An Experimental Methodology to Evaluate Concept Generation Procedures Based on Quantitative Lifecycle Performance

Michel-Alexandre Cardin

Department of Industrial and Systems Engineering, National University of Singapore
Block E1A #06-25, 1 Engineering Drive, Singapore, 117576
macardin@nus.edu.sg

Gwendolyn L. Kolfschoten

Technology, Policy, and Management Faculty, Delft University of Technology
5 Jaffalaan, a2.270, Delft, 2628BX, Netherlands
G.L.Kolfschoten@tudelft.nl

Daniel D. Frey

Department of Mechanical Engineering, Massachusetts Institute of Technology
77 Massachusetts Avenue, 3-449D, Cambridge, MA, 02139, United States
danfrey@mit.edu

Richard de Neufville

Engineering Systems Division, Massachusetts Institute of Technology
77 Massachusetts Avenue, E40-245, Cambridge, MA, 02139, United States
ardent@mit.edu

Olivier L. de Weck

Department of Aeronautics and Astronautics, Massachusetts Institute of Technology
77 Massachusetts Avenue, 33-410, Cambridge, MA, 02139, United States
deweck@mit.edu

David M. Geltner

Department of Urban Studies and Planning, Massachusetts Institute of Technology
77 Massachusetts Avenue, W31-310, Cambridge, MA, 02139, United States
dgeltner@mit.edu

Abstract. This study presents an experimental methodology to measure how concept generation procedures can affect the anticipated lifecycle performance of engineering systems design concepts. The methodology is based on objective and quantitative measurements of anticipated lifecycle performance of the design concepts. It merges cognitive and computer-aided techniques from the fields of collaboration engineering, creativity, and engineering design. It complements the body of existing techniques relying on subjective expert assessments, and other objective metrics not explicitly measuring anticipated lifecycle performance (e.g. development time and cost). Application of the methodology is demonstrated through evaluation of design procedures generating flexibility in engineering systems design. Experiments had ninety participants generate creative design alternatives to a simplified real estate development design problem. Thirty-two teams of two to three participants performed the collaborative design exercise. An online Group-Support System interface enabled efficient data collection and analysis. A computationally efficient mid-fidelity model was used to evaluate flexible design concepts quantitatively based on real options analysis techniques.

Introduction

This study proposes an effective and efficient experimental methodology to evaluate the effect of a design procedure on the anticipated lifecycle performance of engineering systems design concepts generated early in the design cycle. The methodology merges recent developments in cognitive and computer-aided techniques from the fields of collaboration engineering, creativity, and engineering design. Such techniques can demonstrably enhance the anticipated lifecycle performance of design concepts generated at an early stage (Bostrom and Nagasundaram 1998; Shah, Kulkarni, and Vargas-Hernandez 2000).

The proposed methodology is a significant building block of an experimental platform enabling thorough and rigorous evaluation of concept generation procedures in controlled laboratory and real-world settings. It is potentially applicable for comparison and evaluation of conceptual design procedures like Axiomatic Theory (Suh 1990), brainstorming (Osborn 1957), Pahl and Beitz (1984), and TRIZ (Altshuller 1973) all aiming at producing better performing designs.

Development of the approach is motivated by the fact that design procedures are often evaluated based on subjective expert assessments or other objective metrics not explicitly enabling objective and quantitative assessment of anticipated performance (Shah, Vargas-Hernandez, and Smith 2002; Reinig, Briggs, and Nunamaker 2007; Nelson, Wilson, and Rosen 2009). As explained below, evaluation of a concept generation procedure for flexibility requires lifecycle performance assessment based on an objective and quantitative metric (e.g. Net Present Value – NPV). It is reasonable to assume that other concept generation procedures – e.g. robust design by Wu and Hamada (2000) – would also benefit from a more rigorous evaluation mechanism. This is due to inherent complexities in design activities that make it very difficult for any expert to evaluate design concepts from a subjective standpoint. Too many design variables, parameters, and scenarios need to be considered from a holistic standpoint. The methodology presented here builds upon and extends these approaches.

The motivation for the experimental approach pursued here is that typical case or protocol studies evaluating design procedure applications in industry can take a long time – e.g. (Ward et al. 1995), be difficult to reproduce rigorously, and hard to run several times over a short time period (Ullman, Wood., and Craig 1989; Goldschmidt and Weil 1998; Shah, Kulkarni, and Vargas-Hernandez 2000). The system of interest can be different from one study to another, together with experimental conditions. This complicates the task of determining what design procedure is best suited to a particular firm or organization, based on meta-analyses of such studies. The experimental methodology introduced here enables quick replications, and relies on efficient computer-aided techniques for written data collection and analysis, computer modeling, and evaluation of design concepts generated. It can be reproduced easily for different systems, in different environments, and to evaluate different design procedures of interest thoroughly and rigorously (Summers, Anandan, and Teegavarapu 2009).

The methodology is applied as demonstration to evaluate procedures supporting early generation of flexibility – also referred as flexible design concepts – early in engineering systems. The goal is to improve lifecycle performance by means of flexibility and adaptability in the face of uncertainty. The procedures are chosen because they may be particularly valuable to the design process. Indeed, systems designed today, face significant risk and uncertainty in environments, markets, resources, and technology. Yet many design

procedures do not fully exploit uncertainty and the possibility to adapt flexibly to changing circumstances. The underlying assumption is often that design requirements are known *a priori* from customer demands and preferences, and systems are optimized accordingly (Minai, Braha, and Bar-Yam 2006; Eckert et al. 2009).

Related Work

Empirical Procedure Evaluation

Evaluation metrics in empirical studies are often qualitative and subjective in nature. They may not support well quantitative assessment of anticipated lifecycle performance investigated. For instance, Hazelrigg (1998), and Hevner et al. (2004) used concept attributes like cost and weight, which are matched to subjective participants' utility preferences. Kurtoglu et al. (2009) introduced completeness, the level at which a concept variant addresses a sub-function depicted in the function structure. Linsey et al. (2005) used technical feasibility, the degree to which a concept can be implemented practically. Shah et al. (2000; 2002) suggested quality, quantity, novelty, and variety to assess creativity and how much the design space is explored (Guilford 1959; Mullen, Johnson, and Salas 1991). For example, an expert may assess quality for each concept using a 1-10 scale, and the weights are based on the importance of the concept given the design phase, and intended functions. Nelson et al. (2009) integrated these metrics to assess the quality of design space exploration. Briggs et al. (2006) used user satisfaction with the process and results.

As stressed in the real options literature, evaluating procedures for flexibility should be based on objective quantitative measurements of anticipated lifecycle performance. Most metrics above however rely on expert assessments of qualitative nature. Assessing performance in this context can be challenging, even for an expert. So many design variables, parameters, decision rules, long-term strategies and scenarios should be considered. Relying on metrics like cost, weight, and linkography may not indicate how a concept will perform in operations. A concept can be rated as highly complete, feasible, novel or of high quality, but there is no guarantee it will perform as expected once launched. It is not clear whether high quantity and variety of concepts necessarily improves performance. As an indicator however, positive correlations have been found between outcome quantity and quality (Yang 2009). It is possible though that focusing the design effort on fewer but good flexible design concepts may improve performance significantly. Also, it is not clear whether procedures providing good user satisfaction with the process and results lead to better performance. These observations motivate the quantitative performance-based evaluation methodology presented here. Objective and quantitative evaluation of the concepts in operations can be based on financial economics, and explicit scenario modeling.

Flexibility in Engineering Design/Real Options

Definition and Examples. Flexibility in engineering design enables a system to change easily in the face of uncertainty (Fricke and Schulz 2005). It is associated to the concept of Real Option, which provides 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, and/or mixing the above (Trigeorgis 1996). Real Options "in" a system are technical engineering and design components enabling options in operations (Wang and de Neufville 2005). Real Options – referred here interchangeably as *flexible design concepts* –

can be characterized by a *strategy* (or *type*) and *enabler* (or *mechanism*) (Mikaelian et al. 2011).

The Real Options Analysis (ROA) literature focuses on the economic valuation of flexibility (Dixit and Pindyck 1994; Myers 1977; Trigeorgis 1996). It builds upon work in financial options by Black and Scholes (1973) and Cox et al. (1979). Many studies have shown that flexibility in engineering projects brings expected performance improvements ranging between 10% and 30% compared to standard design and evaluation approaches (de Neufville and Scholtes 2011; Copeland and Antikarov 2003; Amram and Kulatilaka 1999). Flexibility improves expected performance by affecting the distribution of possible outcomes. It reduces the effect from downside, risky scenarios, while positioning the system to capitalize on upside, favorable opportunities. Examples in engineering systems design abound: strategically phasing the development of airport terminals over time (de Neufville and Odoni 2003), designing offshore platforms for future capacity expansion (Jablonowski, Wiboonskij-Arphakul, and Neuhold 2008), investing strategically in new nuclear plant construction (Rothwell 2006), adapting supply chains flexibly to uncertain exchange rates (Nembhard, Shi, and Aktan 2005), developing flexible water infrastructures (Zhang and Babovic 2011), etc.

One example of flexibility in real estate is the ability to expand a building vertically (Guma et al. 2009). The HCSC Building in Chicago exploited the strategy to "build small and expand later *if needed*" (Figure 1). This strategy reduced exposure to losses because less capital was required upfront. It also gave access to more profits under favorable market conditions to build more offices, hire personnel, and ultimately generate more profits. This strategy was enabled by carefully engineering the infrastructure for expansion in the early 1990s (e.g. larger elevator shafts, stronger structure). The company exercised the flexibility a few years ago. The expansion phase should be completed by 2011. Other examples in real estate include designing buildings that can be later appropriated for multiple purposes, or anticipating future coupling with other buildings.
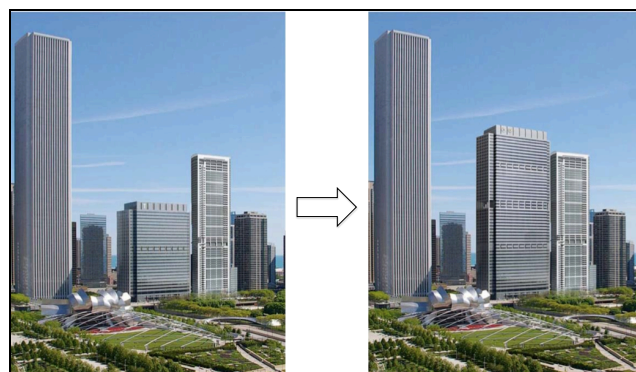

Figure 1 Example flexible engineering system: the HCSC Building in Chicago (Guma et al. 2009).

Overall Approach

The hypothesis is that it is possible to develop an effective and efficient methodology enabling objective and quantitative evaluation a concept generation procedure based on anticipated lifecycle performance of the concepts generated in experiments. An empirical approach is taken to test this hypothesis, by explicitly applying the methodology to evaluate a design procedure of interest. As suggested by Frey and Li (2010), this approach is inspired

from *naturalistic* epistemology, where knowledge is created by gathering evidence through sensorial experience and real experiments (1998). Incidentally, the test of the hypothesis provides information relevant to testing a second hypothesis. It is claimed that a procedure for flexible design concept generation can help designers generate flexible design concepts that improve the overall anticipated performance of the system.

In this paper the concepts of effectiveness and efficiency are operationalized as follows. Effectiveness is referred as the ability to measure the effects that the design procedure intends to produce. Therefore, if the methodology enables objective comparisons of different procedures based on quantitative effects on lifecycle performance of design concepts, it is effective. If the methodology does not enable such measurements, it is ineffective. Efficiency is referred as the ability to replicate experiments and analyze data within a reasonably short time (i.e. a few weeks). If the methodology does not enable a study to be completed within such timeframe, it is inefficient.

Experimental Methodology

Steps 1 to 5 represent the novel elements of the proposed methodology. They complement typical preliminary steps like setting up the Design Of Experiments (DOE), choosing the design procedure of interest (e.g. Axiomatic Theory, Design for Flexibility, Robust Design), breaking down the design procedure into treatment and control factors and levels, and selecting a pool of qualified participants.

Step 1: Design Problem Description

The first step is to setup a benchmark design problem. The design problem is described through a short presentation including clarifications in terms of the design context, market environment, operational conditions, and any other information deemed useful to participants. An example image can be provided for mental conceptualization of the system as a starting point, although not an absolute requirement. A quantitative metric is clearly defined and explained to assess performance of design concepts. An initial benchmark design solution is provided as starting point for idea generation ("ideation"). The task is clearly defined that participants should generate design alternatives that improve overall anticipated performance.

The design problem should be a simple version of a realistic problem. At the same time, it must be complex enough to provide room for creativity and concept generation. It should be chosen carefully so it is accessible to all participants. An engineering, management, and/or design background should be all that is necessary to contribute positively. If the design problem is too specialized, results may be biased in favor of participants that have specialized knowledge. It may also be difficult for other participants to contribute anything useful.

Step 2: Computer Model

The second step is to develop a computer model to measure the anticipated lifecycle performance of design concepts objectively and quantitatively. Its main characteristics should be that it can 1) be developed relatively quickly (e.g. within a few weeks), and 2) run quickly (e.g. within minutes). It should be a mid-fidelity model in contrast to high-fidelity models typically developed in industry, which take months of development, and hours and days of computational time (de Neufville and Scholtes 2011). To evaluate early design concepts, there is no need for high-level complexity and accuracy. It is only important to be able to rank-order different design alternatives objectively and quantitatively.

A quantitative performance metric is chosen to assess anticipated lifecycle performance of design concepts objectively and quantitatively. It can measure an explicit attribute of performance (e.g. service rate emerging from a particular hospital design), economic performance (e.g. financial value stemming from the cash flows generated by a given design), or physical performance (e.g. how fast a robot prototype can run in a virtual environment). The computer model combines design variables, parameters, and operational scenarios to evaluate quickly and explicitly different design concepts. For example, computer models can be developed using a real-time computer-assisted sketching tool – e.g. iCampus MIT-Microsoft Alliance, (2010), a financial spreadsheet running simulations in Excel® (as done here), an optimization algorithm in Matlab® - e.g. Hassan and de Neufville (2006), or a real-time CAD software – e.g. Fumarola et al. (2010). Such performance measurement complements typical dependent variables in design studies summarized below.

Step 3: Online Group-Support System Interface

The third step sets up an online Group-Support System (GSS) interface to structure the collaborative design process, improve efficiency in recording ideation data, help in moderating ideation sessions, and stimulating creativity (Nunamaker et al. 1997; Bostrom and Nagasundaram 1998). The online GSS interface provides an easy and efficient way to record participants' creative responses to a design problem, and enable ideation at distance (i.e. not all designers need to be physically co-located, which is enabled by the online software). It should ideally not require any special facility (other than personal laptops) so ideation sessions can be conducted at any site. This approach is more effective than typical audio/video recording techniques. These may require special facilities and audio/visual equipment for recording, and hours and pages of transcription before further analysis is possible – e.g. as suggested by Johnson and Christensen (2004). The approach suggested here aims at increasing efficiency by skipping the lengthy transcription step, providing the freedom to perform experiments anywhere, and without the burden of having all designers in the same physical location. This is advantageous for companies where departments are physically located in different areas.

The online interface is a particular class of GSS defined more broadly as "socio-technical systems consisting of software, hardware, meeting procedures, facilitation support, and a group of meeting participants engaged in intellectual collaborative work" (de Vreede et al. 2003). It can be used, for instance, to structure ideation sessions by posting different brainstorm topics in a specific order, going through a particular meeting agenda, etc. GSS are used in the collaboration engineering community to stimulate creativity and alleviate concerns about evaluation apprehension – the fear of being judged – and production blocking – losing an idea because someone else is talking (Bostrom and Nagasundaram 1998; Warr and O'Neill 2005). Creativity techniques supported by GSS help people who are not trained professionally to moderate ideation sessions almost as efficiently as trained professionals (Kolfschoten et al. 2009).

Step 4: Data Collection

The fourth step consists of structuring each experiment to maximize statistical significance, by collecting data before-and-after applying the design procedure under evaluation (i.e. pretest-posttest structure). The method controls for variability between the responses that different teams generate when subjected to similar treatment conditions (i.e. within-group variability, where group here refers to treatment, not team). This approach is important because experiments in creativity and collaboration engineering often involve people with different backgrounds, creativity levels, and trainings. There is a high chance the quality of

responses within groups undergoing the same treatment will vary significantly (i.e. the noise, or unexplained variability). For instance, some teams might be more creative and thus naturally generate many ideas compared to the treatment group average. Other teams may know more about the design procedure of interest and generate better ideas on average, even though every participant is screened for such prior knowledge. Focusing on the differential performance within the same groups reduces mean within-group variability. It restricts attention to improvements between different treatments to measure the best possible "signal-to-noise" ratio from experimental sessions.

This step sets up each experimental session to measure an explicit difference "Δ" between the response measured in an initial and a subsequent session. This controls for mean within-group variability by measuring an *improvement* compared to an initial response set by each team. This framework organizes each data collection as in Figure 2. That is, the design problem is described, and then participants are asked in session 1 to generate design concepts that improve anticipated performance without any guidance. They are asked to vote on each design concept generated to discriminate between ostensibly contradictory concepts in the transcript analysis phase. In session 2, the same task is repeated, with the only difference that a treatment of choice is applied. In short, the proposed method collects data on the possible effect of any design procedure, which can then be compared to possible improvements due the passage of time. A debrief explains the rationale of the experiments. A post-experimental survey is used to gather participants' demographics information, and other qualitative responses of interest.
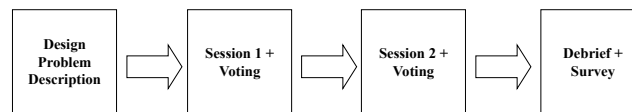

Figure 2: Suggested pretest-postest structure used in experiments.

Step 5: Analysis

Transcript Analysis. This aims at efficiently extracting the concepts generated from the ideation transcripts. This can be done using a standard coding procedure – e.g. Strauss and Corbin (1990), as explained by Trauth and Jessup (2000). Researchers determine first a basic set of criteria that define the "completeness" of an idea – or design concept. An idea is complete if it has all the information necessary for implementation and quantitative analysis by a third person using the computer model. Ideation transcripts are reviewed by at least two independent coders in search for these basic criteria. Once a complete idea is found, it is recorded.

Response Measurements. Transcript analysis gives rise to a set of example responses $\Delta y \geq 0$ generated from session 1 to session 2, measurable for each experimental session:

1. ΔC: the improvement in the number of complete concepts. For example, if one complete concept is generated in session 1, and two new complete concepts are generated in session 2, $\Delta C = C2 - C1 = 3 - 1 = 2$. This measurement is in line with the typical "idea count" metric found in many design studies. It can be used to assess the creativity level pertaining to a particular treatment;

2. ΔG: the improvement in the number of good concepts. This measurement is in line with the "good idea count" metric defined in Reinig et al. (2006). If a complete

concept improves performance compared to the benchmark design solution (or the threshold for good quality), it is considered good;

3. ΔP: the improvement in anticipated performance by implementing only good concepts. This is measured in units of the quantitative metric. For example, assume a benchmark design offers anticipated financial performance of 9.3 millions. If a good concept is generated in session 1 of \$9.5 millions ($P1$ = \$0.2 million) and another good concept is generated is session 2 of \$10.0 millions ($P2$ = \$0.7 million), $\Delta P = P2 - P1 = 0.7 - 0.2 = \$0.5$ millions.

Other complementary responses of interest can be measured using the "Δ" framework, which do not necessarily rely on the quantitative computer model. For example, one may measure the differences in qualitative user satisfaction with the design procedure, satisfaction with results, anticipated quality of design outcomes (Briggs, Reinig, and de Vreede 2006), etc.

Statistical Analysis. The statistical analysis determines whether the design procedure and/or any of its individual factors have main and interaction effects on the dependent variables. Each response is modeled assuming the General Linear Model (GLM) described in Milton and Arnold (1990):

$$\Delta y(x_1, x_2, \ldots, x_n) = \beta_0 + \sum_{i=1}^{n} \beta_i x_i + \sum_{i=1}^{n} \sum_{\substack{j=1 \\ j>i}}^{n} \beta_{ij} x_i x_j + \varepsilon$$

with Δy being the response of interest – for instance ΔC, ΔG, or ΔP – for factor $x_i \in [-1, +1]$ and $i, j = 1, 2, 3, \ldots, n$ assuming a two-level DOE setup, with $j > i$. Coefficient values are calculated using standard least-square minimization regression. β_0 approximates the grand total mean from the dataset, β_i the main effect for factor x_i, and β_{ij} the two-way interaction effect between factors x_i and x_j (higher order interactions are not displayed here for simplicity, but can be considered). The term ε represents the pure experimental error for each response compared to the group mean. It is assumed to have a normal distribution with mean 0 and variance σ^2. The null hypothesis is that H_0: $\beta_i = \beta_{ij} = 0$, $\forall$ $i, j, j > i$.

Concept Generation Procedures

Table 1 summarizes the four procedures – or treatments – evaluated experimentally. To craft simple and intuitive procedures, only two factors were considered: educational training received on flexibility (E), and ideation mechanism (I) used to stimulate creativity.

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

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

Educational Training (E)

Many studies show that educational training and pedagogy play a role in the ability to generate design concepts (Lindemann 2003; Bender and Blessing 2003). A short *explicit training* program may help designers generate flexibility in engineering systems. This is captured by level $E = +1$. The treatment is a short 15-20 minutes lecture on flexibility in large-scale infrastructure systems – a class of engineering systems (ESD 2011). The lecture is

expected to help designers become aware of the effects of uncertainty on lifecycle performance. It should open their mind to the potential of flexibility to deal with uncertainty. It describes generic sources of uncertainty affecting performance, why flexibility can improve performance, and why it must be considered in the early phases of design. It also discusses what important elements form a complete[1] flexible design concept. It provides real-world example applications of these principles in the aerospace and oil industries.[2]

In reality, engineering designers may or may not have received explicit training on flexibility during their educational training and professional experience. This reality is more likely to represent the wider population of designers. It is captured by factor level $E = -1$, and called *current training*. In experiments, this treatment leaves participants address the design problem without emphasis on flexibility. Participants generate concepts with the goal of improving performance based on their background and experience. This treatment serves as the baseline from which to compare the effect of explicit training on flexibility. It assumes that if participants had such training and thought flexibility could improve performance, they would incorporate it naturally in their thinking.

Ideation Mechanism (I)

A *prompting* procedure is suggested to help scaffold the thought process systematically, as captured by level $I = +1$. Prompting is simple, intuitive, and useful to stimulate creativity in collaborative activities (Santanen, Briggs, and de Vreede 2004). Asking direct questions may trigger collective discussions more effectively than relying on industry guidelines alone. Also, prompting is similar to the approach used by researchers working with senior engineers and decision-makers in industrial case studies. Interviews were necessary to elicit sources of uncertainty and flexibility, prior to interpretation, modeling and economic evaluation (Lin 2009; Yang 2009; Suh, de Weck, and Chang 2007; Mikaelian et al. 2011). The prompting mechanism in Appendix builds upon industry and real option guidelines. It is also loosely based on the systematic processes by Babajide et al. (2009), de Neufville and Scholtes (2011), and Walker et al. (2001).

Free undirected brainstorming is a simple and intuitive approach to stimulate creativity based on the work of Osborn (1957). It is widely used in industry and academia, as demonstrated by a survey of U.S. academics and practicing engineers (Yang 2007). Captured in experiments by level $I = -1$, it encourages designers to avoid any barriers to creativity, to express ideas freely and without any influence from a moderator. Free undirected brainstorming represents a good alternative to stimulate flexible concept generation. Given it is widely used, it represents an appropriate baseline for comparison with the prompting mechanism.

Example Application of the Methodology

The methodology above was applied to evaluate the procedures for flexible design concept generation. A detailed version is provided in Cardin (2011). Ninety participants were recruited from professional masters and doctoral programs in engineering systems, design, and management at a U.S. engineering institution. Table 2 summarizes their demographics. They were recruited via class and electronic email announcements at the beginning of two courses on systems design and engineering analysis. The announcement invited voluntary participation to an experiment (off class hours) on flexibility in engineering design. Most

---

[1] Definition specific to this procedure evaluation is provided in step 5 below.
[2] Lecture slides can be downloaded at http://web.mit.edu/macardin/Public/docsResearch/lectureFlexibility-MIT.pdf.

participants were mature graduate students with training in engineering, science, and/or management, and many years of industry experience.

Table 2 Participant demographics

Group Characteristics	Category	Percent (%)
Age	< 25	14
	25-34	67
	> 35	19
Highest Education Level	Bachelor	49
	Master	49
	PhD	2
Gender	Female	19
	Male	81
Work Experience (years)	< 5 years	36
	5-9 years	39
	> 10 years	25

Step 1: Design Problem Description

The design problem consisted of how to best develop and deploy over three years the 309 units of the multi-family residential development project (Figure 3). This system exhibits the technical complexity, social intricacy, and elaborate processes fulfilling important functions for society typical of engineering systems (ESD 2011). One issue was to decide what kind of unit to develop in each phase – either condo (short for condominium) or apartment. Another issue was the number of units to deploy in each phase, and whether to develop/sell these units in each phase. The distinction between unit types was explained as different levels of quality, prices, and construction costs. A condo is typically more luxurious, built from expensive materials, and targets business professionals. Sales price and construction costs are typically higher. An apartment unit is functional, less luxurious, and for students and middle-class families. The sales price and construction costs are lower.

Figure 3: Example figures provided to participants for mental conceptualization of the real estate design problem. Source: http://www.northpointcambridge.com.

The benchmark design was set as a condo-only residential project, with all 309 units developed in year/phase 1. There was no specification how this design was achieved (i.e. how many buildings and units per building, etc.), to minimize fixation issues (Linsey et al. 2010). This choice was justified by providing the highest NPV between a condo-only and an apartment-only project. Even though this choice appears simplistic and rigid, it represents best practices in the real estate industry (Guma et al. 2009; Kronenburg 2007).

Step 2: Computer-Model

Figure 4 shows the computer-based Excel® DCF model used to measure lifecycle performance objectively and quantitatively of the flexible design concepts, based on the model by Geltner and Cardin (2008). A full description of the cost and revenue generating mechanisms is provided in Cardin et al. (2012) and the Appendix section of Cardin (2011).[3] A risk-neutral Expected NPV (ENPV) – or average NPV – metric was used because it balances risk-seeking and risk-averse design decisions. Other metrics could have been used to suit different preference utilities – e.g. standard deviation of the response, as in the Taguchi method (1987). It is left as for future work to replicate this study using different performance metrics.

NPV w Flexible Choice Each Phase:			Phase 1	Phase 2	Phase 3
Year		0	1	2	3
Next Phase Developed As:		CONDO	APT	APT	
Sales Price/Unit			197,947	242,267	249,535
Units Demand			100	82	78
Constr & Sales/Unit			176,259	130,112	126,955
Develop Current Phase?			YES	YES	YES
Planned Capacity Deployment			100	103	106
Expand Capacity this Phase?			NO	NO	NO
Additional Capacity			0	0	0
Total Capacity Added			100	103	106
Units Sold			100	82	78
Sales Revenue			19,794,739	19,984,986	19,472,279
Total Constr & Sales Costs			17,625,907	13,401,562	13,457,232
Net Cash Flow			2,168,832	6,583,424	6,015,047
PV of Cash Flow			2,008,178	5,644,225	4,774,938
NPV (exclu land)		12,427,340			

Figure 4: NPV model for the real estate development design problem.

Design and engineering tradeoffs were in terms of the unit type (i.e. condo vs. apartment), deployment strategy (staged vs. all at once vs. deferred), and unit capacity allocation in each phase. These decisions affected the cash flows (CF_t) at each time $t = [0, T]$ and NPV generated according to the equation below. NPV enabled discriminating between different design alternatives based on quantitative lifecycle arguments. For example, the decision to select condo versus apartments could affect the sales price and construction cost, as they were both higher for condo units than apartments. Also, discounting cash flows at opportunity cost of capital r would imply that timing and unit allocation in each phase mattered from a managerial standpoint. Later cash flows would be more heavily discounted in the model, and weigh less in the NPV. These decisions had engineering and cost implications affecting later phases of the design process. They were modeled and evaluated at the conceptual level.

$$NPV = \sum_{t=0}^{T} \frac{CF_t}{(1+r)^t}$$

For example, the row "*Next Phase Developed As:*" specified the unit type developed next phase, either "*CONDO*" or "*APT*". The decision rule for a *switching* flexibility strategy was implemented here, as prescribed by participants in ideation transcripts. An example decision rule was "*if current cash flows are higher for condos than for apartments, then develop next phase as condo, else develop next phase as apartments*". Following Excel®'s programming language =*IF(logical test, value if true, value if false)*, this could be implemented as:

$$= IF(CF_{t\text{-}condo} > CF_{t\text{-}apt}, "CONDO", "APT")$$

---

[3] Available online at http://ardent.mit.edu/real_options/Real_opts_papers/PhD_Cardin-thesis.pdf

Anticipated lifecycle performance was measured using ENPV for each flexible design concept. This was done by simulating stochastically $M = 2,000$ combined scenarios of price, demand, and cost, based on the Itô formulation of Geometric Brownian Motion (Itô 1951). Each scenario combination would produce different cash flows based on the flexibility strategy and decision rules implemented, and one NPV measurement. ENPV was calculated as:

$$ENPV = E[NPV] = \frac{1}{M}\sum_{m=1}^{M} NPV_m$$

Each run took ~1 second on a standard 2.4 GHz Intel Core 2 Duo MacBook with 4 GB of RAM, running Excel® 2004 on an OSX 10.6 platform.

Step 3: Online Group-Support System Interface

The ThinkTank® GSS online software by GroupSystems® was used (Figure 5). It enabled participants to type in real-time descriptions of their design solutions, similar to that of chatting software. It worked with any standard Internet browser. After the problem description, the moderator posted the brainstorming task (i.e. ideation topic) of improving performance compared to the benchmark design. Each member described their solutions, which were displayed to other members to stimulate creativity and engage discussions. Each member could reply, comment, and append new ideas to a thread. The software also provided the voting module to rate concept quality. The interface provided an easy and efficient way to record participants' creative responses to the design problem – instead of manual transcription. The online feature allowed a few distance students to participate.
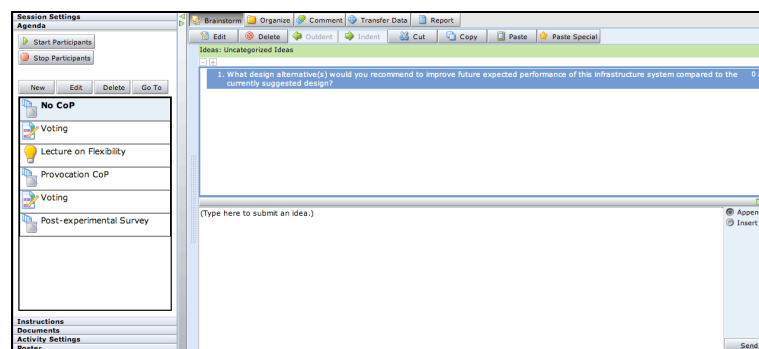

Figure 5: Online GSS interface used in experiments (adapted from ThinkTank® by GroupSystems).

Step 4: Data Collection

At the beginning of each experimental session, the moderator welcomed participants, and described the design problem and context.[4] The participating team was told to represent an internal consulting firm at a renowned multi-family residential real estate firm. The task was assigned to brainstorm and suggest alternative design and development plans potentially improving anticipated lifecycle performance compared to the benchmark design (i.e. NPV). There was no emphasis to complete this task by means of flexibility – although participants were aware the experiment was about flexibility in engineering design.

---

[4] Slides can be downloaded at http://web.mit.edu/macardin/Public/docsResearch/REmodel-MIT.pdf.

Participants repeated the exercise in session 2 for 25 minutes under one of the treatments in Table 1, then followed by another 5 minutes for voting. Treatments 1 and 3 used free undirected brainstorming. Treatments 2 and 4 used the prompting mechanism in Appendix, allocating 5 minutes for discussions on each question. In treatments 3 and 4 only, the short lecture slides were presented before the session began. A debrief explained the purpose of the study after all experiments. Demographics information was collected. Collected data were written ideation transcripts describing conceptual design solutions in plain text, with voting scores, and surveys scores.

Step 5: Analysis

The goal is to evaluate the effects of the design procedure by measuring the improvements on the responses of interest Δy between sessions 1 and 2. This is done first by extracting the relevant design concepts through transcript analysis. Then, the main and interaction effects on the Δy responses are evaluated through statistical analysis.

Transcript Analysis. Written ideation transcripts were analyzed to extract *complete* flexible design concepts using a coding procedure (Strauss and Corbin 1990; Trauth and Jessup 2000). A design concept was considered *complete* if it contained coherent information about the following elements (using the switching example):

1. An *uncertainty source* affecting anticipated performance (e.g. unit demand)
2. A *flexible strategy* to adapt to the above uncertainties in design and operations (e.g. switch between condo and apartments)
3. A conceptual but concrete description of the *flexibility enabler*, considering engineering design, legal, management, and/or financial aspects (e.g. design each unit as empty shells to be finished later as condo or apartments)
4. A *decision rule*, "*if*" statement, or "trigger mechanism" based on observations of the uncertainty sources, determining when it is appropriate to exercise the flexibilities (e.g. if demand is higher for apartments than condos, switch to finishing and selling units as apartments, if not finish and sell as condos.

The switching strategy above contrasts with the benchmark inflexible plan where all units are developed at once as condos. It requires developing units as empty shells to be finished later. This is radically different than designing all units as condos. It is not clear at the conceptual stage what design and decision rules are most profitable given the uncertainties, hence the need for explicit modeling. Other examples of complete concepts from transcripts are: phase the development and deploy capacity over time, expand or reduce unit capacity in each phase whenever appropriate, temporarily abandon the project if market conditions were not suitable, and do not develop a phase if market conditions are unfavorable.

Response Measurements. Each complete concept was implemented using the computer model to determine the subset of good concepts belonging to sessions 1 and 2. Different combinations of good concepts were evaluated to determine the best performance *ENPV*1 and *ENPV*2 achieved in each session compared to the benchmark design. It was important to evaluate different combinations of good concepts, as some may interact positively, and others negatively. The overall improvement response ΔP between sessions 1 and 2 was caused by the implementation of good flexibility concepts, as compared to the initial benchmark design. The performance metric *P* evaluated here was *ENPV*. Measurements of the dependent variable was computed as $\Delta ENPV = ENPV2 - ENPV1$. *ENPV*1 and *ENPV*2 were the best

lifecycle performance achieved in sessions 1 and 2 respectively (expressed in millions $). Each experimental session led to one $\Delta ENPV$ measurement.

Statistical Analysis. The GLM above was modified to suit the response of interest: $\Delta ENPV = \beta_0 + \beta_E E + \beta_I I + \beta_{EI} EI + \varepsilon$. The null hypothesis H_0: $\beta_E = \beta_I = \beta_{EI} = 0$ was tested. Standard least-square regression was used to calculate the main and interaction effects in Matlab®. The p-values of the main and interaction effects were calculated using a non-parametric permutations test (i.e. randomization or exact test) (Fisher 1935; Pitman 1937; Welch 1990). Because $\Delta y \geq 0$ for ΔC, ΔG, and $\Delta ENPV$, sample distributions were truncated about zero. This could not satisfy normality assumptions for standard parametric tests, hence the suggestion to use the non-parametric test.

The mean plot values obtained for each treatment group are shown in Figure 6. The graph shows that prompting had a significant main effect on $\Delta ENPV$ ($\beta_I = 0.98$, $p = 0.00$). This can be seen from Figure 6, where the mean $\Delta ENPV$ values increase consistently from left to right, as one toggles the factor level from -1 to $+1$. The procedure improved anticipated lifecycle performance compared to the benchmark by up to 36% percent (i.e. \$3.34 million/\$9.30 million) in combination with the lecture. These results are in line with those published in the real options literature (Dixit and Pindyck 1994; Trigeorgis 1996; Amram and Kulatilaka 1999; de Neufville and Scholtes 2011). The GLM response obtained was $\Delta ENPV = 1.88 + 0.48E + 0.98I$.
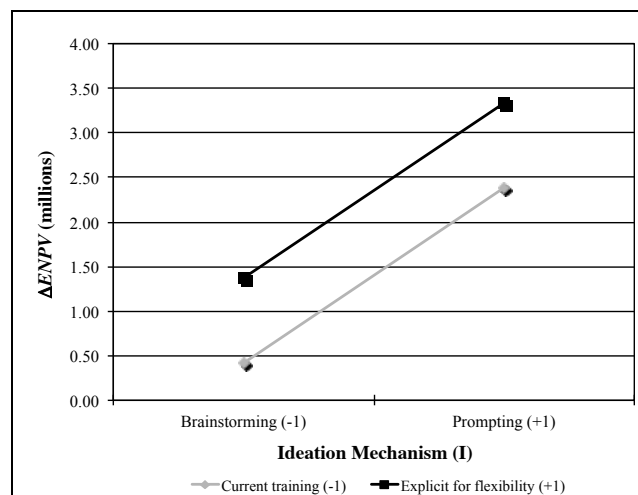

Figure 6: Mean plots for $\Delta ENPV$ (anticipated lifecycle performance). The lower curve shows mean results for current training on flexibility ($E = -1$). The upper curve shows results for explicit training on flexibility ($E = +1$).

Results and Discussion

The demonstration above shows application of the experimental methodology to evaluate concept generation procedures for flexibility. This confirms the first part of our main hypothesis that it is indeed possible to develop a methodology enabling objective and quantitative evaluation of the effects of a design procedure on anticipated performance of design concepts. The fact the demonstration was feasible implicitly suggests it may be applicable to evaluate other concept generation procedures as well. It is actually a weakness of this work that demonstration was made for only one case study. More applications evaluating other design procedures – for flexibility generation, and even perhaps other design activities such as design for robustness – and using other engineering systems would help

validate this claim further. A second application is currently under way, using analogies as a concept generation mechanism, and using an emergency service systems case study.

The test of the main hypothesis also confirmed the second hypothesis that the concept generation procedures helped participants generate valuable flexibility, in this case the prompting mechanism. The concepts in turn demonstrably improved anticipated lifecycle performance of the engineering system. It was interesting to observe that although a similar flexible strategy could be used in many instances, different decision rules would give rise to very different value improvements. For example, the best concept obtained for treatment 1 ($\Delta ENPV$ = $1.5 million) was to develop units according to whichever market demand was highest in each phase (i.e. a switching strategy). In treatment 2, the best concept was also a switching strategy ($\Delta ENPV$ = $5.1 million), although decision was based on which unit type would provide the highest profit in each phase. Under treatment 3, the based concept was also a switching strategy ($\Delta ENPV$ = $5.4 million) based on demand observation in the previous phase, and starting with a 50%-50% mix of condo and apartments in phase 1. In treatment 4, a switching strategy was also suggested in combination to a phasing deployment strategy ($\Delta ENPV$ = $6.7 million). Deployment would stop if costs went up dramatically, while switching would accommodate the market generating most profit opportunities in any given phase.

The results at the end of step 5 demonstrate that the proposed methodology was effective. Indeed, the methodology enabled quantitative measurements of the lifecycle performance improvements brought by the procedures by means of flexibility. One important issue, however, is that it is not clear that the lifecycle improvements measured in experiments would necessarily lead to quantifiable value improvement in real-world operations. Future work could be devoted to compare measurements produced *ex ante* to actual performance realizations in operations. Another difficulty was that the description of design concepts provided by participants may not have been complete, and not reflect the full extent of discussions. One way to improve this would be to record and transcribe discussions, although this would slow down the analysis significantly – which is one of the main advantages of this approach. Speeding up the analysis comes at the expense of precision. This study demonstrated, however, that it is still possible to extract a significant amount of information, even if some details may be left out.

The methodology was efficient for three reasons. First, it took about only one month to set up the experiments, recruit participants, and collect data. This was obviously contingent to having available participants, which a university setting favors. This compares favorably to typical case studies requiring extended fieldwork over several months at a company or organization. This demonstrates the well-known advantage of controlled user studies as compared to individual case studies (Summers, Anandan, and Teegavarapu 2009). On the other hand, it is true that no participant had expertise in this area. As further improvement, this experimental setting could be used in industry to determine the best ways to generate flexible design concepts in practice. Second, it only took about two weeks to develop the case study and computer model. This was also efficient compared to other research where the modeling part is more time consuming, and may require months. Third, transcript analysis and computer evaluation took a total of about seventy-five hours. Much transcription time was saved by having participants transcribed their answers themselves using the GSS interface. The approach avoided the intensive word for word transcription that is typically required if audio/video material is used.

Conclusion

This paper introduced and demonstrated application of an experimental methodology to evaluate objectively design procedures based on their effects on anticipated lifecycle performance of early design concepts. This methodology integrates a set of computer-aided techniques to promote effectiveness and efficiency based on GSS technology, and mid-fidelity computer modeling. It is effective and efficient because of the relatively low cost, time, and resources required for experimentation, data collection, and analysis. As demonstrated, it enabled measuring quantitatively the anticipated lifecycle performance of the concepts generated in experiments. The methodology was applied in a controlled setting to evaluate concept generation procedures for flexibility using a simplified real estate development design problem. Ninety experienced graduate students split among thirty-two collaborative design teams participated in these experiments.

The proposed methodology is an important component of an experimental platform to evaluate early conceptual design procedures thoroughly, rigorously, and efficiently in controlled laboratory. It complements the body of existing methodologies relying mostly on subjective expert performance assessments, or other objective metrics that do not explicitly measure anticipated performance. The platform can be used, for instance, to determine the potential of a particular design procedure before beginning a lengthy and deeper field study involving application in industry.

The paper addresses a need outlined by Frey et al. (2006), Reich (2010), and Tomiyama et al. (2009) for more tools and methodologies enabling rigorous evaluation of design procedures. It provides grounds for future research in terms of comparing design procedures to one another, such as Pahl and Beitz, Robust Design, and TRIZ. Other computer-aided tools can be used to develop a wider portfolio like CAD software, computer-assisted sketching, and optimization algorithms. Rapid prototyping should be considered as an avenue to complement computer modeling techniques suggested here.

Acknowledgments

References

Altshuller, G. 1973. *The Innovation Algorithm*. Worcester, MA, United States: Technical Innovation Center.

Amram, M., and N. Kulatilaka. 1999. *Real Options: Managing Strategic Investment in an Uncertain World*. Cambridge, MA, United States: Harvard Business School Press.

Audi, R. 1998. *Epistemology: A Contemporary Introduction to the Theory of Knowledge*. New York, NY, United States: Routledge.

Babajide, A., R. de Neufville, and M.-A. Cardin. 2009. Integrated Method for Designing Valuable Flexibility in Oil Development Projects. *SPE Projects, Facilities, and Construction* 4:3-12.

Bender, B., and L. T. M. Blessing. 2003. Re-Interpretation of Conceptualization: A Contribution to the Advance of Design Theory. In *Human Behaviour in Design: Individuals, Teams, Tools*, edited by U. Lindemann. Berlin Heidelberg, Germany: Springer-Verlag.

Black, F., and M. Scholes. 1973. The Pricing of Options and Corporate Liabilities. *Journal of Political Economy* 81 (3):637-654.

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

Briggs, R. O., B. A. Reinig, and G.-J. de Vreede. 2006. Meeting Satisfaction for Technology-Supported Groups: An Empirical Validation of a Goal Attainment Model. *Small Group Research* 37 (6):585-611.

Cardin, M.-A. 2011. Quantitative Performance-Based Evaluation of a Procedure for Flexible Design Concept Generation. Doctoral Dissertation in Engineering Systems, Massachusetts Institute of Technology, Cambridge, MA, United States.

Cardin, M.-A., G. L. Kolfschoten, R. de Neufville, D. D. Frey, O. L. de Weck, and D. M. Geltner. 2012. Empirical Evaluation of Procedures to Generate Flexibility in Engineering Systems and Improve Lifecycle Performance. *Submitted for review to Research in Engineering Design*.

Copeland, T., and V. Antikarov. 2003. *Real Options: A Practitioner's Guide*. New York, NY, United States: Thomson Texere.

Cox, J. C., S. A. Ross, and M. Rubinstein. 1979. Options Pricing: A Simplified Approach. *Journal of Financial Economics* 7 (3):229-263.

de Neufville, R., and A. Odoni. 2003. *Airport Systems: Planning, Design, and Management*. New York, NY, United States: McGraw-Hill Companies Inc.

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

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

Dixit, A. K., and R.S. Pindyck. 1994. *Investment under Uncertainty*. NJ, United States: Princeton University Press.

Eckert, C. M., O. L. de Weck, R. Keller, and P. J. Clarkson. 2009. Engineering Change: Drivers, Sources and Approaches in Industry. In *17th International Conference on Engineering Design*. Stanford, CA, United States.

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

Fisher, R. A. 1935. *The Design of Experiment*. New York, NY, United States: Hafner.

Frey, D. D., and C. L. Dym. 2006. Validation of Design Methods: Lessons from Medicine. *Research in Engineering Design* 17:45-57.

Frey, D. D., and X. Li. 2010. Model-Based Validation of Design Methods. Cambridge, MA, United States: Massachusetts Institute of Technology.

Fricke, E., and A. P. Schulz. 2005. Design for Changeability (DfC): Principles to Enable Changes in Systems Throughout Their Entire Lifecycle. *Systems Engineering* 8 (4):342-359.

Fumarola, M., M. Seck, and A. Verbraeck. 2010. A DEVS Component Library for Simulation-based Design of Automated Container Terminals. Netherlands: Delft University of Technology.

Geltner, D., and M.-A. Cardin. 2008. *Engineering Economy Module*. Cambridge, MA, United States: Massachusetts Institute of Technology. Lecture Notes.

Goldschmidt, G., and M. Weil. 1998. Contents and Structure in Design Reasoning. *Design Issues* 14 (3):85-100.

Guilford, J. P. 1959. *Personality*. New York, NY, United States: McGraw-Hill.

Guma, A., J. Pearson, K. Wittels, R. de Neufville, and D. Geltner. 2009. Vertical Phasing as a Corporate Real Estate Strategy and Development Option. *Journal of Corporate Real Estate* 11 (3):144-157.

Hassan, R., and R. de Neufville. 2006. Design of Engineering Systems Under Uncertainty via Real Options and Heuristic Optimization. In *Real Options Conference*. New York, NY, United States.

Hazelrigg, G. A. 1998. A Framework for Decision-Based Engineering Design. *Journal of Mechanical Design* 120:653-658.

Hevner, A. R., S. T. March, J. Park, and S. Ram. 2004. Design Science in Information Systems Research. *Management Information Systems Quarterly* 28 (1):75-105.

iCampus MIT-Microsoft Alliance. 2010. *Magic Paper: Introducing Natural Interaction* 2010 [cited June 25 2010]. Available from http://icampus.mit.edu/magicpaper/.

Itô, K. 1951. On Stochastic Differential Equations. *Memoirs, American Mathematical Society* 4:1-51.

Jablonowski, C., C. Wiboonskij-Arphakul, and M Neuhold. 2008. Estimating the Cost of Errors in Estimates Used During Concept Selection. *SPE Projects, Facilities, and Construction* 3:1-6.

Johnson, B., and L. Christensen. 2004. *Educational Research: Quantitative and Qualitative Approaches*. 2nd ed. Boston, MA, United States: Allyn and Bacon.

Kolfschoten, G., R. O. Briggs, G. P. J. Duivenvoorde, and G.-J. de Vreede. 2009. Practitioners vs. Facilitators a Comparison of Participant Perceptions on Success. In *42nd Annual Hawaii International Conference on System Sciences*. Big Island, HI, United States: IEEE Computer Society.

Kronenburg, R. 2007. *Flexible - An Architecture that Responds to Change*. London, United Kingdom: Laurence King.

Kurtoglu, T., M. I. Campbell, and J. S. Linsey. 2009. An Experimental Study on the Effects of a Computational Design Tool on Concept Generation. *Design Studies* 30:676-703.

Lin, J. 2009. Exploring Flexible Strategies in Engineering Systems Using Screening Models – Applications to Offshore Petroleum Projects. Doctoral Dissertation in Engineering Systems, Massachusetts Institute of Technology, Cambridge, MA, United States.

Lindemann, U. 2003. *Human Behaviour in Design: Individuals, Teams, Tools*. Berlin Heidelberg, Germany: Springer-Verlag.

Linsey, J. S., M. G. Green, M. van Wie, K. L. Wood, and R. Stone. 2005. Functional Representations in Conceptual Design: A First Study in Experimental Design and Evaluation. Paper read at Proceedings of the American Society for Engineering Education Annual Conference and Exposition.

Linsey, J. S., I. Tseng, K. Fu, J. Cagan, J. L. Wood, and C. Schunn. 2010. A Study of Design Fixation, Its Mitigation and Perception in Engineering Design Faculty. *Journal of Mechanical Design* 132.

Mikaelian, T., D. J. Nightingale, D. H. Rhodes, and D. E. Hastings. 2011. Real Options in Enterprise Architecture: A Holistic Mapping of Mechanisms and Types for Uncertainty Management. *IEEE Transactions on Engineering Management* 54 (3):457-470.

Milton, J. S., and J. C. Arnold. 1990. *Introduction to Probability and Statistics – Principles and Applications for Engineering and the Computing Sciences*. United States: McGraw-Hill.

Minai, A. A., D. Braha, and Y. Bar-Yam. 2006. Complex Engineered Systems: A New Paradigm. In *Complex Engineered Systems: Science Meets Technology*, edited by A. A. Minai, D. Braha and Y. Bar-Yam. Netherlands: Springer.

Mullen, B., C. Johnson, and E. Salas. 1991. Productivity Loss in Brainstorming Groups: A Meta-Analytic Integration. *Basic and Applied Social Psychology* 12 (1):3-23.

Myers, S. C. 1977. Determinants of Corporate Borrowing. *Journal of Financial Economics* 5:147-175.

Nelson, B. A., J. O. Wilson, and D. Rosen. 2009. Refined Metrics for Measuring Ideation Effectiveness. *Design Studies* 30:737-743.

Nembhard, H. B., L. Shi, and M. Aktan. 2005. A Real-Options-Based Analysis for Supply Chain Decisions. *IIE Transactions* 37:945-956.

Nunamaker, J. F., R. O. Briggs, D. D. Mittleman, D. R. Vogel, and P. A. Balthazard. 1997. Lessons from a Dozen Years of Group Support Systems Research: A Discussion of Lab and Field Findings. *Journal of Management Information Systems* 13 (3):163-207.

Osborn, A. F. 1957. *Applied Imagination: Principles and Procedures of Creative Thinking.* 1st ed. New York, NY, United States: Scribeners and Sons.

Pahl, G., and W. Beitz. 1984. *Engineering Design.* London, United Kingdom: Springer Design Council.

Pitman, E. J. G. 1937. Significance Tests which May Be Applied to Samples from Any Population. *Royal Statistical Society Supplement* 4:119-130 (Part I) and 225-32 (Part II).

Reich, Y. 2010. My Method is Better! [Editorial]. *Research in Engineering Design* (April).

Reinig, B. A., R. O. Briggs, and J. F. Nunamaker. 2007. On the Measurement of Ideation Quality. *Journal of Management Information Systems* 23 (4):143-161.

Rothwell, G. 2006. A Real Options Approach to Evaluating New Nuclear Power Plants. *Energy Journal* 27:37-54.

Santanen, E. L., R. O. Briggs, and G.-J. de Vreede. 2004. Causal Relationships in Creative Problem Solving: Comparing Facilitation Interventions for Ideation. *Journal of Management Information Systems* 20 (4):167-197.

Shah, J. J., S. V. Kulkarni, and N. Vargas-Hernandez. 2000. Evaluation of Idea Generation Methods for Conceptual Design: Effectiveness Metrics and Design of Experiments. *Journal of Mechanical Design* 122:377-384.

Shah, J. J., N. Vargas-Hernandez, and S. M. Smith. 2002. Metrics for Measuring Ideation Effectiveness. *Design Studies* 24 (2):111-134.

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

Suh, E. S., O. L. de Weck, and D. Chang. 2007. Flexible Product Platforms: Framework and Case Study. *Research in Engineering Design* 18:67-89.

Suh, N. P. 1990. *The Principles of Design.* United Kingdom: Oxford University Press.

Summers, J. D., S. Anandan, and S. Teegavarapu. 2009. Introduction of Design Enabling Tools: Development, Validation, and Lessons Learned. In *Tools for Innovation*, edited by A. B. Markman and K. L. Wood. United Kingdom: Oxford University Press.

Taguchi, G. 1987. *The System of Experimental Design Engineering Methods to Optimize Quality and Minimize Cost*. Vol. 1 and 2. Dearborn, MI, United States: American Supplier Institute.

Tomiyama, T., P. Gu, Y. Jin, D. Lutters, Ch. Kind, and F. Kimura. 2009. Design Methodologies: Industrial and Educational Applications. *CIRP Annals – Manufacturing Technology* 58:543-565.

Trauth, E. M., and L. M. Jessup. 2000. Understanding Computer-Mediated Discussions: Positivist and Interpretive Analyses of Group Support System Use. *Management Information System Quarterly* 24 (1):43-79.

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

Ullman, D., S. Wood., and D. Craig. 1989. The Importance of Drawing in the Mechanical Design Process. *Computer and Graphics* 14 (2):263-274.

Walker, W. E., S. A. Rahman, and J. Cave. 2001. Adaptive Policies, Policy Analysis, and Policy-Making. *European Journal of Operational Research* 128:282-289.

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

Ward, A. C., J. K. Liker, J. J. Cristiano, and D. K. II Sobeck. 1995. The Second Toyota Paradox: How Delaying Decisions Can Make Better Cars Faster. *Sloan Management Review* 36 (3):43-61.

Warr, A., and E. O'Neill. 2005. Understanding Design as a Social Creative Process. In *Proceedings of the 5th Conference on Creativity and Cognition*. London, United Kingdom.

Welch, W. J. 1990. Construction of Permutation Tests. *Journal of American Statistical Association* 85:693-698.

Wu, C. F. J., and M. Hamada. 2000. *Experiments: Planning, Analysis, and Parameter Design Optimization*. New York, NY, United States: John Wiley and Sons.

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

Repeated Author. 2009. Observations on Concept Generation and Sketching in Engineering Design. *Research in Engineering Design* 20:1-11.

Yang, Y. 2009. A Screening Model to Explore Planning Decisions in Automotive Manufacturing Systems Under Demand Uncertainty. Doctoral Dissertation in

Engineering Systems, Massachusetts Institute of Technology, Cambridge, MA, United States.

Zhang, S. X., and V. Babovic. 2011. A Real Options Approach to the Design and Architecture of Water Supply Systems Using Innovative Water Technologies Under Uncertainty. *Journal of Hydrodynamics*:In press.

Appendix

Prompting Procedure

Uncertainty

What are the major sources of uncertainty affecting the future performance of this system? Examples:
- Exogenous uncertainties (e.g. demand markets, natural catastrophes, etc.)
- Endogenous uncertainties (e.g. technology failure rates, etc.)
- Scenarios where things go really bad (e.g. prices drop, economic crisis, etc.)
- Scenarios where things go really well (e.g. demand rises suddenly, etc.)

Flexibility

What flexible strategies would enable the system to change and adapt if the uncertainty scenarios you just discussed occur during operations? Examples:
- Defer the initial capital investment until favorable market conditions
- Abandon the project to get out of bad, negative market situations
- Invest in R&D to support growth and future opportunities
- Phase capacity deployment over time instead of deploying initially all capacity at once
- Alter operating scale by expanding or reducing production capacity depending on market conditions
- Switch production output and/or input depending on observed demand

Design

How should you prepare, engineer, and design this particular system to enable the flexibilities you just discussed? Think about how to best engineer the system so it can react to:
- Negative or bad scenarios (e.g. start with a smaller initial design, and reduce risk of over-capacity and losses)
- Positive or good scenarios (e.g. engineer ability to switch product output easily, write legal contract to enable physical expansion later on if needed)
- Completely unexpected scenario (e.g. plan ahead for emergency procedure in case of hurricane)

Management

How should you manage and decide when it is appropriate to use, or exercise, the flexibilities in this system? Examples:
- If demand is lower than capacity for two years, I will shutdown operations for 6 months
- If market price gets above a certain threshold, I will expand production capacity

Biographies

Michel-Alexandre Cardin is Assistant Professor of Industrial and Systems Engineering at the National University of Singapore (NUS). His research focuses on the design of engineering systems for uncertainty and flexibility. Prior to joining NUS, he received a Ph.D. in Engineering Systems and Master of Science in Technology and Policy from the Massachusetts Institute of Technology (MIT) in the United States. He received a Master of Applied Sciences in Aerospace Engineering from the University of Toronto, and an Honours Bachelor's degree in Physics from McGill University in Canada.

Gwendolyn L. Kolfschoten is Assistant Professor at Delft University of Technology (TU Delft) in the Netherlands. Her research focuses on how teams, experts or stakeholders can effectively collaborate, and use collaboration support tools, in problem solving tasks. She is an experienced facilitator of Group Support System workshops and sessions, having worked with numerous public and private organizations. She has organized minitracks and tutorials at HICSS for the past five years, and organized the GDN 2010 conference in Delft. Her research has been published in the Journal of Management Information Systems, International Journal of Human-Computer Studies, Journal of the AIS, Computers and Education, and Group Decision and Negotiation.

Daniel D. Frey is Associate Professor of Mechanical Engineering and Engineering Systems at MIT. His research concerns robust design of engineering systems. He is actively involved in design of engineering devices for the developing world. He is co-director of the Singapore-MIT International Design Center. He is a member of the American Society of Mechanical Engineers (ASME), the American Statistical Association (ASA), the International Council on Systems Engineering (INCOSE), and the American Society of Engineering Education (ASEE). He holds a Ph.D. in Mechanical Engineering from MIT, an MS in Mechanical Engineering from the University of Colorado, and a BS in Aeronautical Engineering from Rensselaer Polytechnic Institute.

Olivier L. de Weck is Associate Professor of Aeronautics and Astronautics and Engineering Systems at MIT. He was born in Switzerland and holds degrees in industrial engineering from ETH Zurich (1993) and aerospace systems engineering from MIT (2001). Before joining MIT he was a liaison engineer and later engineering program manager on the F/A-18 aircraft program at McDonnell Douglas (1993-1997). Prof. de Weck is a leader in systems engineering research. He focuses on how complex man-made systems such as aircraft, spacecraft, automobiles, printers and critical infrastructures are designed and how they evolve over time. His main emphasis is on strategic properties that have the potential to maximize lifecycle value (a.k.a the "ilities").

Richard de Neufville is Professor of Civil and Environmental Engineering and Engineering Systems at MIT. He is an engineer and system designer. His research and teaching focuses on inserting flexibility into the design of technological systems. His book *Flexibility in Engineering Design* (co-authored with Stefan Scholtes of the University of Cambridge) is being published by the MIT Press in 2011. He has a Ph.D. from MIT and a Dr. h.c. from TU Delft. In 1965 he served as a first White House Fellow for President Lyndon Johnson. He is known worldwide for his applications in Airport Systems Planning, Design, and Management. He has been associated with major airport projects in North America, Europe, Asia, Australia – as well as others in Africa and Latin America.

David Geltner is the Director of Research for the Center for Real Estate, as well as the George Macomber Professor of Real Estate Finance in the Department of Urban Studies and Planning at MIT. He served as Director of the MIT/CRE 2003-08, in which position he was responsible for the overall planning and management of the Center, and headed MIT's Master of Science in Real Estate Development (MSRED) program. His research has focused on real estate investment performance measurement and the related areas of asset valuation and private asset market functioning, fields in which he has published widely in academic and trade journals. He received his PhD in 1989 from the MIT Department of Civil & Environmental Engineering in the field of infrastructure finance and economics. He also has degrees in urban studies from Carnegie-Mellon University and the University of Michigan