of the operating (C_{Ot}), construction/deployment (C_{Ct}), and upgrading costs (C_{Ut}) (Equations 4):

$$NPV = \sum_{t=0}^{T} PV_t = \sum_{t=0}^{T} \frac{-C_t}{\left(1+r\right)^t} \tag{3}$$

$$C_t = C_{Ot} + C_{Ct} + C_{Ut} \tag{4}$$

EART was chosen as a measure of a design’s effectiveness. It was calculated by averaging the average response time for each simulation m and each year in each simulation t (Equation 5):

$$EART = \sum_{m=1}^{M}\sum_{t=1}^{T} ART_{m,t} \tag{5}$$

$ART_{m,t}$ was calculated assuming no queuing in the system (i.e. rate of fire low enough that no calls accumulate). Occurrence of each type of fire in each sector was modeled as an independent Poisson process, with an arrival rate equal to the rate of fire. The independence assumption led to the total arrival rate of fires being the sum of the fire rates in each sector. For example, the total rate of fire in the city $\lambda_{T,t}$ $t > 0$, including rate of minor ($\lambda_{T1,t}$) and major ($\lambda_{T2,t}$) fires was calculated as (Equations 6-8):

$$\lambda_{T,t} = \lambda_{T1,t} + \lambda_{T2,t} \tag{6}$$

$$\lambda_{T1,t} = \lambda_{N1,t} + \lambda_{S1,t} + \lambda_{E1,t} + \lambda_{W1,t} + \lambda_{C1,t} \tag{7}$$

$$\lambda_{T2,t} = \lambda_{N2,t} + \lambda_{S2,t} + \lambda_{E2,t} + \lambda_{W2,t} + \lambda_{C2,t} \tag{8}$$

The proportion of fires of a given type occurring in a sector at time t was equal to the rate of that type of fire in that sector, divided by the total rate of fires in the city $\lambda_{T,\,t}$. The fastest response time was determined for each fire type. It was assumed that the response time was the lowest traveling time required to travel between a fire station and a fire. For instance, if one assumed there were two enhanced fire stations, one in the North and another in the South, it was first determined which of the two stations would deliver the fastest response sector. As an example, t_{NS} was the time to travel from the North to the South. $ART_{m,t}$ was then calculated for time t and simulation m by summing up the product of the proportion of fires, and the fastest response time as shown in Equation 9:

$$\begin{aligned} ART_{m,t} = {} & \left(\frac{\lambda_{N1,t}}{\lambda_{T,t}} + \frac{\lambda_{N2,t}}{\lambda_{T,t}}\right)t_{NN} + \left(\frac{\lambda_{S1,t}}{\lambda_{T,t}} + \frac{\lambda_{S2,t}}{\lambda_{T,t}}\right)t_{SS} + \left(\frac{\lambda_{E1,t}}{\lambda_{T,t}} + \frac{\lambda_{E2,t}}{\lambda_{T,t}}\right)t_{NE} + \dots \\ & \dots + \left(\frac{\lambda_{W1,t}}{\lambda_{T,t}} + \frac{\lambda_{W2,t}}{\lambda_{T,t}}\right)t_{SW} + \left(\frac{\lambda_{C1,t}}{\lambda_{T,t}} + \frac{\lambda_{C2,t}}{\lambda_{T,t}}\right)t_{NC} \end{aligned} \tag{9}$$

---

[1] For convenience, subscript m will be omitted from here on.

4.6 Analysis

4.6.1 Coding Analysis

Two independent treatment-blind coders reviewed each ideation transcript in a randomized order to extract and count complete concepts using a standard coding procedure (Strauss and Corbin, 1990), with 92% average inter-rater agreement. Concepts retained for implementation, evaluation, and statistical analysis, were the ones agreed upon by both reviewers. A design concept was considered *complete* if it contained coherent information about the following elements (using the upgrading example above): a) *uncertainty source* affecting lifecycle performance (e.g. time when factory is built), b) *flexible strategy* to adapt to the above uncertainties in design and operations (e.g. upgrade from basic to enhanced facility), c) conceptual but concrete description of the *flexibility enabler*, considering engineering design, legal, management, and/or financial aspects (e.g. build a basic station with the ability to upgrade), and d) *decision rule*, an "`IF`" statement or "trigger mechanism" based on observations of the uncertainty sources, determining when it is appropriate to exercise the flexibilities (e.g. first build a basic station, and `IF` a factory is built, `THEN` upgrade to an enhanced station).

4.6.2 Dependent Variables

The null hypothesis of no main and interaction effects of factors E and I was tested on the following dependent variables: a) lifecycle performance of flexible design concepts ($ENPC$ and $EART$), and b) subjective impressions of satisfaction with the procedures/processes (PS). A response $\Delta y = y2 - y1$ was measured for each experiment, where $y1$ was the response of interest in session 1 only, and $y2$ in both sessions combined. For example, considering the benchmark design leading to $EART = 10.5$ minutes, if a complete concept from session 1 led to $EART = 10.4$ minutes ($EART1 = 10.5 - 10.4 = 0.1$ minutes), and another concept in session 2 led to $EART = 9.8$ minutes ($EART2 = 0.7$ minutes), then $\Delta EART = 0.7 - 0.1 = 0.6$ minutes. Only the best combinations of flexible design concepts producing the highest $ENPC$ and $EART$ values were considered in $\Delta ENPC$ and $\Delta EART$ measurements.

4.6.3 Survey Analysis

Survey responses were analyzed to measure improvements in ΔPS. Responses recorded the differences in user impressions between sessions 1 and 2, using a discrete 7-scale Likert mechanism. Each construct was evaluated using five or six questions (maximum score 35 or 42). A positive (negative) score meant PS improvement (worsening) from session 1 to 2.

4.6.4 Statistical Analysis

Each response Δy was modeled using a general linear model (Equation 10). Coefficient β_0 approximated the total mean, β_E and β_I modeled the main effects, while β_{EI} modeled the first order two-way interaction between factors E and I. Variable ε accounted for the mean experimental error. Least-square regression was used to calculate the main and interaction effects, and p-values of each coefficient in Excel®. The null hypothesis tested was H_0: $\beta_E = \beta_I = \beta_{EI} = 0$.

$$\Delta y = \beta_0 + \beta_E E + \beta_I I + \beta_{EI} EI + \varepsilon \tag{10}$$

5 RESULTS

5.1 Improvement in Lifecycle Performance ($\Delta ENPC$ and $\Delta EART$)

Overall, twenty-five complete concepts were generated across all experiments: fifteen across all session 1 replicates, and ten more across all session 2 replicates. Figure 1 presents the mean plot graphs for $\Delta ENPC$ (left) and $\Delta EART$ (right), showing no statistically significant main effect for any of the procedures. The apparent interaction effect in $\Delta EART$ was in fact weak ($\beta_{EI} = -0.01$, $p = 0.78$). A combination of brainstorming and explicit training, however, produced better lifecycle cost improvement (i.e. cost savings) on average ($\Delta ENPC$ = \$335,566), and one of the two best response times ($\Delta EART = 0.11$ minutes or ~7 seconds) as compared to the benchmark design.

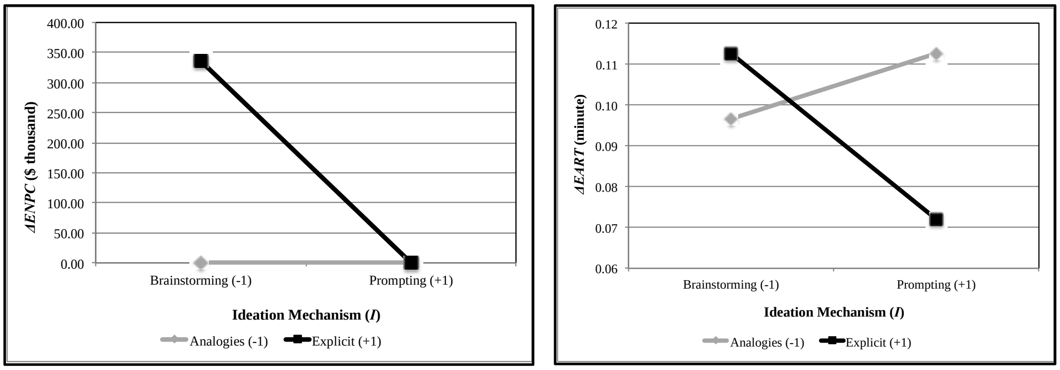

Figure 1. Mean plots for improvements in terms of ΔENPC (left) and ΔEART (right).

5.2 Improvement in Process Satisfaction (Δ*PS*)

Figure 2 shows that the analogies lecture had a main effect on ΔPS ($\beta_I = -0.87$, $p = 0.09$). This implies that participants were more satisfied on average after the analogies lecture than they were after explicit training on flexibility. This may be because analogies were like stories and perhaps more intuitive, as opposed to explicit training, which was more abstract and conceptual. There was also a noticeable interaction effect ($\beta_{EI} = 0.60$, $p = 0.24$). Combining analogies with brainstorming led to the best overall improvement ($\Delta PS = 2.20$ points) but combined with prompting, it led to even negative improvement ($\Delta PS = -0.75$ points). It may be that participants preferred the freedom provided by free undirected brainstorming, while they did not appreciate as much the structure imposed by prompting.

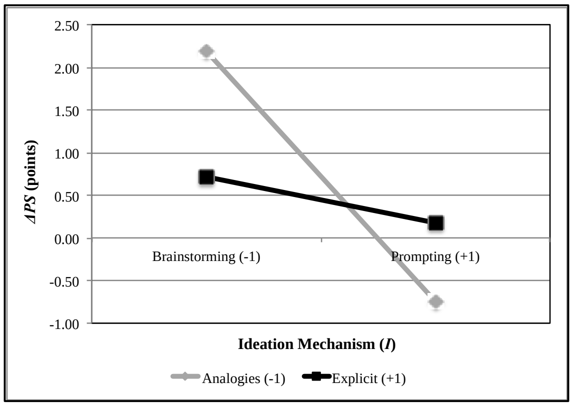

Figure 2. Mean plots for improvements in terms of ΔPS (process satisfaction).

6 DISCUSSION AND CONCLUSION

The results suggest that procedure evaluation requires both quantitative lifecycle performance and qualitative user impressions, since none alone can highlight completely the strengths and weaknesses of different concept generation procedures. Here, analogies + brainstorming was best to improve ΔPS, but not lifecycle performance. The procedure combining explicit training + brainstorming led to the best overall results in terms of lifecycle performance. Student *t*-tests assuming unequal variances between all treatment and control responses revealed a two-tail statistically significant difference for $\Delta EART$ ($p = 0.05$), but not for $\Delta ENPC$ ($p = 0.33$). Hence, the treatments were helpful to improve average response time, but because no main effect was observed, no particular treatment stood out.

Flexible design concepts contributed mainly to $\Delta EART$ improvement compared to the benchmark. It is possible that the procedures led to little $\Delta ENPC$ improvement because of the structure of the design problem. It may be easier for students with a strong background in industrial and systems engineering to think about operational flexibility (i.e. related to daily operations, of which *EART* is a good performance measure) then to think about strategic flexibility (i.e. related to long-term planning and deployment, for which *ENPC* is a good measure). For example, many flexibility ideas improving *EART* exploited the idea of upgrading a fire station from basic to enhanced. While this improved the responsiveness in case of major fires, it did not improve lifecycle cost enough to make a significant difference. Few students suggested deploying many future fire stations in phases, as opposed to all at once, an example of strategic flexibility. This strategy would have taken more advantage of the time value of money, deferring additional costs to later, and further reducing *ENPC*.

6.1 Results Validity, Limitations, and Future Work

Many strategies were exploited to limit threats to internal validity of results. The strategies described in Section 4.3.2 helped control for undesired effects from exogenous factors. Two qualified, treatment-blind, and independent coders reviewed ideation transcripts to enhance interpretive validity, measurement reliability, and reduce researcher bias. Random assignments to treatment groups diffused the possibility that some participants would try harder for one treatment over another. Threats to external validity included the fact that participants were graduate students, as opposed to practicing engineers. The design problem did not capture the full complexity of a real-world system. Nevertheless, participants were mature graduate students with many years of experience in different industries. Even if the design problem was simplified, it was modeled through close interactions with the main emergency service provider in a major Asian city. The benchmark solution captured best current practice. The same market and stochastic parameter assumptions were used to evaluate all flexible design concepts to improve measurement reliability. Building upon a validated survey enhanced response reliability. Cronbach α values between 0.93 and 0.99 showed that survey items reliably measured constructs within and across participants (Cortina, 1993).

Designing complex systems for flexibility is not an easy process. More effort is needed to understand the effects of procedures on flexible design concept generation, depending on the kind of engineering system under study, sample population, and context. Measuring both quantitative performance and qualitative user impressions matters, since procedures may well improve performance, but may be too cumbersome for use in practice, or vice-versa. Two very different procedures imposing different constraints to the creativity process were evaluated here – prompting being more rigid, brainstorming more open – realizing that design processes in general lie somewhere in between. The study by Cardin et al. (2012) using similar procedures (except analogies) but a fundamentally different design problem (i.e. real estate) showed different results, partially attributable to the different nature of the system. These observations call for more work to gain better understanding of the conditions under which concept generation procedures work, together with their effects depending on the problem and setting.

ACKNOWLEDGMENTS

The authors are thankful for the financial support provided by the National University of Singapore (NUS) Faculty Research Committee via MOE AcRF Tier 1 grants WBS R-266-000-061-133. We are thankful to Erlin Wiyono Wong and Mehdi Bourani Ranjbar for their help with experiments. We thank the intellectual support provided by all ISE colleagues, and active participation by IE5003 students.

REFERENCES

Bostrom, R. P. and Nagasundaram, M. (1998) Research in Creativity and GSS, *Proceedings of the 31st Annual Hawaii International Conference on System Sciences*, Kohala Coast, HI, United States: IEEE Computer Society Press.

Briggs, R. O., Reinig, B. A. and de Vreede, G.-J. (2006) Meeting Satisfaction for Technology-Supported Groups: An Empirical Validation of a Goal Attainment Model. *Small Group Research,* Vol. 37, No. 6**,** pp. 585-611.

Cardin, M.-A., Kolfschoten, G. L., Frey, D. D., Neufville, R., Weck, O. L. and Geltner, D. M. (2012) Empirical Evaluation of Procedures to Generate Flexibility in Engineering Systems and Improve Lifecycle Performance. *Research in Engineering Design,* Vol., No.**,** pp. 1-19.

Cortina, J. M. (1993) What is Coefficient alpha? An Examination of Theory and Applications. *Journal of Applied Psychology,* Vol. 78, No.**,** pp. 98-104.

de Neufville, R. and Scholtes, S. (2011) *Flexibility In Engineering Design.* Cambridge, MA, United States: MIT Press.

de Vreede, G.-J., Vogel, D., Kolfschoten, G. and Wien, J. (2003) Fifteen Years of GSS in the Field: A Comparison Across Time and National Boundaries, *36th Annual Hawaii International Conference on System Sciences*, Big Island, HI, United States: IEEE Computer Society.

Engel, A., Reich, Y., Browning, T. R. and Schmidt, D. M. (2012) Optimizing System Architecture for Adaptability, Proceedings of DESIGN 2012, the 12th International Design Conference, 2012 Dubrovnik, Croatia: The Design Society.

Eppinger, S. D., Fine, C. H. and Ulrich, K. T. (1990) Interdisciplinary Product Design Education. *IEEE Transactions on Engineering Management,* Vol. 37, No. 4**,** pp. 301-305.

ESD (2011) Engineering Systems Division Strategic Report, Cambridge, MA, United States: Massachusetts Institute of Technology.

Frey, D. D. and Dym, C. L. (2006) Validation of Design Methods: Lessons from Medicine. *Research in Engineering Design,* Vol. 17, No.**,** pp. 45-57.

Fricke, E. and Schulz, A. P. (2005) Design for Changeability (DfC): Principles to Enable Changes in Systems Throughout Their Entire Lifecycle. *Systems Engineering,* Vol. 8, No. 4**,** pp. 342-359.

Guma, A., Pearson, J., Wittels, K., de Neufville, R. and Geltner, D. (2009) Vertical Phasing as a Corporate Real Estate Strategy and Development Option. *Journal of Corporate Real Estate,* Vol. 11, No. 3**,** pp. 144-157.

Knoll, S. W. and Horton, G. (2010) Changing the Perspective: Improving Generate ThinkLets for Ideation, *43rd Annual Hawaii International Conference on System Sciences*, Koloa, Kauai, HI, United States: IEEE Computer Society.

Kurtoglu, T., Campbell, M. I. and Linsey, J. S. (2009) An Experimental Study on the Effects of a Computational Design Tool on Concept Generation. *Design Studies,* Vol. 30, No.**,** pp. 676-703.

Larson, R. C. and Odoni, A. R. (1981) *Urban Operations Research.* NJ, United States: Prentice-Hall.

Mikaelian, T., Rhodes, D. H., Nightingale, D. J. and Hastings, D. E. (2012) A Logical Approach to Real Options Identification With Application to UAV Systems. *IEEE Transactions on Systems, Man, and Cybernetics: Part A,* Vol. 42, No. 1**,** pp. 32-47.

Ong, M. E. H., Ng, F. S. P., Overton, J., Yap, S., Andresen, D., Yong, D. K. L., Lim, S. H. and Anantharaman, V. (2009) Geographic-Time Distribution of Ambulance Calls in Singapore: Utility of Geographic Information System in Ambulance Deployment (CARE 3). *Annals Academic of Medicine,* Vol. 38, No. 3**,** pp. 184-191.

Osborn, A. F. (1957) *Applied Imagination: Principles and Procedures of Creative Thinking.* New York, NY, United States: Scribeners and Sons.

SCDF (2012) *RE: Personal Communication with Singapore Civil Defence Force.*

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.**,** pp. 377-384.

Spitas, C. (2011) Analysis of Systematic Engineering Design Paradigms in Industrial Practice: Scaled Experiments. *Journal of Engineering Design,* Vol. 22, No. 7**,** pp. 447-465.

Strauss, A. and Corbin, J. (1990) *Basic of Qualitative Research: Grounded Theory Procedures and Techniques.* Newbury Park, CA, United States: Sage Publications.

Suh, E. S., de Weck, O. L. and Chang, D. (2007) Flexible Product Platforms: Framework and Case Study. *Research in Engineering Design,* Vol. 18, No.**,** pp. 67-89.

Trigeorgis, L. (1996) *Real Options.* Cambridge, MA, United States: MIT Press.

Wang, T. and de Neufville, R. (2005) Real Options 'In' Projects, *Real Options Conference*, Paris, France.

Yang, M. (2007) Design Methods, Tools, and Outcome Measures: A Survey of Practitioners, *ASME International Design Engineering Technical Conferences and Computers and Information in Engineering Conference*, Las Vegas, NV, United States: IDETC/CIE 2007.

Yang, M. (2009) Observations on Concept Generation and Sketching in Engineering Design. *Research in Engineering Design,* Vol. 20, No.**,** pp. 1-11.