QUANTITATIVE PERFORMANCE-BASED EVALUATION OF A PROCEDURE FOR FLEXIBLE DESIGN CONCEPT GENERATION

by
Michel-Alexandre Cardin

Honors B.Sc. in Physics, McGill University, 2001
M.A.Sc. in Aerospace Science and Engineering, University of Toronto, 2003
S.M. in Technology and Policy, Massachusetts Institute of Technology, 2007

Submitted to the Engineering Systems Division
in Partial Fulfillment of the Requirements for the Degree of
Doctor of Philosophy in Engineering Systems
at the Massachusetts Institute of Technology
February 2011

Signature of Author…………………………………………………………………………………………..
Engineering Systems Division
January 7th, 2011

Certified by…………………………………………………………………………………………………..
Prof. Richard de Neufville
Professor of Engineering Systems and Civil and Environmental Engineering
Thesis Supervisor

Certified by…………………………………………………………………………………………………..
Prof. Daniel D. Frey
Associate Professor of Engineering Systems and Mechanical Engineering
Thesis Committee Member

Certified by…………………………………………………………………………………………………..
Prof. David M. Geltner
Professor of Engineering Systems and Real Estate Finance
Thesis Committee Member

Certified by…………………………………………………………………………………………………..
Dr. Ir. Gwendolyn L. Kolfschoten
Assistant Professor of Collaboration Engineering, Delft University of Technology
Thesis Committee Member

Certified by…………………………………………………………………………………………………..
Prof. Olivier L. de Weck
Associate Professor of Engineering Systems and Aeronautics and Astronautics
Thesis Committee Member

Accepted by…………………………………………………………………………………………………..
Prof. Nancy Leveson
Professor of Engineering Systems and Aeronautics and Astronautics
Chair, Engineering Systems Division Education Committee

À tous ceux qui me sont chers

QUANTITATIVE PERFORMANCE-BASED EVALUATION OF A PROCEDURE FOR FLEXIBLE DESIGN CONCEPT GENERATION

by
Michel-Alexandre Cardin

Submitted to the Engineering Systems Division on January 7th, 2011
in Partial Fulfillment of the Requirements for the Degree of
Doctor of Philosophy in Engineering Systems

Abstract

This thesis presents an experimental methodology for objective and quantitative design procedure evaluation based on anticipated lifecycle performance of design concepts, and a procedure for flexible design concept generation. The methodology complements existing evaluation methodologies by measuring anticipated performance via efficient computer-modeling techniques. The procedure, in contrast to others, stimulates flexible design concept generation by packaging a short lecture on flexibility, and a prompting ideation mechanism.

Controlled collaborative experiments had participants suggest alternative solutions to a design problem under different treatment conditions. Experimental conditions used the procedure for flexibility, while control conditions relied on prior training in science and engineering only, and free undirected ideation. Measures included the quantity of flexible design concepts generated, anticipated economic performance improvements compared to a benchmark design, participants' subjective impressions of satisfaction with the process and results, and results quality assessments. Seventy-one designers divided among twenty-six teams performed the experiments involving a simplified real estate infrastructure design problem.

Application of the methodology demonstrated effective and efficient evaluation of the design procedure based on anticipated performance of design concepts. The lecture and prompting mechanism significantly improved anticipated performance compared to the benchmark design, by nearly thirty-six percent. The prompting mechanism significantly improved generation of valuable flexible design concepts. Lecturing improved significantly user satisfaction with the process and results, as well as results quality assessments. Even though prompting demonstrably improved anticipated performance and concept generation, it had no effect on participants' satisfaction with the process and results – unless combined with the lecture. Also, prompting did not lead participants to expect better results quality. This demonstrates the need for thorough and rigorous procedure evaluations based both on subjective user impressions and objective quantitative measurements. A preliminary analysis suggests that the proposed experimental platform can be used to study the influence of uncertainty and flexibility related words on discussion content, although more work is necessary to fully validate the approach.

Thesis supervisor: Richard de Neufville
Title: *Professor of Engineering Systems and Civil and Environmental Engineering*

Acknowledgments

I would like to thank gratefully Prof. Richard de Neufville for his mentorship, intellectual support, and for the extremely enriching opportunities he gave me during my journey at MIT since September 2005. Following your footsteps, and working with you on a daily basis has been the most engaging, rewarding, and determining experience in my life – besides being a great honor. I have been most lucky to receive and count on your advice on how to think about research, as not many students are this fortunate. Thank you for being so interested in my success, and for creating all these opportunities – both professional and academic – over the last five years. I hope I will continue honoring the tradition and legacy you have set forward.

I would like to thank my doctoral committee members Profs. Daniel Frey, David Geltner, Gwendolyn Kolfschoten, and Olivier de Weck for their intellectual support, generosity, and time. Thank you Dan for providing the backbone intellectual framework behind this work, which involves rigorous evaluation and validation of design procedures. Thank you for your insightful guidance, for your help with the experimental and statistical analyses, and for inviting me to join the SUTD team last summer. Thank you David for exercising such rigorous and critical thinking about the economic side of my thesis, and for all the exposure you gave me to the world of real estate development and economics. Due to that spark you helped create, I am looking forward to use real estate as one of the important application areas for my academic career. Thank you Gwen for welcoming me to your team last February in Delft. This has been one of the most important phases of my doctoral work. This thesis has essentially been shaped up due to our stimulating interactions; there is much influence of your work here. Finally, thank you Oli for your creative advice on my research, for seeding some of the key ideas behind this thesis, and for encouraging me to pursue this experimental approach. Thank you as well for allowing me to interact with your students during experiments, this has been vital to the research. Thank you all for being such a supportive, intellectually challenging and stimulating team. I am looking forward to future collaborations with you, as I enter a new stage of my career.

I am most grateful to all the groups and organizations that generously provided funding over the last five and half years. This includes the Engineering Systems Division, the MITRE Corporation, the MIT Center for Real Estate, the National Science and Engineering Research Council of Canada, the Fonds Québécois de la Recherche sur la Nature et les Technologies, the Systems Design and Management program, the MIT Portugal Program, the SUTD-MIT International Design Center, the Cambridge-MIT Institute, and the Fondation Desjardins.

Thank you to my close collaborators at MIT and overseas for the help and support provided at various stages. I very much look forward our continued interactions. From the University of Cambridge, many thanks to William Nuttall from the Judge Business School for introducing me to the exciting topic of innovative nuclear technology development, and for welcoming me to Cambridge during two summers. It is always a pleasure working with you Bill, I hope we will resume working together soon. Thank you to Geoff Parks and Steven Steers from the Department of Engineering for extending my views and allowing me to apply various flexibility valuation techniques to ADSR technology. Many thanks to Alexander Verbraeck from Delft University of Technology for demonstrating such excitement about my research proposal while in Delft, and supporting my experiments by having your own students involved. Many thanks to Tony Ciochetti from the Center for Real Estate for having me involved in courses at the Professional Development Institute. This has been a greatly enjoyable experience. Thank you to Beth Milnes from ESD and to Sydney Miller from TPP for making my time at MIT so incredibly smooth on the administrative side. Thank you to Dava Newman and Frank Field from TPP for giving me this first chance at MIT, and mentoring me throughout this experience.

Thanks to all my friends who made this time the greatest: Yves S. Boussemart, Laure-Anne Ventouras, Arzum "the Legend" Akkas, Tommy Rand-Nash, Rhonda "Ms." Jordan, Matt and Sarah Silver, Marie-Ève Rancourt, Blandine Antoine, Lara "Ms." Pierpoint and Tristan Weir, Kat Donnelly, Jason Leaf, Rebecca "Ms." Weiss, Nick "the Nuge" Nugent, Regina Clewlow, Katherine "Ms." Dykes, Erica Gralla and Sam, Valerie Karplus, Dave Keith, Dan Livengood, Brian and Tuco Mekdeci, Sidharth Rupani, Vivek Sakhrani, Nidhi Santen, Lynette Cheah, Yingxia Yang, Mike Hanowsky, Kate Steel, Carlos Osorio, Jason Bartolomei, Chris Lawson, Catarina Bjelkengren, John Pendray, John MacNeil, Vishal Doshi, Anna Nicholson, Hsiao-Lan Lin, Gayle Ratliff, JunJay Tan, Nicole Jordan, Pragnya Alekal, Samuel Chow, Cécilia Gérard, Sala McGuire, Laura Escudero, Ellie Ereira, Alda Mendes, Sabra Loewus, Laurent Guérin, Alexandra Bell, Stephan Feilhauer, Jérémie Berteau and Raphaële Thévenin, Romain Lacombe, Sébastien Uzel, Marie-Claude "MC" Nadeau, Elza Brunelle-Yeung, Mélissa Ouellet, Ramis Movassagh, Anshul Mohnot, Ambika Goel, Allin Resposo, Jen Tran, Audrey Simon and Matt Laplante, Ali Motamedi, Jonathan Birge, Michael Koeris, Roberto Franco, Ian Coveny, Timea Pal, Caroline Nowlan, Dave Bradwell, Amandine Martin, Howard Yue, and Linda Gregorio.

Thanks to all my hockey buddies and friends John McBean, Nate Tedford, Kenzie MacIsaac, Desmond Adler, Jesse Edgerton, Dave Shirokoff, Tim DeSmet, Mike Georgas, Dave Tax, Ken Hu, Jeff Defoe, Andrew Piekos, Scott McKay, Dave Allen, Gordon Sims, and Sebastian Stirling. You guys contributed to a slimmer me at the end of this trip, I am most grateful for this!

Merci à ma belle et douce Manon Blouin qui a su me comprendre, me supporter, m'aimer et me donner tout ce dont j'ai eu besoin pendant la dernière moitié de ce trajet. Ton sourire et ta joie de vivre ont été les remèdes à beaucoup de maux ma Belle. J'ai très hâte de m'embarquer avec toi dans cette nouvelle aventure à plus de douze fuseaux horaires d'ici !

J'aimerais garder mes derniers remerciements pour ma famille qui m'a supporté et encouragé pendant toutes ces années. Ça n'a pas toujours été facile de suivre mes changements d'adresse mais je vous suis infiniment reconnaissant de l'avoir fait et de ne pas avoir abandonné mon cas ! Merci à mon père Michel Cardin, ma mère Marjolaine Coutu, mes sœurs Marie-Ève et Sophie-Catherine, mon beau-frère de toujours Éric Chouteau et à la toute dernière venue, la petite Charlotte. Comme vous pouvez le voir mes chers parents, j'ai trouvé le moyen de combiner vos deux passions en une : l'immobilier et l'enseignement ! Mes chères sœurs, je me sens choyé d'être près de deux si merveilleuses femmes, continuons à profiter d'être ensemble ou aux quatre coins du monde encore pendant plusieurs années ! J'espère qu'Éric continuera de nous enchanter avec ses blagues et ses barbots encore longtemps. Espérons que Charlottine continuera aussi de nous démontrer ses prouesses de miroir – et autres trucs rigolos à venir !

Table of Contents

ABSTRACT **3**

ACKNOWLEDGMENTS **4**

NOMENCLATURE **10**

LIST OF ACRONYMS 10
LIST OF SYMBOLS 11

LIST OF FIGURES **13**

LIST OF TABLES **16**

CHAPTER 1 – INTRODUCTION **17**

1.1 MOTIVATIONS 18
1.1.1 Current Considerations of Uncertainty and Flexibility in Design *18*
1.1.2 Current Evaluation Methodologies for Design Procedures *21*
1.2 INTENDED AUDIENCE AND APPLICATION 22
1.3 RESEARCH APPROACH 24
1.4 THESIS CONTENT SUMMARY 25

CHAPTER 2 – LITERATURE REVIEW **26**

2.1 ENGINEERING DESIGN RESEARCH 26
2.1.1 Proposed Order in Engineering Design Research *26*
2.1.2 Thesis Positioning Within Engineering Design Research *30*
2.2 CURRENT DESIGN PROCEDURES 30
2.2.1 Established Design Procedures *31*
2.2.2 Design Procedures for Concept Generation *43*
2.2.3 Design Procedures for Flexibility *44*
2.3 DESIGN PROCEDURE EVALUATION METHODOLOGIES 62
2.3.1 Categories of Methodologies *62*
2.3.2 Evaluation Metrics *63*
2.3.3 Example Case Studies *65*
2.3.4 Example Controlled User Studies *66*
2.3.5 Example Protocol Studies *68*
2.4 RESEARCH OPPORTUNITIES 68
2.4.1 Related to Established Design Procedures *68*
2.4.2 Related to Design Evaluation and Selection *69*
2.4.3 Related to Concept Generation Procedures *70*

2.4.4 Related to Design for Flexibility *71*
2.4.5 Related to Design Procedure Evaluation Methodologies *74*
2.5 Anticipated Contributions 76
2.5.1 A Procedure for Flexible Design Concept Generation *76*
2.5.2 A Methodology for Quantitative Performance-Based Design Procedure Evaluation *77*

CHAPTER 3 – RESEARCH QUESTIONS AND APPROACHES 79

3.1 Area 1: Procedure for Flexible Design Concept Generation 79
3.2 Area 2: Performance-Based Design Procedure Evaluation 81
3.3 Area 3: Design Procedure Influence on Discussion Content...83

CHAPTER 4 – DESIGN PROCEDURE FOR FLEXIBILITY..85

4.1 Education Factor (*E*) 86
4.2 Ideation Factor (*I*)...87
4.3 Using the Design Procedure for Flexibility in Engineering Practice 90
4.3.1 A Complementary Design Procedure for Industry..*90*
4.3.2 Required Training for Application *91*
4.3.3 Complementary Tools *92*

CHAPTER 5 – EXPERIMENTAL METHODOLOGY 94

5.1 Step 1: Design Problem Description 95
5.2 Step 2: Computer Model 95
5.3 Step 3: Online Group-Support System Interface 96
5.4 Step 4: Data Collection 97
5.5 Step 5: Analysis 99
5.5.1 Coding Analysis and Response Measurements *99*
5.5.2 Survey Analysis *100*
5.5.3 Content Analysis..*101*
5.5.4 Statistical Analysis *102*

CHAPTER 6 – SPECIFIC EXPERIMENTAL IMPLEMENTATION 105

6.1 Preliminary Setup...105
6.2 Step 1: Design Problem Description 106
6.3 Step 2: Computer-Model 108
6.4 Step 3: Online Group-Support System Interface 111
6.5 Step 4: Data Collection 112
6.6 Step 5: Analysis 113
6.6.1 Coding Analysis and Response Measurements *114*

6.6.2 Survey Analysis *116*
6.6.3 Content Analysis *117*
6.6.4 Statistical Analysis *119*

CHAPTER 7 – RESULTS **120**

7.1 IMPROVEMENT IN COMPLETE IDEAS (ΔC) 120
7.2 IMPROVEMENT IN GOOD IDEAS (ΔG) 123
7.3 IMPROVEMENT IN ENPV ($\Delta ENPV$) 124
7.4 IMPROVEMENT IN PROCESS SATISFACTION (ΔPS) 125
7.5 IMPROVEMENT IN RESULTS SATISFACTION (ΔRS) 127
7.6 IMPROVEMENT IN QUALITY ASSESSMENT (ΔQA) 128
7.7 IMPROVEMENT IN UNCERTAINTY INFLUENCE (ΔUI) 129
7.8 IMPROVEMENT IN FLEXIBILITY INFLUENCE (ΔFI) 131

CHAPTER 8 – FINDINGS AND DISCUSSION **133**

8.1 AREA 1: PROCEDURE FOR FLEXIBLE DESIGN CONCEPT GENERATION 133
8.2 AREA 2: PERFORMANCE-BASED DESIGN PROCEDURE EVALUATION 136
8.3 AREA 3: DESIGN PROCEDURE INFLUENCE ON DISCUSSION CONTENT 138
8.4 RESULTS VALIDITY AND LIMITATIONS 139
8.4.1 Internal Validity *139*
8.4.2 External Validity *144*

CHAPTER 9 – CONCLUSION **146**

9.1 EXTENDING CURRENT APPROACHES TO DESIGN 148
9.2 PURSUING REAL IMPACT ON DESIGN PRACTICE 149
9.3 FUTURE RESEARCH OPPORTUNITIES 150

BIBLIOGRAPHY **152**

APPENDIX A – SLIDES FOR LECTURE ON FLEXIBILITY **164**

APPENDIX B – PROMPTS FOR IDEATION MECHANISM **172**

APPENDIX C – SLIDES FOR DESIGN PROBLEM DESCRIPTION **174**

APPENDIX D – ASSUMPTIONS FOR DISCOUNTED CASH FLOW MODEL **178**

APPENDIX E – DEBRIEF MATERIAL AND EXAMPLE SURVEY QUESTIONS **184**

APPENDIX F – EXAMPLE ORIGINAL TRANSCRIPT CODING ANALYSIS **188**

APPENDIX G – SUMMARY OF COMPLETE IDEAS FROM TRANSCRIPT ANALYSIS **189**

APPENDIX H – TRANSCRIPT ANALYSIS AND RESPONSE MEASUREMENTS DATASET 190

APPENDIX I – COMPLEMENTARY RESPONSE MEASUREMENTS DATASETS 191

APPENDIX J – LIST OF FLEXIBILITY AND UNCERTAINTY RELATED WORDS 193

APPENDIX K – CODE FOR PERMUTATION-BASED STATISTICAL ANALYSIS 195

Nomenclature

List of Acronyms

Accelerator-Driven Subcritical Reactor (ADSR) 47
adaptive One Factor At a Time (aOFAT) . 61
Centering Resonance Analysis (CRA)...... 84
Computer-Aided Design (CAD) 77
Coupled-DSM (C-DSM)........................... 56
Cumulative Density Function (CDF)........ 52
Customer Attribute (CA) 37
Decision-Based Design (DBD)................. 34
Design for X (DfX) 34
Design Of Experiment (DOE) 32
Design Parameter (DP) 37
Design Structure Matrix (DSM) 40
Discounted Cash Flow (DCF)................... 52
Engineering Systems Matrix (ESM)......... 56
Expected NPV (ENPV)............................. 51
Failure Mode & Effect Analysis (FMEA) 35
Functional Requirement (FR) 37
General Design Theory (GDT) 31
General Linear Model (GLM) 102
Geometric Brownian Motion (GBM) 50
Group Support System (GSS)................... 66
Idea Generation (IG) 43
LINear ACcelerator (LINAC)................... 47
Multi-Attribute Tradespace Exploration (MATE) 55
Net Present Value (NPV).......................... 17
Opportunity Cost of Capital (OCC)........ 108
Process Variable (PV)............................... 37
Pugh Controlled Convergence (PuCC)..... 41
Quality Function Deployment (QFD)....... 17
Real Options Analysis (ROA) 47
Research and Development (R&D) 54
sensitivity DSM (sDSM) 56

List of Symbols

A_m = desirable design attribute m

α = Cronbach's alpha for inter-item reliability measure

β_0 = total mean of dataset

β_i = main effect for factor x_i

β_{ij} = interaction effect between factors x_i and x_j

BC_t = base construction cost at time t

C_{Ct} = total construction and sales cost at time t for condo-only design

CC_{Ct} = construction cost at time t for condo-only design

CC_{Ct}^S = stochastic construction cost at time t for condo-only design

CF_{Ct} = cash flow at time t for condo-only design

ΔC = improvement in quantity of complete design concepts between sessions 1 and 2

$\Delta ENPV$ = ENPV improvement of design concepts between sessions 1 and 2

ΔFI = flexibility influence improvement between sessions 1 and 2

ΔG = improvement in quantity of good design concepts between sessions 1 and 2

ΔP = improvement in performance design concepts between sessions 1 and 2

ΔPS = process satisfaction improvement between sessions 1 and 2

ΔQA = quality assessment improvement between sessions 1 and 2

ΔRS = results satisfaction improvement between sessions 1 and 2

ΔUI = uncertainty influence improvement between sessions 1 and 2

Δy = improvement for response y between sessions 1 and 2

dZ_t = standard Wiener process random variable

D_{Ct} = unit demand at time t for condo-only design

D_{Ct}^S = stochastic unit demand at time t for condo-only design

E = education factor $\in \{-1, +1\}$

$ENPV_C$ = Expected Net Present Value for condo-only design

FC_{Ct} = finishing cost at time t for condo-only design

g_{CC} = projected annual growth rate for construction cost

g_{CCt}^S = stochastic growth rate for construction cost at time t

g_D = projected annual growth rate for demand

$g_{Dt}{}^{S}$ = stochastic growth rate for demand at time t

g_P = projected annual growth rate for unit price

$g_{Pt}{}^{S}$ = stochastic growth rate for unit price at time t

g_{jk} = number of shortest paths connecting the j^{th} and k^{th} words

$g_{jk}(i)$ = number of paths between words j and k containing word i

K = number of questions/items used in a survey to study a construct of interest

K_{Ct} = planned capacity deployment at time t for condo-only design

I = ideation factor $\in \{-1, +1\}$

I_i^T = influence of word i in text T

M = maximum number of samples in Monte Carlo simulation

N = number of words in influence network

NPV_C = Net Present Value for condo-only design

P_{Ct} = unit price for condo at time t

$P_{Ct}{}^{S}$ = stochastic unit price for condo at time t

PV_{Ct} = present value of cash flow at time t for condo-only design

r = opportunity cost of capital or discount rate

R_{Ct} = revenue at time t for condo-only design

R_k = k^{th} response or replicate to a treatment group

σ_{CC} = uncertainty factor around annual construction cost projections

σ_D = uncertainty factor around annual demand projections

σ_P = uncertainty factor around annual unit price projections

$\sigma_X{}^2$ = variance of observed total survey scores

$\sigma_{Yi}{}^2$ = variance of survey scores obtained for item i among all participant responses

S_{mn} = score for attribute m of design alternative n

T = maximum time value t

W_m = weight of design attribute m

WS_n = weighted score for design alternative n

List of Figures

FIGURE 2.1: ORGANIZATION OF ENGINEERING DESIGN RESEARCH ACCORDING TO HORVÁTH (2004). 28
FIGURE 2.2: EXAMPLE KEPNER-TREGOE DECISION MATRIX. A_M IS A PARTICULAR DESIRABLE ATTRIBUTE, W_M IS THE WEIGHT OF AN ATTRIBUTE BASED ON DESIGNERS' PREFERENCES, S_{MN} IS THE SCORE OF A GIVEN CONCEPT RELATIVE TO ATTRIBUTE A_N, AND WS_N GIVES THE WEIGHTED SCORE OF A DESIGN CONCEPT CONSIDERING ALL ATTRIBUTES. 34
FIGURE 2.3: EXAMPLE HOUSE OF QUALITY PROCESS USED TO MAP CUSTOMER DEMAND QUALITY TO ATTRIBUTES TO A COMPANY'S PRODUCT DEVELOPMENT CAPABILITIES. 36
FIGURE 2.4: PROCESS FLOW IN AXIOMATIC DESIGN (TOMIYAMA ET AL., 2009). 38
FIGURE 2.5: PROCESS FLOW FOR PAHL AND BEITZ (TOMIYAMA ET AL., 2009). 39
FIGURE 2.6: A SCREENING MODEL SHOULD BE USED TO PRECEDE AND COMPLEMENT A MORE DETAILED ANALYSIS OF THE DESIGN (DE NEUFVILLE & SCHOLTES, 2011). 40
FIGURE 2.7: EXAMPLE DSM (BROWNING, 2001). ELEMENTS IN THE ROW PROVIDE INFORMATION TO ELEMENTS IN THE COLUMNS, WHILE ELEMENTS IN THE COLUMNS RECEIVE INFORMATION (OR DEPEND) FROM ELEMENTS IN THE ROWS..........41
FIGURE 2.8: OVERVIEW OF FORMAL IG METHODS (SHAH ET AL., 2000; 2002). 43
FIGURE 2.9: THE FLEXIBLE DESIGN OF THE BOEING B-52 STRATOFORTRESS'S ALLOWED ADAPTATION TO CHANGING WARFARE CONDITIONS. 45
FIGURE 2.10: PHASE I (A) AND PHASE II (B) OF THE FLEXIBLE HEALTH CARE SERVICE CORPORATION DEVELOPMENT PROJECT (GUMA, 2008). 46
FIGURE 2.11: EXAMPLE DECISION TREE COMPARING A FLEXIBLE ADSR DESIGN TO AN INFLEXIBLE DESIGN ALTERNATIVE (CARDIN ET AL., 2010). 49
FIGURE 2.12: COPPER PRICE EVOLUTION BASED ON THE BINOMIAL LATTICE APPROACH. 50
FIGURE 2.13: EXAMPLE BINOMIAL LATTICE DEPICTING THE FOLDING BACK EVALUATION OF THE ENPV OF CASH FLOWS TO EVALUATE THE FLEXIBILITY TO SHUT DOWN A COPPER MINE IN CASE COPPER PRICE IS TOO LOW (DE NEUFVILLE, 2010). LIGHT FIGURES ARE NEGATIVE ENPV OUTCOMES. 51
FIGURE 2.14: CORRESPONDING OPTIMAL DECISIONS FOR EACH OF THE DECISION NODE IN FIGURE 2.13. YES REPRESENTS THE STATES WHERE THE FLEXIBILITY TO SHUT DOWN SHOULD BE EXERCISED, WHILE NO REPRESENTS STATES WHERE THE MINE SHOULD REMAIN IN OPERATIONS. 51
FIGURE 2.15: EXAMPLE OF CDF – OR TARGET CURVE – USED TO ASSESS THE VALUE OF FLEXIBLE CAPACITY EXPANSION FOR A VERTICALLY BUILT PARKING GARAGE (DE NEUFVILLE & SCHOLTES, 2011). THE INFLEXIBLE "FIXED" DESIGNS ARE FIVE AND SIX LEVEL DESIGNS RESPECTIVELY, WHILE THE FLEXIBLE DESIGN ENABLES CAPACITY EXPANSION, STARTING WITH FOUR LEVELS. 53
FIGURE 2.16: VISUAL REPRESENTATION OF THE DBD METHOD TO EXPLORE THE DESIGN SPACE OF FLEXIBLE SYSTEMS (OLEWNIK & LEWIS, 2006). 55

Figure 2.17: Representation of the CPA algorithm (E. S. Suh et al., 2007). ΔE_{IN} and ΔE_{OUT} represent input and output changes respectively. CPI means Change Propagation Index. 56
Figure 2.18: Representation of the sDSM algorithm for processing and identifying interesting sources of flexibility (Kalligeros, 2006). Design variables are represented by x_I, functional requirements by FR_J. 57
Figure 2.19: Visual representation of the ESM approach (Bartolomei, 2007). 58
Figure 2.20: Example of dependencies among an enterprise view used in the C-DSM framework (Nightingale & Rhodes, 2007). 59
Figure 2.21: Suggested screening framework suited for flexibility analysis (Lin, 2009). 61
Figure 2.22: Representation of aOFAT to explore the design space (Frey & Wang, 2006). 61
Figure 5.1: Flow chart summarizing the experimental methodology. 94
Figure 5.2: Suggested pretest-posttest experimental structure to control for inherent creativity levels, and for possible prior experience with the design procedure of interest. 98
Figure 5.3: Example distribution of simulated coefficients β_I for factor X_I main effects, obtained from five thousand random permutations of the original dataset of a hypothetical response ΔY. The location of the test statistic $\beta_I = 0.75$ is shown as the vertical dashed line. 103
Figure 6.1: Example figures provided for mental conceptualization of the real estate development problem assigned to participants (http://www.northpointcambridge.com). 107
Figure 6.2: DCF model measuring NPV for the real estate development design problem. 109
Figure 6.3: CDFs or "target curves" for the condo-only, apartment (apts)- only, and flexible designs. The curves show NPV outcomes from simulations of the DCF model in Excel®, as well as ENPV (vertical dashed lines). The dark curves for the inflexible designs are very similar, thus almost indistinguishable. The light curve represents the flexible case. 110
Figure 6.4: Example online GSS interface by GroupSystems® used in this study. 111
Figure 7.1: Mean plots for ΔC, for all treatments. The lower light curve depicts the means with prior training only (factor E, level -1). The upper dark curve depicts the means when a lecture on flexibility is provided (factor E, level $+1$). 122
Figure 7.2: Mean plots for ΔG, for all four treatments. The lower light curve depicts the means with prior training only (factor E, level -1). The upper dark curve depicts the means when a lecture on flexibility is provided (factor E, level $+1$). 124
Figure 7.3: Mean plots for $\Delta ENPV$, for all treatments. The lower light curve depicts the means with prior training only (factor E, level -1). The upper dark curve depicts the means when a lecture on flexibility is provided (factor E, level $+1$). 125
Figure 7.4: Mean plots for ΔPS, for all four treatments. The lower light curve depicts the means with prior training only (factor E, level -1). The upper dark curve depicts the means when a lecture on flexibility is provided (factor E, level $+1$). 126

Figure 7.5: Mean plots for ΔRS, for all four treatments. The lower light curve depicts the means with prior training only (factor E, level -1). The upper dark curve depicts the means when a lecture on flexibility is provided (factor E, level $+1$). 128
Figure 7.6: Mean plots for ΔQA, for all four treatments. The lower light curve depicts the means with prior training only (factor E, level -1). The upper dark curve depicts the means when a lecture on flexibility is provided (factor E, level $+1$). 129
Figure 7.7: Mean plots for ΔUI, for all four treatments. The light curve depicts the means with prior training only (factor E, level -1). The dark curve depicts the means when a lecture on flexibility is provided (factor E, level $+1$). 130
Figure 7.8: Mean plots for ΔFI, for all four treatments. The lower light curve depicts the means with prior training only (factor E, level -1). The upper dark curve depicts the means when a lecture on flexibility is provided (factor E, level $+1$). 132
Figure 8.1: Conceptualization of the spectrum of "amount of direction" provided in prompting mechanisms to stimulate creativity. The extremes range from "no direction" to "complete direction" where answers are effectively given away. 143

List of Tables

TABLE 2.1: EXAMPLES OF WELL-KNOWN PROCEDURES FOR EARLY CONCEPTUAL DESIGN ACTIVITIES. 33
TABLE 2.2: EXAMPLE DESIGN PROCEDURES FOR FLEXIBILITY. 48
TABLE 2.3: EXAMPLE STUDIES MAKING USE OF DIFFERENT METRICS FOR EVALUATION OF PROCEDURES FOR DESIGN CONCEPT GENERATION. 64
TABLE 4.1: 2 X 2 DOE SETUP FOR EVALUATING THE DESIGN PROCEDURE FOR FLEXIBILITY. 86
TABLE 6.1: 2 X 2 DOE SETUP FOR EVALUATING THE DESIGN PROCEDURE FOR FLEXIBILITY. 105
TABLE 6.2: PARTICIPANT DEMOGRAPHICS. 106
TABLE 6.3: COMPLETE DATASET OF ΔC MEASUREMENTS FOR ALL TREATMENT GROUPS WITH $X_1 = E$ AND $X_2 = I \in \{-1, +1\}$. R_K IS THE K^{TH} REPLICATE OF A TREATMENT. EXPRESSED IN UNITS OF COMPLETE IDEAS. 116
TABLE 6.4: SAMPLE DATASET OF EIGHT ΔPS MEASUREMENTS FOR EACH TREATMENT GROUP WITH $X_1 = E$ AND $X_2 = I \in \{-1, +1\}$. R_K IS THE K^{TH} PARTICIPANT FOR A TREATMENT. EXPRESSED IN PS POINTS...117
TABLE 6.5: CRONBACH'S α MEASURED FOR CONSTRUCTS ΔPS, ΔRS, AND ΔQA IN SESSIONS 1 AND 2. 117
TABLE 6.6: COMPLETE DATASET OF ΔUI MEASUREMENTS FOR ALL TREATMENT GROUPS WITH $X_1 = E$ AND $X_2 = I \in \{-1, +1\}$. R_K IS THE K^{TH} REPLICATE OF A TREATMENT. EXPRESSED IN UNCERTAINTY INFLUENCE POINTS. 118
TABLE 7.1: MEAN VALUES FOR ΔC FOR ALL FOUR TREATMENTS, INCLUDING MARGINAL MEANS FOR EACH FACTOR, AND THE TOTAL MEAN ON THE BOTTOM-RIGHT CORNER. EXPRESSED IN UNITS OF COMPLETE IDEAS. 121
TABLE 7.2: MEAN VALUES FOR ΔG FOR ALL FOUR TREATMENTS, INCLUDING MARGINAL MEANS FOR EACH FACTOR, AND THE TOTAL MEAN ON THE BOTTOM-RIGHT CORNER. EXPRESSED IN UNITS OF GOOD IDEAS. 124
TABLE 7.3: MEAN VALUES FOR $\Delta ENPV$ FOR ALL FOUR TREATMENTS, INCLUDING MARGINAL MEANS FOR EACH FACTOR, AND THE TOTAL MEAN ON THE BOTTOM-RIGHT CORNER. EXPRESSED IN MILLIONS. 125
TABLE 7.4: MEAN VALUES FOR ΔPS FOR ALL FOUR TREATMENTS, INCLUDING MARGINAL MEANS FOR EACH FACTOR, AND THE TOTAL MEAN ON THE BOTTOM-RIGHT CORNER. EXPRESSED IN POINTS OF SATISFACTION. 126
TABLE 7.5: MEAN VALUES FOR ΔRS FOR ALL FOUR TREATMENTS, INCLUDING MARGINAL MEANS FOR EACH FACTOR, AND THE TOTAL MEAN ON THE BOTTOM-RIGHT CORNER. EXPRESSED IN POINTS OF SATISFACTION. 127
TABLE 7.6: MEAN VALUES FOR ΔQA FOR ALL FOUR TREATMENTS, INCLUDING MARGINAL MEANS FOR EACH FACTOR, AND THE TOTAL MEAN ON THE BOTTOM-RIGHT CORNER. EXPRESSED IN QUALITY POINTS. 129
TABLE 7.7: MEAN VALUES FOR ΔUI FOR ALL FOUR TREATMENTS, INCLUDING MARGINAL MEANS FOR EACH FACTOR, AND THE TOTAL MEAN ON THE BOTTOM-RIGHT CORNER. EXPRESSED IN INFLUENCE POINTS. 130
TABLE 7.8: MEAN VALUES FOR ΔFI FOR ALL FOUR TREATMENTS, INCLUDING MARGINAL MEANS FOR EACH FACTOR, AND THE TOTAL MEAN ON THE BOTTOM-RIGHT CORNER. EXPRESSED IN INFLUENCE POINTS. 131

Chapter 1 – Introduction

"Fear comes from uncertainty. When we are absolutely certain, we are almost impervious to fear." – William Congreve (1670 – 1729)

"It is not the strongest of the species that survive, nor the most intelligent, but the one most responsive to change." – Charles R. Darwin (1809 – 1882)

This thesis is concerned with the experimental evaluation of a design procedure supporting early conceptual activities for engineering systems design operating in uncertain environments. The thesis presents an experimental methodology to evaluate a design procedure objectively and quantitatively based on anticipated lifecycle performance of design concepts, together with a design procedure stimulating generation of flexible design concepts in face of uncertainty. The methodology presented here hopes to complement existing evaluation methodologies by measuring anticipated performance of design concepts via computationally efficient computer-modeling techniques. The design procedure for uncertainty and flexibility, in contrast to others used in cognitive science, collaboration engineering, and engineering design research, stimulates flexible design concept generation by packaging a short lecture on flexibility in design, and a prompting ideation mechanism.

A design procedure is referred as "a technique or method supporting the design process and/or artifact production". Well-known examples of design procedures are Axiomatic Design (N. P. Suh, 1990), Pahl and Beitz (1984), Quality Function Deployment (QFD) (Mizuno & Akao, 1993), Robust Design (Taguchi, 1987), Total Design (Pugh, 1991), and TRIZ (Altshuller, 1973). Such procedures incorporate creativity techniques to generate design concepts, structured mechanisms to explore the design space once design alternatives are generated, and analytical tools to represent the system, and manage the collaborative design process. Anticipated performance of a design concept is defined as the "anticipated capabilities of a design, as observed under particular conditions" (The New Oxford American Dictionary, 2006). Anticipated performance can be measured using financial metrics like Net Present Value (NPV), or non-financial metrics like patient service rate at a hospital, or Mission Technology Readiness Level at NASA (Avnet, 2009).

The design of engineering systems poses a new challenge to today's design activities. One reason is that engineering systems – such as airports, bridges, communication and transportation networks, healthcare facilities, power plants and electricity distribution grids, real estate development projects – are typically long-lived, require significant upfront investments (and therefore design efforts), and will operate under significantly uncertain economic, environmental, political, and technological influences. Customer demands, preferences, needs, as well as regulatory and environmental conditions will inevitably change over such long time periods (Eckert, de Weck et al., 2009). Design requirements will change as well during, and even after the design phase. Such changes have to be accommodated during the design cycle, and after the engineering system is launched in operations.

This thesis aims to address the new challenges of modern engineering design by encouraging designers to consider more explicitly uncertainty and flexibility in the early conceptual phases. The goal is to push the boundaries of engineering systems performance one step further. To do this, a novel and complementary concept generation technique is introduced help them do so. Similarly, a novel experimental methodology is proposed to evaluate rigorously and thoroughly the proposed design procedure for flexibility. The following Section 1.1 explains the specific motivations underlying this work. Section 1.2 describes the intended audience and limits of applicability of this thesis, while Section 1.3 summarizes the research approach. Section 1.4 provides a summary of the overall structure of the thesis.

1.1 Motivations

1.1.1 Current Considerations of Uncertainty and Flexibility in Design

Uncertainty is defined here as "anything affecting the future performance of an engineering system". The case of the Iridium cell-phone system is one example demonstrating how market uncertainty can severely affect the performance of an engineering system. In the 1990s, Motorola® supported the development of a satellite infrastructure to provide wireless cell-phone communications at any geographical point on the planet. This endeavor was completed in 1999.

As explained by de Weck, de Neufville et al. (2004), designers and managers underestimated demand for land-based cell phones, and overestimated demand for the Iridium technology. This resulted in the ambitious launch of sixty-six satellites, at a development cost of $4 billion. Even though the system functioned perfectly from a technological standpoint – it won several recognition awards – anticipated demand for Iridium phones never materialized. Because the system was optimized for some fixed capacity forecast, it never generated the revenues necessary to cover development costs. The system sold in bankruptcy for half of one percent of the original investment (Hesseldahl, 2001). This illustrates how uncertainty can impact design activities, and ultimately performance of engineering systems.

It is often the case in engineering design – as for the Iridium example – that considerations of uncertainty are simplified by considering one (or too few) future scenario for operational environment, market, regulatory conditions, and technology. This is understandable: it is already difficult to design a large-scale system considering all possible combinations of design variables, parameters, objective and utility functions, as well as operational conditions, it is even more difficult to design for a range of possible outcomes! As the first introductory quote suggests, designers may feel more comfortable with a simplified approach to uncertainty where design requirements are “frozen” early on (Eckert et al., 2009). This approach enables selecting a pinpoint, more certain, and clearly defined design through optimization techniques.

Even though engineering design has been extremely successful in the past, designs selected under this approach may become sub-optimal very quickly after launch. This is because the system is optimized for one (or too few) manifestation of futures conditions, which can easily change. For instance, platform designs at BP are often optimized assuming a fixed price of oil, and most likely quantity of original oil in place (Lin, 2009). Mining operating plans at Codelco – the national Chilean company – are optimized assuming an average copper price over many years (Cardin, de Neufville et al., 2008). The supersonic bomber Convair B-58 Hustler was optimized for high-altitude bombing in the 1960s, even though intercontinental ballistic missiles were developed at the same time – thus making it obsolete (Saleh & Hastings, 2000).

Even though engineers recognize uncertainty and how it affects performance, they must resist the tendency to simplify such considerations so that more agile and flexible engineering systems can be developed. Uncertainty, whether it comes from exogenous or endogenous sources, must be recognized early in the process, and dealt with pro-actively. Computer technologies developed over the last decades can support such thinking in the design process, making it more accessible than ever.

Design for flexibility enables dealing pro-actively with uncertainty, to improve expected performance over a range of uncertainty scenarios. Flexibility is defined here as "the ability, but not the obligation, to change the system configuration at a later time in light of some uncertainty realization". The intended effect of designing for flexibility is to enable practitioners to craft strategies positioning managers to adapt the system in face of changing circumstances, both in the design process and operations. Expected value improvements ranging between ten and thirty percent compared to initial assessments are routinely shown in the aerospace, airport, automotive, defense, energy, healthcare, public infrastructure, management, mining, real estate, and transportation industries[1].

Design for flexibility is however not a widespread approach. Among other reasons, this is because uncertainty is not easy to recognize and deal with in design, as there are many sources to consider potentially affecting performance. There is no "one fits all" solution: flexible alternatives are different from one system to another. Cultural issues are such that engineering typically operates in "silos", so that very little crossover occurs between engineering, marketing, and other disciplines.

This thesis is motivated by the need to recognize uncertainty and flexibility more explicitly in conceptual collaborative design of engineering systems. The hope is to devise a simple and user-friendly procedure that can complement existing ones, with the goal of bringing significant improvements to systems in terms of anticipated performance.

---

[1] See example case studies at http://ardent.mit.edu/real_options/Common_course_materials/papers.html and http://strategic.mit.edu/publications.php.

1.1.2 Current Evaluation Methodologies for Design Procedures

This thesis is also motivated by the need to measure both objectively and quantitatively the effects on anticipated performance of flexible design concepts. It was discovered during the research process that a quantitative performance-based methodology could complement very well existing methodologies for design concept generation procedures.

Typical evaluation methodologies for concept generation procedures often rely on subjective expert assessments of design concept quantity, quality, novelty, and/or variety (Shah, Kulkarni et al., 2000; Shah, Vargas-Hernandez et al., 2002) to study the effects of design procedures. Evaluations are complemented by measurements of participants' and users' impressions of satisfaction with the process and results, to ensure that the design procedure can be used satisfactorily in industry practice.

Evaluating concept generation procedures based on subjective evaluation metrics like idea quantity, quality, novelty, and variety has limitations. It may lead to inconsistent results, depending on the specific metrics used (Reinig, Briggs et al., 2007). This raises questions regarding the overall internal validity of this methodological approach. Also, it is difficult for any expert to assimilate all the complexities of design activities into a holistic quality scoring – using for instance a discrete Likert scoring mechanism between 1 and 10. Designers have to consider many design variables, parameters, objective functions, and operating scenarios. These cannot possibly be accounted for completely and satisfactorily in judging design concepts from a subjective standpoint. One needs finer resolutions to discriminate anticipated performance of design concepts than provided by a typical quality scoring mechanism. Similarly, it is not because a design procedure makes participants satisfied with the process and results, or because it is deemed of good quality, that the concepts will necessarily lead to good performance.

The specific issue of measuring anticipated performance is not well addressed in the literature on cognitive science, collaboration engineering, and research in engineering design. This need is however recognized, and the field is evolving in that direction (Kurtoglu & Campbell, 2009; Zuo, Leonard et al., 2010). This thesis aims to address this issue more specifically by introducing an objective and quantitative approach for design procedure evaluation based on anticipated performance of design concepts.

Another aspect of existing evaluation methodologies relates to the type of study that are performed. Even though they provide detailed and in depth knowledge, case studies evaluating design procedures in industry can take a long time, be difficult to reproduce rigorously, and hard to run several times over a short time period. 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 same difficulties arise in protocol studies where a designer is invited to think out loud during design activities, so that underlying cognitive processes can be studied and understood better.

The thesis wishes to augment current evaluation methodologies by proposing a controlled user experimental methodology for more thorough and rigorous performance-based evaluation of a proposed procedure for flexible design concept generation. The methodology hopes to enable quick replications, and relies on efficient computer-aided techniques for data collection and analysis, computer modeling, and evaluation of design concepts generated. The demonstration shown here suggests that the methodology can be reproduced for different systems, in different environments, and to evaluate different design procedures of interest thoroughly and rigorously – although this aspect needs to be validated further.

1.2 Intended Audience and Application

This thesis targets two audiences: industry practitioners and researchers interested in engineering design. One the one hand, the design procedure for flexibility targets designers who may be increasingly collaborating with experts in different areas (e.g. conceptual and detailed design, system architecture, marketing, manufacturing process and operations, end of life considerations, etc.). These realities require more agile thinking about design, calling for more explicit considerations of uncertainty and flexibility early in the collaborative design cycle. In particular, the design procedure is crafted for engineering systems design, with particular focus on critical infrastructures for communication (e.g. land and satellite-based telecommunication), defense (e.g. unmanned aero vehicle deployment), energy production (e.g. power plants and grids), healthcare (e.g. hospitals), transportation (e.g. airports, car manufacturing, highways, public transit), real estate (e.g. commercial and residential), and/or resource extraction (e.g. mining activities, oil production).

The procedure does not aim at identifying the optimal design configuration, in contrast to what is often done through use of traditional optimization techniques. Under the design for flexibility paradigm, a design may very well optimize mean performance (e.g. mean NPV), but can still become sub-optimal if the worst scenario occurs (e.g. demand and prices for the product plummet, like for the Iridium system). This highlights the fact that a flexible design better suited for a range of possible scenarios cannot be optimal for all scenarios that actually occur in reality. These ideas differ subtly from the traditional engineering paradigm, which relies heavily on "finding the optimal design".

The procedure differs from typical scenario planning often employed in design activities. It focuses on how to enable early on the best course of actions through design, depending on the different scenarios that may arise. It promotes a pro-active rather than reactive approach to design. The procedure aims to extend traditional design thinking by nudging practitioners to consider a range of possible outcomes as the basis for design activities – as opposed to a narrow set of scenarios. What the procedure guarantees is to provide a structured mechanism relying on designers' expertise to anticipate possible futures, craft strategies to deal with them more effectively, and communicate better about the system. The net desired effect is to improve expected performance in light of a range of scenarios – and not identifying a single optimal design. Even in the case where the worst possible outcome occurs, a flexible system may be in a better position than an inflexible one to react and assuage the downside effects from bad economic or environmental downturns. Similarly, if the future is better than expected, flexibility enables capitalizing on such upside opportunities.

The design procedure intends to complement existing design procedures, prior to a more detailed design analysis. The early phase is the point where most influential decisions can be made to alter future system cost, performance, and schedule. At the detailed design phase, many requirements are already locked in, and much freedom is lost. The cost of enabling and exercising flexibility later on may be much higher (Silver & de Weck, 2007). Therefore, the design procedure is not expected to be useful at later stages of design analysis and system architecting. Given it relies on collaboration between designers in different areas of expertise, the design procedure is not expected to make much difference at organizations structured around the "silo" model, where different areas of the product development cycle have little interactions with one another. This was the case at BP, as outlined by Lin (2009), where platform engineers had little to no interactions with reservoir engineers due to a sequential design approach. Designers were imposed a fixed price per oil barrel by the management team, making the process of designing for uncertainty and flexibility tedious – if not impossible. This reality is unfortunately widely spread in practice, especially those specializing in infrastructure and civil engineering.

For researchers, the experimental methodology hopes to extend the set of metrics available when evaluating and validating design procedures experimentally. It is best suited for research interested in early generation of design concepts. It may not be appropriate for other areas of engineering design research. Even though this thesis demonstrates application for a simplified design problem, the methodology is ostensibly applicable in a real-world setting, in collaboration with an engineering firm working on a realistic system. The methodology may be useful to help the firm decide between different design procedures, by providing a platform for quick and efficient evaluation, prior to more widespread dissemination at the firm. Verifying this assertion is out of the scope of this thesis, but is certainly an interesting opportunity for future research.

The design procedure for flexibility is also interesting for researchers, as it aims to extract and present general principles that can be used more widely in design, and applied to different systems. This may stimulate research interested in systematically incorporating flexibility and more agile design thinking into real-world design practice.

1.3 Research Approach

A novel design procedure is suggested for flexible design concept generation early in the design cycle. The procedure aims to improve overall system performance, and provide users with subjective impressions of satisfaction with the process and results, and of better results quality.

An experimental methodology is developed to evaluate whether the proposed procedure adequately incorporates these attributes. Existing evaluation methodologies for design procedure can be used to study the effects on creativity levels, user satisfaction with the process and results, and anticipated quality of results assessed subjectively. As explained above and detailed in Section 2.4, these methodologies cannot be used satisfactorily to assess anticipated performance improvements objectively and quantitatively. Existing methods typically rely on subjective assessments by experts and participants, which is not sufficient in this context.

Thus, it is necessary to develop an experimental platform and methodology for objective and quantitative performance-based evaluation of early design concepts. This evaluation methodology complements existing ones focusing on subjective assessments. It extends the set of metrics enabling rigorous and thorough evaluation, and can ostensibly be used and reproduced quickly and efficiently.

1.4 Thesis Content Summary

The thesis is structured as follows. Chapter 2 surveys the literature to position the thesis as part of engineering design research, and to determine the contributions from existing design procedures in early conceptual design activities. The chapter also assesses the current state of the art in terms of design procedure evaluation methodologies, and what is needed to evaluate the proposed procedure for flexibility. Chapter 3 presents the research questions and overall research approaches. Chapter 4 presents the procedure for flexible design concept generation. Chapter 5 describes the experimental methodology used for evaluation from a generic standpoint, while Chapter 6 describes the specific experimental implementation to evaluate the procedure for flexibility. Chapter 7 presents evaluation results, which are discussed in Chapter 8. Chapter 9 concludes and opens on upcoming research opportunities in related areas.

Chapter 2 – Literature Review

"Knowledge is the food of the soul." – Plato (c. 428 – 348 BC)

This chapter has five objectives. The first objective is to situate this work as part of the broader field of engineering design research. An overview of the current state of the field is provided in Section 2.1, with indications where the thesis best fits best within this intellectual framework. The second objective is to identify potential contribution to existing design procedures by considering uncertainty and flexibility more explicitly in the design process. To this end an overview of existing design procedure is provided in Section 2.2, describing the purpose, uses, and origins of established design procedures, including concept generation and design for flexibility. The third objective is to determine how existing design procedure evaluation methodologies can be improved to evaluate a procedure for flexibility. Section 2.3 provides an overview of existing methodologies from the literature on cognitive science, collaboration engineering, and research in engineering design. The fourth objective emerges from these literature reviews, and consists of identifying research gaps and opportunities, as explained in Section 2.4. The fifth objective fulfilled in Section 2.5 is to explain how this thesis intends to address these issues, and contribute to the wider body of engineering design knowledge.

2.1 Engineering Design Research

In order to situate these intellectual contributions, the section summarizes a recent organization of engineering design research by Horváth (2004). This summary shows the extent to which engineering design research is a wide field capturing ideas and methodologies from engineering, management, and social sciences. The rationale is provided as to why the thesis best fits within the research categories of design theory and methodology – even though it could fit partially within many areas.

2.1.1 Proposed Order in Engineering Design Research

Research in engineering design developed rather chaotically since the first documents by Reuleaux (1861, 1875) formalizing and structuring design activities (Horváth, 2004). Many countries and companies pioneered their own design traditions over decades, without necessarily communicating with each another, most likely due to cultural and language barriers. For instance, Taguchi (1987) formulated the basis for robust design in the 1950s in Japan. Japanese manufacturers developed QFD around the 1960s, while TRIZ was being created in Russia. Pahl and Beitz developed their approach – one of the most widely known approaches – in Germany around the 1970s.

Finger and Dixon (1989a, 1989b) were among the first attempts at organizing the field of engineering design research. In Part I, they categorized design activities as descriptive, prescriptive, and computer-based models of design processes. In Part II, they described languages, representations, and environments for design, analytical tools to support design decisions, as well as design for manufacturing and other life cycle issues such as reliability, serviceability, etc. Their description however focused mostly on the United States. Also, it did not focus as much on the methodologies employed at large in engineering design studies, but rather on the procedural activities of design.

Horváth (2004) provided a more extensive organization of the field, considering as well the different research methodologies employed to study engineering design in a broader context. The author also extended the overview far beyond the methods used in the United States. According to the review summarized in Figure 2.1, engineering design research is organized into three categories: source, channel, and sink research.

Source research studies the fundamental human capacities required for engineering design activities, along four dimensions: human assets, design knowledge, artifacts knowledge, and processes knowledge. Human assets – the repository of mental and physical capacities to produce new values – studies how designers go about design. Such studies involve design psychology (the mind and behaviors of designers), design cognition (thought and physiological processes occurring during design activities), design ethnography (describing designers culture and culture-sensitive artifacts), design aesthetics (appearance and perceptions of shape, functions, attributes, and behaviors of products), design ergonomics (interactions between humans and products/environments), and product marketing (marketing of design artifacts and related technical services).

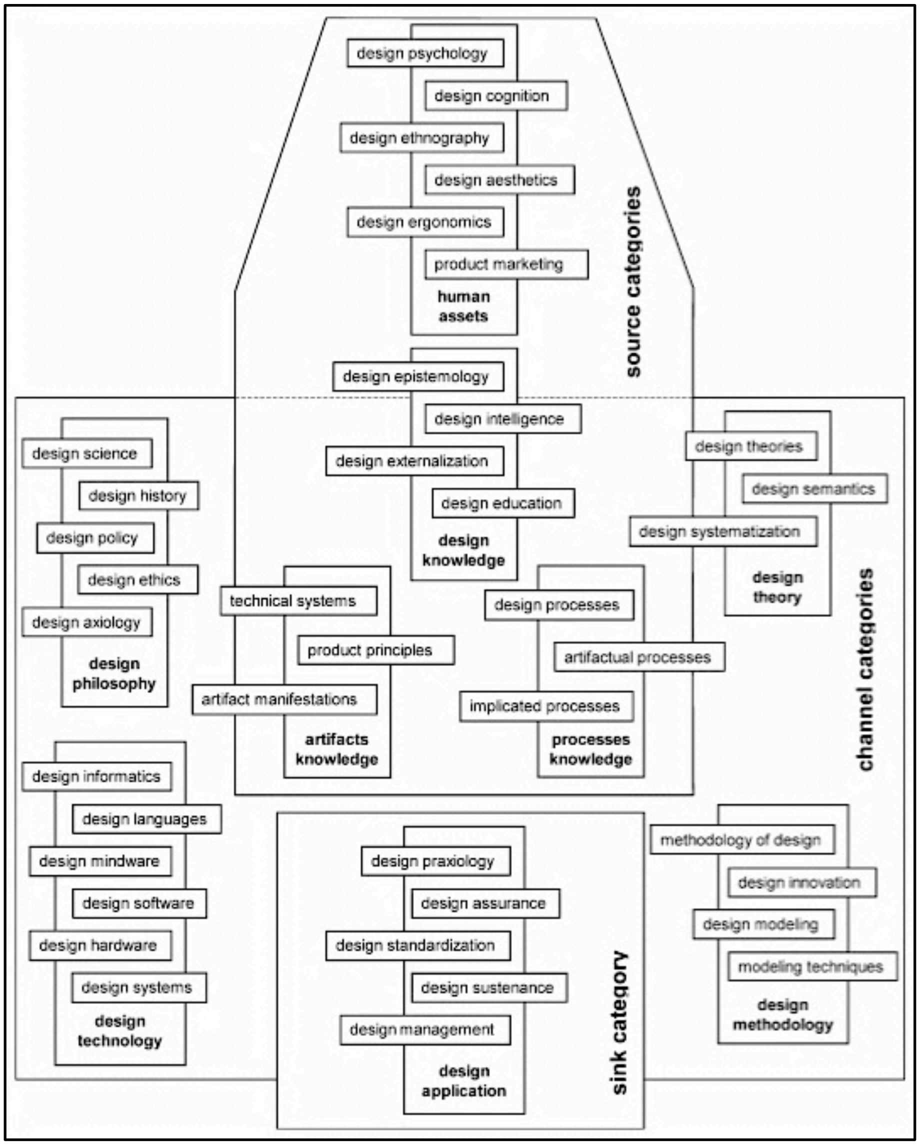


Figure 2.1: Organization of engineering design research according to Horváth (2004).

Studies of design knowledge – knowledge about design and of design – involve design epistemology (knowledge of the origins, nature, forms, structure of design, as well as validation and methods), design intelligence (ability to think, learn, and cope creatively with the unexpected), design externalization (generation of mental images, representation, and communication of design ideas), and design education (how to teach and perform design activities). Research in artifacts knowledge – understand the artifacts generated historically over time – studies technical systems (mechatronic devices, mechanisms, mechanical engineering machines, thermal engineering equipment, etc.), product principles (product paradigms, structures, technologies, materials, and product intelligence), and artifact manifestations (producing artifact taxonomies, catalogues, properties, and evaluations). Studies in processes knowledge focus on design processes (modeling and resources to improve quality), artifactual processes (existential, operations, application, and service processes of products), and implicated processes (realization and exploitation of products).

Channel research is concerned with making connections between scientific/theoretical knowledge of design, and pragmatic/technical design knowledge. This is done through studies involving design philosophy, design theory, design methodology, and design technology. Research focusing on design philosophy – existence and manifestation of design, role and position in society, historical evolution, foundational basis for design thinking – looks into design science (ontology, phenomenology, teleology of design), design history (chronological development of design knowledge and sub-disciplines), design policy (execution of complex research projects), design ethics (ethical dimensions in engineering design, such as man made changes to nature, moral rules of designing, etc.), and design axiology (measures of value of products, reflection about value, and passion of design). Research in design theory – organization of design knowledge to serve practical purposes – is concerned with design theories (general, specific, descriptive, and prescriptive theories of design processes, as described by Finger and Dixon, 1989a, 1989b), design semantics (meanings and intentions in design), and design systematization (techniques supporting design automation, decision-making, instrumentation, and optimization). Design methodology research – design methods, activities, and techniques providing guidelines for design – involve methodologies of design (methodological systematization of design processes and employment of modeling, representation, analysis, simulation, evaluation, and/or physical testing techniques for researchers), design innovation (rationalizing multi-disciplinary product development and creative concept generation), design modeling (how to model artifacts and processes), and modeling techniques (mathematical, symbolic, textual, verbal, and visio-spatial approaches for representing design artifacts, knowledge, and processes). Research in design technology – how engineering design is used to produce technology – focuses on design informatics (handling of design data, information, and human knowledge of design), design languages (procedural and visio-spatial grammars to structure design functions and constraints), design mindware (structuring and archiving design data in digital repositories), design software (modeling, analysis, simulation software to support design activities), design hardware (facilitating design support on the computational side), and design system (integration of various design tools into holistic approaches).

The only sink research category – research in design application – focuses on generating knowledge necessary to deploy engineering design knowledge through concrete applications. It comprises research in design praxeology (efficient design action and problem solving), design assurance (quality of design actions and deliverables), design standardization (generating codes and norms for efficient and quality design technology production), design sustenance (how to sustain and manage design projects), and design management (low-level organization to support design activities, exploitation of particular design tools for particular products).

2.1.2 Thesis Positioning Within Engineering Design Research

This thesis fits best within the two channel research categories of design theory and methodology – see right side of Figure 2.1. The rationale is that the thesis presents a prescriptive approach to support the specific issue of designing for flexibility, which fits best within design theory. The thesis also proposes a novel experimental methodology to evaluate objectively and quantitatively the design procedure for flexible design concept generation, which is in line with the definition of design methodology.

2.2 Current Design Procedures

This section provides an overview of current design procedures to narrow down where the proposed procedure for flexibility can contribute for engineering practice. This overview is based on the classifications prepared by Tomiyama et al. (2006; 2009) as well as Finger and Dixon (1989a, 1989b). The overview related to design concept generation procedures is inspired from the work by Shah et al. (2000; 2002), as well as Knoll and Horton (2010). The section also provides an overview and classification of existing design procedures focusing on the issue of flexibility. The research issues and opportunities for contributions identified out of this overview are summarized in Section 2.4.

On a semantic note, the definition of "design theory and methodology" by Tomiyama et al. (2006; 2009) is closer conceptually to how Horváth (2004) defines "design theory". This is because Horváth's definition focuses mostly on the generic, specific, prescriptive, and descriptive theories of design processes, which is what Tomiyama et al. (2006; 2009) describe and categorize. "Design methodology" according to Horváth (2004) incorporates the methodologies researchers may use to evaluate different design procedures, which is not part of the definition by Tomiyama et al. (2006; 2009).

To resolve this ambiguity, the term *design procedure* is used in this thesis to replace what Horváth (2004) refers to as design theory, and to what Tomiyama et al. (2006; 2009) calls design theory and methodology. It also encompasses conceptually what Finger and Dixon (1989a) call "prescriptive models for design". Examples of design procedures are Axiomatic Design (N. P. Suh, 1990), the Taguchi Method (1987), or TRIZ (Altshuller, 1973). On the other hand this thesis uses the term "methodology" when referring to the methodology for design procedure evaluation. This is closer conceptually to Horváth's definition of design methodology.

2.2.1 Established Design Procedures

Tomiyama et al. (2006; 2009) developed a classification of design procedures based on General Design Theory (GDT) (Tomiyama & Yoshikawa, 1987; Yoshikawa, 1981). GDT assumes that design knowledge can be formalized mathematically, and operated upon. It comprises a set of three axioms: axiom of recognition (design entities can be recognized and described by attributes and/or design concepts), axiom of correspondence (entity set and set of entity concepts have a one-to-one correspondence), and axiom of operations (the set of abstract concepts is a topology of the set of entity concepts that can be operated upon through logical operations like union and intersection). In GDT, a design procedure maps design function space to design attribute space.

According to these authors, design procedures are classified depending on their intended use:

1. Generate a new design solution;
2. Enrich functional and attributive information of design solutions;
3. Manage design and represent design knowledge.

The first category of design activities is referred as *concept generation* in this thesis for simplicity. The second category is called *design space exploration*, as it allows designers to investigate, evaluate, and select neighboring solutions to the design concepts elicited with the goal of improving some attribute and/or function. The third category is called *management and representation*. This enables capturing, representing, modeling, and codifying design knowledge and information about design processes, objects, environments, and life cycle issues. It also enables managing more easily the design process.

Table 2.1 provides an overview of established design procedures fitting along the three categories of design activities, inspired from the work by Tomiyama et al. (2006; 2009), and in chronological order from earliest to latest (insofar as information is available on development times). Each design procedure is summarized, and justifications are provided as to why a given procedure fits particular categories. For a more extensive review of design procedures, the reader is directed to the original review papers.

Robust Design

Gen'ichi Taguchi (1987) pioneered the basis of robust design in Japan in the 1950s. The Taguchi method now forms the basis of modern quality engineering, and is tightly linked to statistical theory through Design Of Experiment (DOE) (Wu & Hamada, 2000). The design procedure was introduced in North America and Europe around the 1980s. It explicitly recognizes exogenous uncertainties affecting value, and suggests dealing with them by designing the system to operate under a wide variety of conditions without modifying the design in operations.

Robust design introduces the notion of loss function, which consists of losses due to product function variations and other losses due to cost and side effects. An ideal robust design is one that is less sensitive to variations in manufacturing conditions, customer use, and natural degradation over time (Frey & Dym, 2006). The goal of robust design is often to minimize the volatility of a given system's response.

Table 2.1: Examples of well-known procedures for early conceptual design activities.

	Design Activities		
Design Procedures	Concept Generation	Management and Representation	Design Space Exploration
1950s			
Robust Design (Taguchi, 1987)			✓
1960s			
Design Decision-Making (Simon, 1960)			✓
Decision Matrix (Kepner & Tregoe, 1981)		✓	✓
Quality Function Deployment (Mizuno & Akao, 1993)			✓
TRIZ (Altshuller, 1973)	✓		
1970s			
Axiomatic Design (N. P. Suh, 1990)	✓	✓	✓
Pahl and Beitz (Pahl & Beitz, 1984)	✓		
Screening Models (Jacoby & Loucks, 1972)		✓	✓
Concurrent Design		✓	
1980s			
Design Structure Matrix (Steward, 1981)		✓	
Total Design (Pugh, 1991)		✓	✓
1990s			
Design for X			✓

Design Decision-Making

Herbert Simon (1960) developed the basis for Design Decision-Making methods. These methods typically rely on the generation of design alternatives, and provide mechanisms to evaluate and select design alternatives. This is why they fit best within the design space exploration category. These procedures require that designers know exactly what they want to achieve, and have a clearly defined metric for evaluating and comparing design alternatives. Design for X (DfX) procedures described below offer a similar paradigm, as designers must know what they want the design to accomplish (e.g. maintainability, low cost, reliability, etc.). Other approaches involve determining designers' utility function as the basis for decision-making. For instance, Hazelrigg (1998) proposed using decision analysis and utility theory to evaluate and select the most valuable design configuration – a framework known as Decision-Based Design (DBD). DBD also incorporates surveys of population choices to determine the best design. Keeney (2009) outlined the difficulties and limitations of relying on utility theory in the context of design.

Decision Matrix

The Decision Matrix developed by Kepner and Tregoe (1981) in the 1960s has columns representing the different design alternatives, and rows specifying desired attributes (e.g. ease of fabrication, speed, stiffness, etc.). In Figure 2.2, each attribute A_m is weighed according to designers' utility and preferences (W_m values). Each alternative receives a score (S_{mn}) relative to a particular attribute. Design alternatives are then evaluated based on total weighted scores (WS_n).

Attributes	Weight	Concept 1	Concept 2	Concept 3	…	Concept n
A_1	W_1	S_{11}	S_{12}	S_{13}	…	S_{1n}
A_2	W_2	S_{21}	S_{22}	S_{23}	…	S_{2n}
A_3	W_3	S_{31}	S_{32}	S_{33}	…	S_{3n}
…	…	…	…	…	…	…
A_m	W_m	S_{m1}	S_{m2}	S_{m3}	…	S_{mn}
Weighted Score		WS_1	WS_2	WS_3	…	WS_n

Figure 2.2: Example Kepner-Tregoe decision matrix. A_m is a particular desirable attribute, W_m is the weight of an attribute based on designers' preferences, S_{mn} is the score of a given concept relative to attribute A_n, and WS_n gives the weighted score of a design concept considering all attributes.

The decision matrix is most useful to explore the design space collectively, thus helping managing the design process as well. It is one of many tools integrated by Ullman (1995) for systematic mechanical design and experience. It has been used and applied extensively in product development, and at various companies like Boeing and Hewlett Packard (Tomiyama et al., 2009).

Quality Function Deployment

QFD is a design procedure developed in Japan in the 1960s by Mizuno and Akao (1993) to consider more closely customer demands throughout the product development life cycle. QFD was originally developed as a step-by-step process to control product quality throughout the manufacturing process. The four steps are:

1. Quality deployment, involving mapping customer demands (i.e. Voice of Customers) into measurable product quality, structures, and components;
2. Technology deployment, requiring mapping product structure and components into technology items and manufacturing processes;
3. Cost deployment, requiring enumeration of cost items according to technology deployment, and;
4. Reliability deployment, using Failure Mode and Effect Analysis (FMEA) (McDermott, Mikulak et al., 1996) to lower the probability of defects and to improve efficiency.

A matrix formulation is used in each step of the process to map quality items into other quality items – referred to as House of Quality (see example in Figure 2.3). QFD can also be used in conjunction with other design procedures, such as TRIZ and Taguchi.

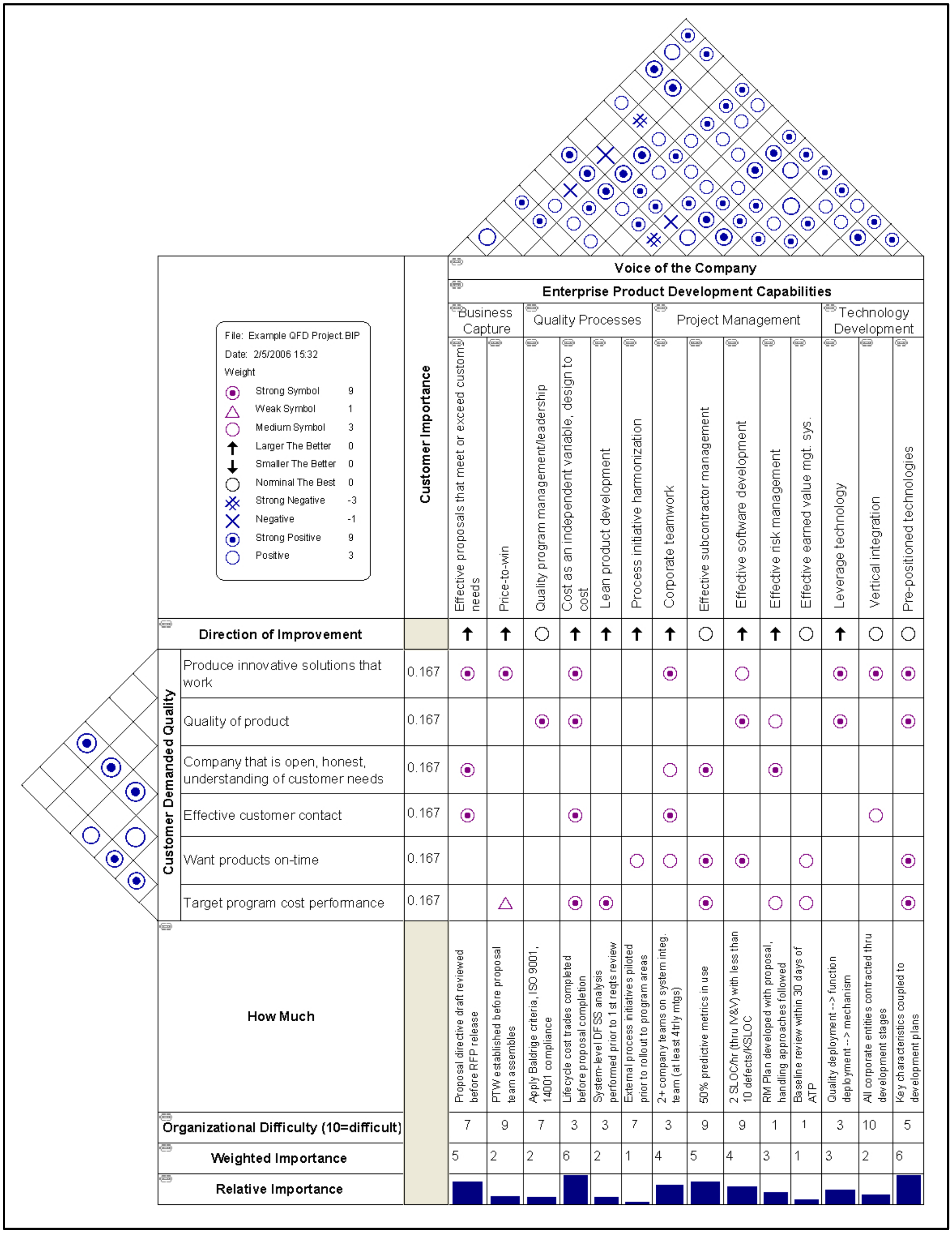


Figure 2.3: Example House of Quality process used to map customer demand quality to attributes to a company's product development capabilities.

TRIZ

The *Teoriya Resheniya Izobretatelskikh Zadatchis* in Russian, or Theory of Inventive Problem Resolution (TRIZ) is one of the most widely known approaches to design. Its development is attributed to Altshuller (1973) in the former Soviet Union. TRIZ is based on a set of fundamental design principles extracted inductively from intensive studies of technical object evolution and analysis of forty thousand innovative patents. Straker and Rawlinson (2003) provide forty examples of such design principles. For instance, the principle of segmentation promotes design modularity to ease the manufacturing process in terms of repair, transportation, and assembly. The principle of extraction suggests removing bad parts in a system, and placing them somewhere else to have a different function, and potentially different results. The principle of local quality encourages designers to consider that any part in a given design can be changed, with the goal of improving quality. TRIZ has been widely used in industry in Germany, Japan, Korea (by Samsung in particular), and of course Russia. It requires several months of intensive training, and hence students are typically exposed only partially to the whole procedure (Tomiyama et al., 2009).

Axiomatic Design

Axiomatic Design (N. P. Suh, 1990; 1998) is among the most cited procedure in engineering design literature. It relies on two axioms: axiom of maximum independence of functional elements, and axiom of minimum information content. The first axiom encourages design artifacts having a one-to-one correspondence between each functional requirement and design component. This ensures that designs are adjustable, controllable, and potentially avoiding unintended consequences. The second axiom encourages simple designs that are robust to exogenous operating conditions. Design concepts are generated from customer preferences, determining design requirements. The process summarized in Figure 2.4 flows from the customer domain where Customer Attributes (CA) (i.e. demands and preferences) are established. These attributes are mapped onto the functional domain by defining the Functional Requirements (FR) fulfilling these demands and attributes. FRs are mapped onto the physical domain by choosing appropriate Design Parameters (DP) fulfilling the functional requirements. DPs are mapped onto the process domain to determine the Process Variables (PV) to further implement the design. The design space exploration process consists of “zigzagging” between the different domains linking the customer, functional, physical, and process domains. Zigzagging between different domains also helps structuring and managing the design process. Axiomatic design has been used extensively in the product and manufacturing industries (Gu, Rao et al., 2001; N. P. Suh, 1990, 2001).

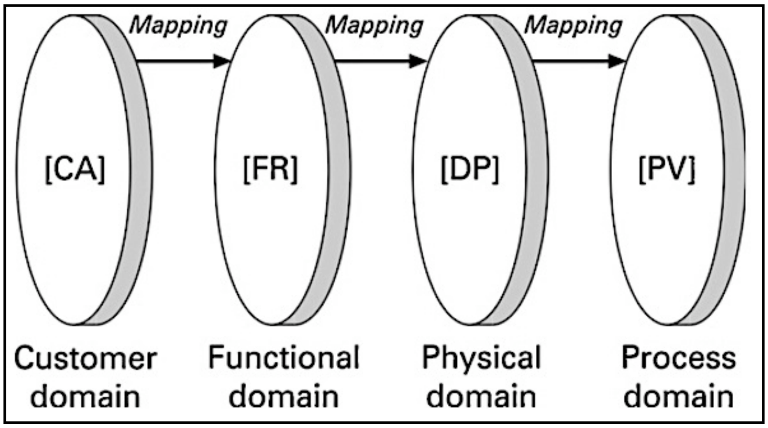


Figure 2.4: Process flow in Axiomatic Design (Tomiyama et al., 2009).

Pahl and Beitz

The design procedure by Pahl and Beitz (1984) is by far the most known and used in both industry and academia. It was developed in Germany in the 1970s through an elaborate analysis of technical systems fundamentals, systematic approaches, and problem solving processes. It is a step-by-step process guiding designers through the whole product life cycle, summarized in Figure 2.5. It includes tasks such as planning and clarification of the task, as determined from market demands and preferences, or other requirements. Conceptual design follows to determine the solution principle and functions required to accomplish the design goal. Embodiment and detailed designs then provide the overall layout and concrete arrangements of forms, dimensions, materials and properties of the product, with cost estimates. This method best represents the traditional approach to design taught at engineering institutions.

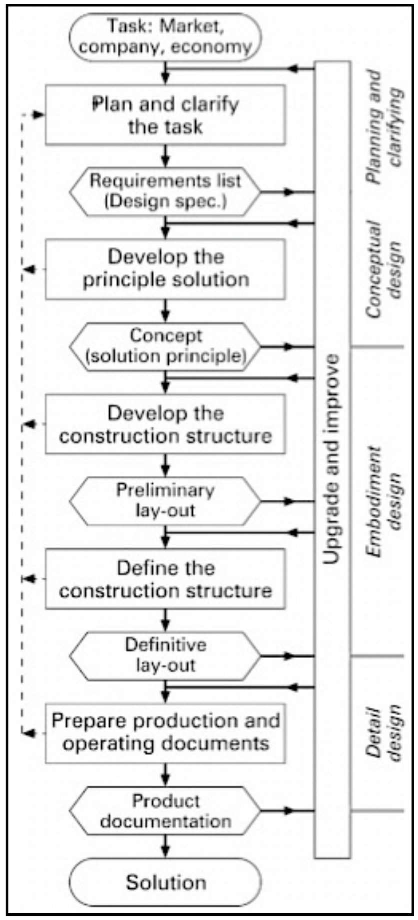


Figure 2.5: Process flow for Pahl and Beitz (Tomiyama et al., 2009).

Screening Models

Proposed first in the 1970s by Jacoby and Loucks (1972), screening models are used to explore the design space rapidly and efficiently to identify interesting areas of the design space for more detailed analysis. The typical process is shown graphically in Figure 2.6. Screening models may rely on optimization algorithms and DOE techniques to search – and screen – the design space for interesting candidate designs. Interesting design candidates are analyzed further with a more detailed model of the system. de Neufville and Scholtes (2011) identify three types of screening models: bottom-up, simulators, and top-down. Bottom-up screening models use simplified versions of a complex, detailed model used for design. They typically simplify and integrate each model component in a coherent whole. Simulators incorporate statistical techniques and/or fundamental principles to mimic the response of the complex model. Top-down screening models use representations of major relationships between the parts of the system to understand possible system responses. In addition to enabling quick and efficient exploration of the design space, screening models are useful to represent the design, as they require designers to produce a model of the system for quick and detailed analyses, before proceeding to more detailed analysis.

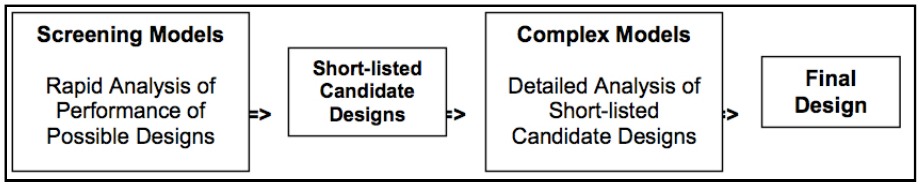


Figure 2.6: A screening model should be used to precede and complement a more detailed analysis of the design (de Neufville & Scholtes, 2011).

Concurrent Engineering

Concurrent Engineering is an integrated design procedure for managing product development for the entire life cycle, from planning, design, to production, delivery, and end-of-life considerations. It is shown to dramatically improve product development lead-time, and overall product quality (National Research Council, 1991). It was developed by the Japanese automotive industry and brought to U.S. engineers' attention in the 1980s (Womack, Jones et al., 1991). This procedure typically involves tight communication between product and process development teams, collective target sharing, identification of difficult issues, and integrated design and analysis activities. It differs in that respect from traditional engineering typically breaking down the design problem into subsystem components, with teams working on subsystems in isolation and with little to no contact with other teams (Stagney, 2003).

Design Structure Matrix

The Design Structure Matrix (DSM) was introduced by Steward (1981) as a way to represent design tasks as a sequence of network interactions. A DSM is a square matrix where the rows and columns list all the relevant design and management components of a system (Browning, 2001). The DSM can encode and represent graphically an engineering system, which is useful to manage the design process. The matrix entries represent how the design and management components are connected, and how the information flows from one another. Although used early on to represent design task activities (Eppinger, Whitney et al., 1994), the DSM has been used for the analysis of technical artifacts (Malmstrom & Malmquist, 1998; Pimmler & Eppinger, 1994).

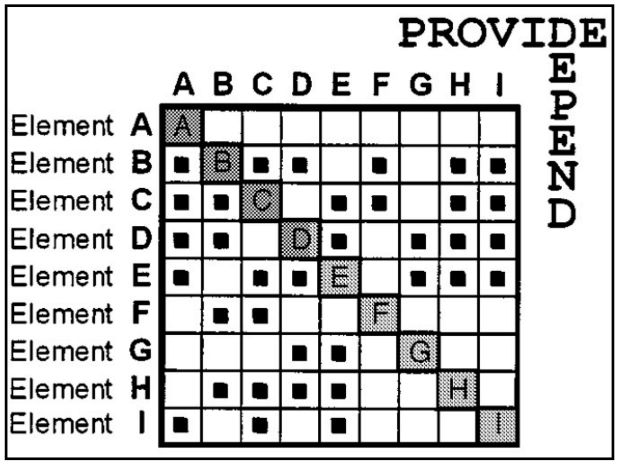


Figure 2.7: Example DSM (Browning, 2001). Elements in the row provide information to elements in the columns, while elements in the columns receive information (or depend) from elements in the rows.

Total Design

The Total Design approach developed in the 1980s by Pugh (1991) offers a set of tools to support design concept evaluation and selection by design teams, very early in the conceptual stages. Given the tools are easily applicable, Total Design is gaining more popularity in the product development industry. General Motors used the approach in development of the Saturn project (Tomiyama et al., 2009). One of Pugh's main contributions is the Pugh Controlled Convergence (PuCC) mechanism helping design teams converge iteratively towards a smaller set of design alternatives (Frey, Herder et al., 2009). This mechanism can be used to evaluate overall design solutions, but also for concept selection for system architecture, as well as subsystem and individual components. The approach is similar to the Kepner-Tregoe Decision Matrix, in that design concepts are listed in the columns, and desired attributes are listed in the rows. The main difference lies in using a datum, or benchmark design alternative against which all other design alternatives are compared. The approach uses +, S, or – signs when a design concept is better, similar, or worst than the datum for a given attribute. Summing the pluses and minuses gives an idea of least performing design concepts, which eliminates design concepts iteratively, and enabling convergence towards a smaller set of alternatives. One of the benefits of the PuCC mechanism is to complement any design procedure after design concepts are generated. It provides a structured mechanism for design teams to explore the design space collectively and efficiently, based on subjective performance assessments.

Design for X

Design for X represents a family of design procedures, all aiming at optimizing a particular attribute of the design process. DfX procedures are typically used in the product manufacturing industry. Tichem (1997) explains that X can represent a specific property to improve (e.g. cost, efficiency, lead time, quality), or a particular life-cycle phase of the product (e.g. manufacturing, assembly, distribution). The Six Sigma initiative can be seen as design for manufacturing reliability, as it aims at minimizing the probability of product defects (iSixSigma, 2010). Design for flexibility can be a desirable attribute of the manufacturing process, and hence could be considered as part of the DfX family. Design for flexibility in this context usually relates to making products more flexible to suit different customer needs (Rajan, Wie et al., 2005). It also refers to strategies employed to structure the manufacturing process more flexibly (Sethi & Sethi, 1990).

The procedure for flexibility investigated in this thesis intends to have wider applicability than for product manufacturing. It aims to apply to complex engineering systems like product manufacturing systems, but also including other large-scale systems like critical infrastructures. The goal is not necessarily to be more flexible, but to improve performance by means of flexibility. As it is referred to in this thesis – and as detailed in Section 2.2.3 – design for flexibility involves recognizing a distribution of possible outcomes upfront in the design cycle, and creating designs adapting pro-actively to changing circumstances. This is a significant conceptual departure from traditional engineering paradigms focusing on optimizing to deterministic forecasts. Also, there is whole set of design tools necessary to design for flexibility, ranging from concept generation, to management and representation, and design space exploration. This contrasts with the DfX family, which aims mainly at improving some attributes and/or functions of the design.

For all these reasons, design for flexibility is not considered a mere component of the DfX family. As explained below, designing for flexibility represents rather a complementary approach to established design procedures, standing on its own.

2.2.2 Design Procedures for Concept Generation

This thesis is interested in a procedure for flexible design concept generation. In essence, this work is focusing on the first category of design procedures dedicated to concept generation presented in Table 2.1. This thesis aims to extend these procedures to focus on the specific issues of uncertainty and flexibility. This subsection presents a more refined overview of such procedures, based on the work by Shah et al. (2000; 2002), and summarized in Figure 2.8.

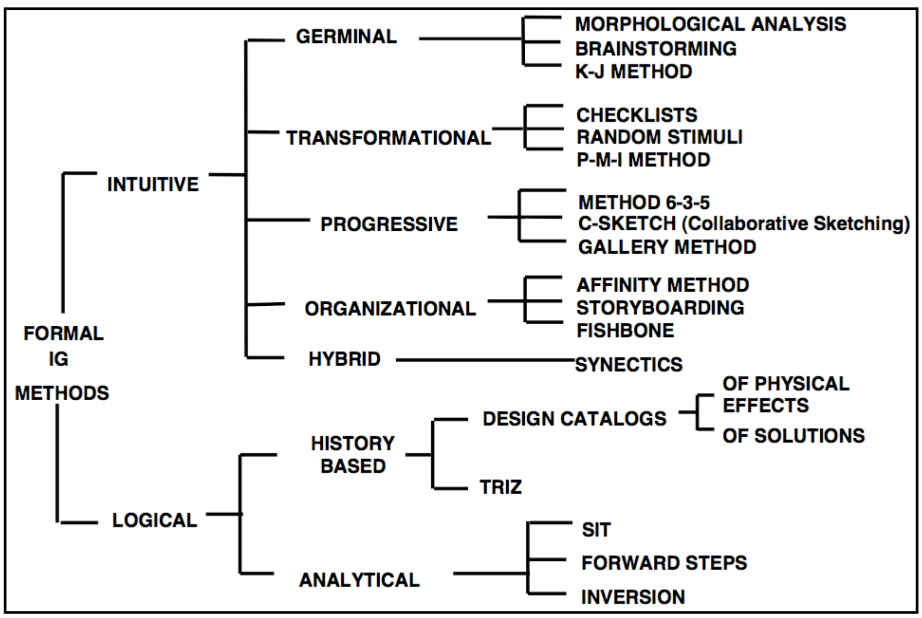


Figure 2.8: Overview of formal IG methods (Shah et al., 2000; 2002).

Idea Generation (IG) techniques are classified generically as either intuitive, or logical. Intuitive techniques include germinal (i.e. creating new design concepts from scratch), transformational (i.e. evolving an idea from an initial concept), progressive (i.e. repeating the same steps many times to generate new ideas), organizational (i.e. getting people together to create ideas), and hybrid methods (i.e. mixing one or another approach). For instance, brainstorming (Osborn, 1957) is one example of germinal technique, while relying on itemized checklists is an example of transformational approach. The 6-3-5 method is an example of progressive approach, as it involves a team of six designers generating up to three design ideas each through five cycles. Storyboarding is an organizational approach using, for instance, diagrams showing upcoming events in sequence as a way to generate new design concepts. Logical methods can be classified into history-based (i.e. relying on past solutions and principles catalogued and archived) and analytical (i.e. systematically analyzing basic relations causal chains, and desirable/undesirable attributes) methods. TRIZ is an example of history-based approach.

To complement this review, Knoll and Horton (2010) reviewed one hundred and one techniques to stimulate creative ideation based on changing participants' perception of the problem. They asserted that creativity techniques typically rely on three general different cognitive principles: Analogy, Provocation, and Random changes of perspective. Analogy searches for similar situations and uses existing knowledge about these situations to generate new ideas. Provocation challenges the underlying assumptions of the creative task to generate a new perspective. Random relies on external stimuli that are unrelated to the creative task.

2.2.3 Design Procedures for Flexibility

Designing for uncertainty and flexibility can be particularly valuable to complement established design procedures described above. Systems designed today grow in complexity, and 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. As explained below in Section 2.4.1, many design procedures often assume that design requirements are known *a priori* from customer demands and preferences, and systems are optimized to satisfy a set of deterministic forecasts. While these procedures to design have been extremely successful, more agile and iterative design approaches are gaining terrain in various engineering disciplines. In order to consider the growing complexities of todays engineering design, one needs complementary design procedures considering uncertainty and flexibility more explicitly over the entire lifecycle of the project.

The need to design engineering systems with wider considerations of economic, environmental, political, and technological uncertainty and flexibility is being increasingly recognized (de

Neufville & Scholtes, 2011; Eckert et al., 2009; Saleh, Mark et al., 2008). Studies are emerging to understand basic underlying principles to design for flexibility. There has been much work in the manufacturing industry and product development sector (de Toni & Tonchia, 1998; Ferguson, Siddiqi et al., 2007; Sethi & Sethi, 1990) to extract general principles and lessons to enable flexible design thinking – not as much for engineering systems in general (Saleh et al., 2008). Designing engineering systems for uncertainty and flexibility is still not a widespread approach.

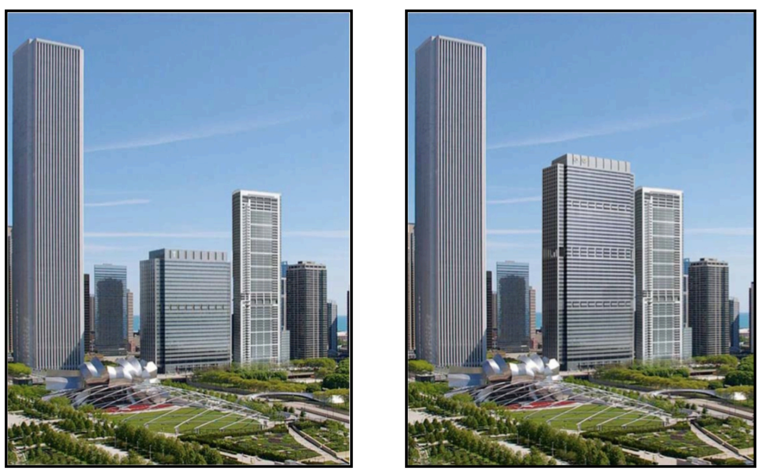
The aim of designing for uncertainty and flexibility is to help designers make better decisions early on to provide contingencies that will improve overall anticipated performance of systems over time, as uncertainty is resolved. This is done by reducing the negative impacts from downside scenarios, while enabling contingencies to capture upside opportunities.

Many systems have been designed flexibly in the past. Whether flexibility was actively pursued or not remains unclear, but nonetheless these examples are informative. The Boeing B-52 Stratofortress is one such example (Figure 2.9). It could be reconfigured to suit different missions and purposes, depending on exogenous operating conditions (Saleh & Hastings, 2000).

Figure 2.9: The flexible design of the Boeing B-52 Stratofortress's allowed adaptation to changing warfare conditions.

Developed in the 1950s, the bomber's configuration was changed a number of years later and in several occasions. The aircraft was originally designed to carry heavy and cumbersome nuclear warheads at high altitude (Montulli, 1986). The aircraft's large-scale belly was one of the main design features to accomplish this. A few years later, the Soviet air defense incorporated surface-air missiles, which forced the aircraft to fly at lower altitude. The belly was then reconfigured to carry air-launched cruise missiles to defend the aircraft through such mission (Boyne, 2001; Dorr & Peacock, 1995). This low-altitude capability was used later on during the Vietnam War to assist ground troop operations.

The Health Care Service Corporation headquarter building in Figure 2.10 is another example of flexibility in design (Guma, Pearson et al., 2009). Since the management board was not clear on its future needs for office space, it developed the project into two vertical phases. It designed Phase I explicitly so that it could accommodate another twenty-four stories, should the board decide to go on with expansion when needed. Phase I completed in 1997 consisted of thirty stories above ground. Phase II is currently under construction, aiming to complete the fifty-four story building before the end of 2010.



(a) (b)

Figure 2.10: Phase I (a) and Phase II (b) of the flexible Health Care Service Corporation development project (Guma, 2008).

Table 2.2 provides an overview of the design procedures developed over the last decades to support design for uncertainty and flexibility. The overview borrows from the surveys by Ferguson et al. (2007), Saleh et al. (2008), and Sethi and Sethi (1990). The taxonomy is similar to the one suggested by Tomiyama et al. (2006; 2009) and reflected in Table 2.1. There are design procedures developed to support flexible design concept generation, represent and manage the design process for a flexible engineering system, and explore the design space for flexible design alternatives. The main difference however is that design for flexibility integrates techniques to recognize, describe, and characterize uncertainty. It also integrates valuation techniques from the field of Real Options Analysis (ROA) to assess the economic value of flexible design alternatives.

Real Options Analysis

Evaluation techniques from financial options analysis (Arnold & Crack, 2003; Black & Scholes, 1973; Cox, Ross et al., 1979) and ROA (Dixit & Pindyck, 1994; Myers, 1977; Trigeorgis, 1996) have been adapted more recently to suit the needs of engineering design, and integrated to the design process over the last decade. These techniques are used to model uncertainty, and to evaluate performance of flexible design alternatives subject to a range of scenarios. They rely on economic metrics like NPV – measuring the sum of discounted cash flows over the project lifecycle – although evaluations can be done as well using non-financial metrics (e.g. service rate at a hospital). They can be integrated easily within screening models to evaluate and rank order different design alternatives, and to explore the design space efficiently.

Analytical tools better suited for ROA in an engineering context are binomial lattice, decision analysis, and Monte Carlo simulations (de Neufville, 2010). Binomial lattice and decision analysis both rely on the folding back principle from dynamic programming to assess the value of flexibility, and determine the best decision at each stage.

In decision analysis, a decision tree structure is created to represent uncertainty scenarios and associated decisions as time evolves. Figure 2.11 gives an example decision tree used to analyze the value of a flexible Accelerator-Driven Subcritical Reactor (ADSR) design compared to an inflexible alternative (Cardin, Steer et al., 2010), subject to uncertainty in technological development. An ADSR is an innovative nuclear technology suggested to produce emissions-free electricity by coupling a standard nuclear reactor core and a LINear ACcelerator (LINAC).

Table 2.2: Example design procedures for flexibility.

	Design Activities			
Design Procedures for Flexibility	Uncertainty Recognition	Concept Generation	Management and Representation	Design Space Exploration
During or before the 1990s				
Real Options Analysis (Dixit & Pindyck, 1994; Myers, 1977; Trigeorgis, 1996)	✓	✓		✓
2000s				
Design Decision-Making (Olewnik, Brauen et al., 2001; 2006; Ross, 2006)				✓
Design Structure Matrix (Bartolomei, 2007; Kalligeros, 2006; Mikaelian, 2009; E. S. Suh, 2005)			✓	
Industry Guidelines (Fricke & Schulz, 2005; Rajan et al., 2005; Skiles, Singh et al., 2006; Slack, 2005)		✓		
Screening Models (Hassan & de Neufville, 2006; Lin, 2009; Wang, 2005; Y. Yang, 2009)			✓	✓

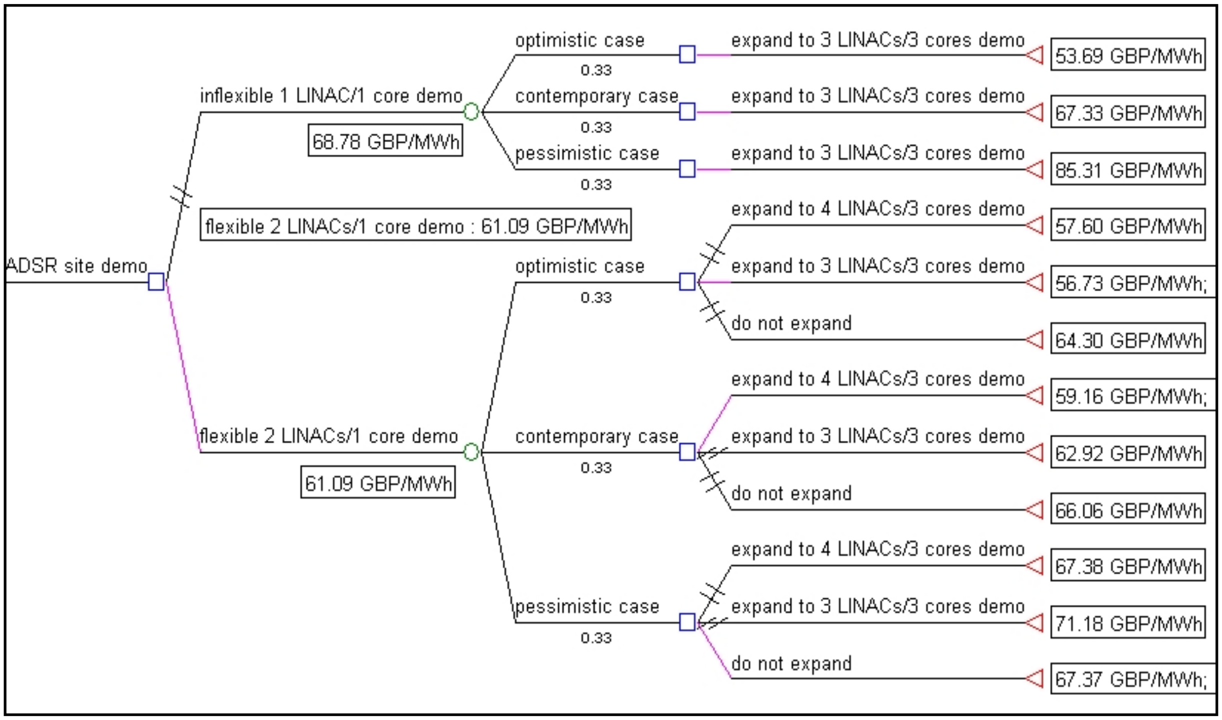


Figure 2.11: Example decision tree comparing a flexible ADSR design to an inflexible design alternative (Cardin et al., 2010).

Without focusing on the technical details of this system, one notices the structure of the tree as a sequence of decision (square) and chance (circle) nodes going from left to right. A stage consists of one sequence of decision and chance nodes. For instance, the first stage consists of choosing between the inflexible and flexible designs, and then observing manifestations of uncertain technology scenarios (optimistic, contemporary, pessimistic development scenarios). A state corresponds to each of the possible value outcomes at the end of the stage. The second stage involves no decision for the inflexible case (it is not possible to modify the design, and hence it is subject to the uncertainties without possibility to adapt), and the decision to change the system configuration in light of how technology evolved. The cost outcome for each path is shown at the far right of the decision tree. The analysis of a decision tree is done from right to left, applying a dynamic programming folding back process. The best decision is made at each decision point by calculating the expected value outcome, and pruning out suboptimal decisions. Hash marks in Figure 2.11 indicate suboptimal paths pruned out during the folding back process.

The main conceptual benefit of decision analysis in an engineering context is to have engineers go through the process of recognizing and factoring in uncertainty and flexibility explicitly. One analytical benefit is the freedom that the structure provides. Stationary and non-stationary stochastic processes can be depicted, different decisions can be used and evaluated at different stages, and different uncertainty sources can be considered in the same model. One important drawback is the curse of dimensionality. The number of paths in a decision tree typically explodes quickly, making it difficult to go beyond two or three stages, even with a minimum of decision and chance outcomes. On the other hand, a decision tree is useful as a first-pass analysis to recognize and incorporate flexibility in the design process. It can be complemented with a more detailed screening or simulation-based model of the system.

A binomial lattice is similar to a decision-tree, however the uncertainty (e.g. demand, price) evolves each time period to go up or down relative to the previous state (i.e. hence the name binomial). Figure 2.12 shows an example binomial lattice depicting the stochastic evolution of copper price – based on Geometric Brownian Motion (GBM) – over six years, starting at $2,000/ton.

Year	0	1	2	3	4	5	6
Price	2000.00	2210.34	2442.81	2699.72	2983.65	3297.44	3644.24
		1809.67	2000.00	2210.34	2442.81	2699.72	2983.65
			1637.46	1809.67	2000.00	2210.34	2442.81
				1481.64	1637.46	1809.67	2000.00
					1340.64	1481.64	1637.46
						1213.06	1340.64
							1097.62

Figure 2.12: Copper price evolution based on the binomial lattice approach.

Each time period (i.e. year) represents a stage, while a state is a particular price outcome as part of a stage. The stochastic process is described from left to right, as time evolves. In year 1, price can either go up to $2,210 or down $1,810 relative to the initial price. To reduce the number of possible outcomes, path independence is assumed, such that lattice nodes can recombine. This means the value of the system after an “up-down” sequence is the same as that after a “down-up” sequence – inspired from financial options analysis.

Figure 2.13 shows the Expected NPV (ENPV) of cash flows obtained when evaluating the flexibility to shut down a copper mine when copper price is too low – inspired from a case study in Chile (de Neufville, 2010). At each node a folding back process (from right to left) is used to determine the optimal decision based on the expected values of future outcomes. The numbers in each stage represent the optimal ENPV of cash flows generated by the mine, considering the possibility to shut down if the price of copper is too low. Optimal decisions to shut down (YES) or remain open (NO) are depicted for each state in Figure 2.14.

Year	0	1	2	3	4	5	6
ENPV(Cash Flow)	763,158	2,421,144	5,995,974	8,657,050	9,930,365	9,304,581	6,221,188
WITH SHUTDOWN OPTION		3,844,483	2,892,857	68,027	2,278,813	3,500,545	2,918,247
Dynamic programming			4,705,550	3,844,483	2,892,857	1,251,397	214,028
approach				5,484,675	4,705,550	3,844,483	2,000,000
(check next year)					6,189,657	5,484,675	3,812,692
						6,827,551	5,296,800
							6,511,884

Figure 2.13: Example binomial lattice depicting the folding back evaluation of the ENPV of cash flows to evaluate the flexibility to shut down a copper mine in case copper price is too low (de Neufville, 2010). Light figures are negative ENPV outcomes.

Year	0	1	2	3	4	5	6
Shut Down?	NO	NO	NO	NO	NO	NO	
WITH SHUTDOWN OPTION		YES	YES	NO	NO	NO	
Dynamic programming			YES	YES	YES	NO	
approach				YES	YES	YES	
(check next year)					YES	YES	
						YES	

Figure 2.14: Corresponding optimal decisions for each of the decision node in Figure 2.13. YES represents the states where the flexibility to shut down should be exercised, while NO represents states where the mine should remain in operations.

The binomial lattice is particularly useful when valuing sources of flexibility "in" and "on" systems that are similar to call options (e.g. capacity expansion, phasing) or put options (abandonment, temporary shutdown). The main benefit comes from the path independence assumption, which reduces the number of possible paths to a linear function of the number of stages. It is computationally simple and efficient, and therefore can be used as a first-pass analysis before a more detailed analysis. A lattice is essentially a discrete binomial formulation of the Black-Scholes formula (Black & Scholes, 1973) used to value financial options (Cox et al., 1979). Because of its discrete structure, it offers more flexibility for analysis in an engineering context than the Black-Scholes formula. On the other hand, the rigidity of the lattice structure makes it difficult to analyze more than one uncertainty sources at once – although it is feasible using a quadranomial lattice, described by Copeland and Antikarov (2003). The path independence assumption may not be realistic in an engineering context, since the value of a system after an up-down movement may be significantly different than for a down-up movement, because the sequence of decisions may differ. It is difficult to model complex managerial decision rules with a lattice. Also, the lattice evolution assumes a stationary process, which may not be realistic in many cases.

de Neufville, Scholtes et al. (2006) also suggest an approach for valuing flexibility based on Monte Carlo simulation. To provide transparency to practitioners, the method typically involves three steps:

1. A standard Discounted Cash Flow (DCF) analysis of an inflexible design using deterministic projections of the exogenous factor(s) affecting value;
2. A stochastic process simulating exogenous factor fluctuations over the project lifetime. Several stochastic scenarios can be simulated, and a DCF analysis is performed on the inflexible system for each scenario. This approach provides a distribution of value outcomes measured using a financial metric like NPV;
3. Flexibility is incorporated in the DCF model and valued using simple spreadsheet programming and logical statements (e.g. IF, ELSE, etc.). Under each stochastic scenario, the spreadsheet computes a NPV under the flexible design and decision rules incorporated in the model. This step also creates a Cumulative Density Function (CDF) – or target curve – of NPV outcomes, one upon which the designer may act by selecting different designs. The goal is to act on desirable properties of the entire distribution to take advantage of upside opportunities and reduce exposure to downside scenarios.

Figure 2.15 shows examples of target curves for three parking garage designs (de Neufville & Scholtes, 2011). The two curves for five and six "fixed" levels depict simulation outcomes for inflexible garage designs. The curve "Flexible Starting with 4 Levels" shows the range of outcomes for a flexible design starting with four initial levels, but having the capacity to expand to more levels as demand for parking space increases. This design captures both features of the two inflexible designs. It reduces downsides losses by limiting the initial capital investment, and captures upside opportunities by enabling expansion when needed.

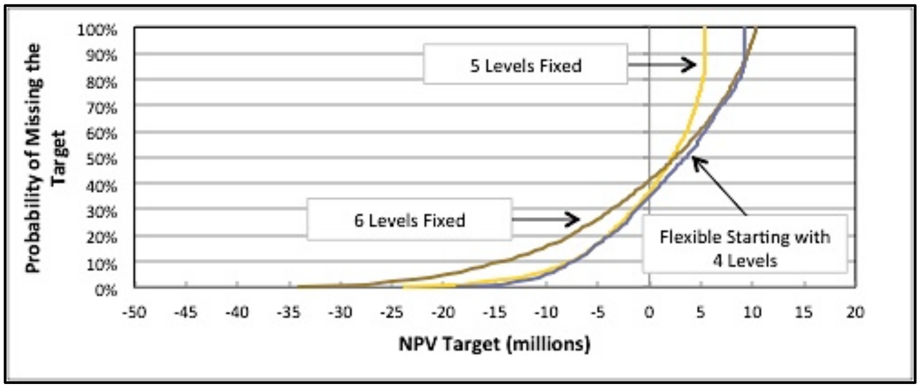


Figure 2.15: Example of CDF – or target curve – used to assess the value of flexible capacity expansion for a vertically built parking garage (de Neufville & Scholtes, 2011). The inflexible "fixed" designs are five and six level designs respectively, while the flexible design enables capacity expansion, starting with four levels.

The main benefit of Monte Carlo simulation is to provide more freedom in terms of the decision rules, design variables, and parameters to be modeled. In essence, it offers the possibility of a more detailed analysis. On the other hand, simulations may be more demanding computationally, especially when a high fidelity and detailed model of the system is developed. This is a tendency in engineering design that screening models help alleviate.

In terms of flexible design concept generation, Trigeorgis (1996) presents six canonical real options strategies useful to design flexible engineering systems design:

1. Defer investment and wait favorable market conditions to commit capital;
2. Change time-to-build, involving staged asset deployment over time;
3. Alter operating scale by expanding or contracting output production capacity;
4. Abandon the project with the possibility of reselling the physical asset at salvage value;

5. Switch production output and/or input;
6. Grow by providing future opportunities, such as investing in Research and Development (R&D).

These categories can be further divided as sources of flexibility "in" and "on" the system, as suggested by Wang and de Neufville (2005). Flexibility "in" the system requires in depth technical knowledge of the system design components (see examples below). It differs from flexibility "on" the system, which provides managerial flexibility without necessarily requiring technical inputs from engineers. From the categories developed by Trigeorgis (1996), investment deferral, abandonment, and growth options can be categorized as sources of flexibility "on" the system, as they do not require in depth engineering knowledge of the infrastructure design.

A time-to-build real option, as well as an option to alter operating scale and switch input/output can be associated to flexibility "in" engineering systems. Technology needs to be considered explicitly to enable such kinds of flexibility. For example, de Weck et al. (2004) show that deploying the Iridium satellite constellation in phases – upon observing actual demand realizations – rather than all at once – would have saved up nearly twenty percent in expected lifecycle cost. This might have saved the company from bankruptcy. This flexible strategy would however require designing each satellite with the capability to change orbital configuration as the constellation grows, to cover different geographical areas. This flexibility requires in-depth technical knowledge of the engineering system. Another example is the flexibility to alter oil production capacity of an offshore tension leg platform. Babajide, de Neufville et al. (2009) shows that designing additional slots in the platform to connect more direct vertical access wells allows production capacity expansion if more oil is discovered than originally expected.

Design Decision-Making

To support Design Decision-Making related to flexibility, Olewnik et al. (2001; 2006) extended Hazelrigg's DBD framework to incorporate considerations of flexibility in design space exploration. The framework in Figure 2.16 shows that the system design vector is already chosen from initial performance requirements (steps 1 and 2), and that design variables are made flexible to fulfill these requirements. The process evaluates possible design configurations, considering flexibility, and searches the one(s) providing the highest utility as proposed by Hazelrigg (1998).

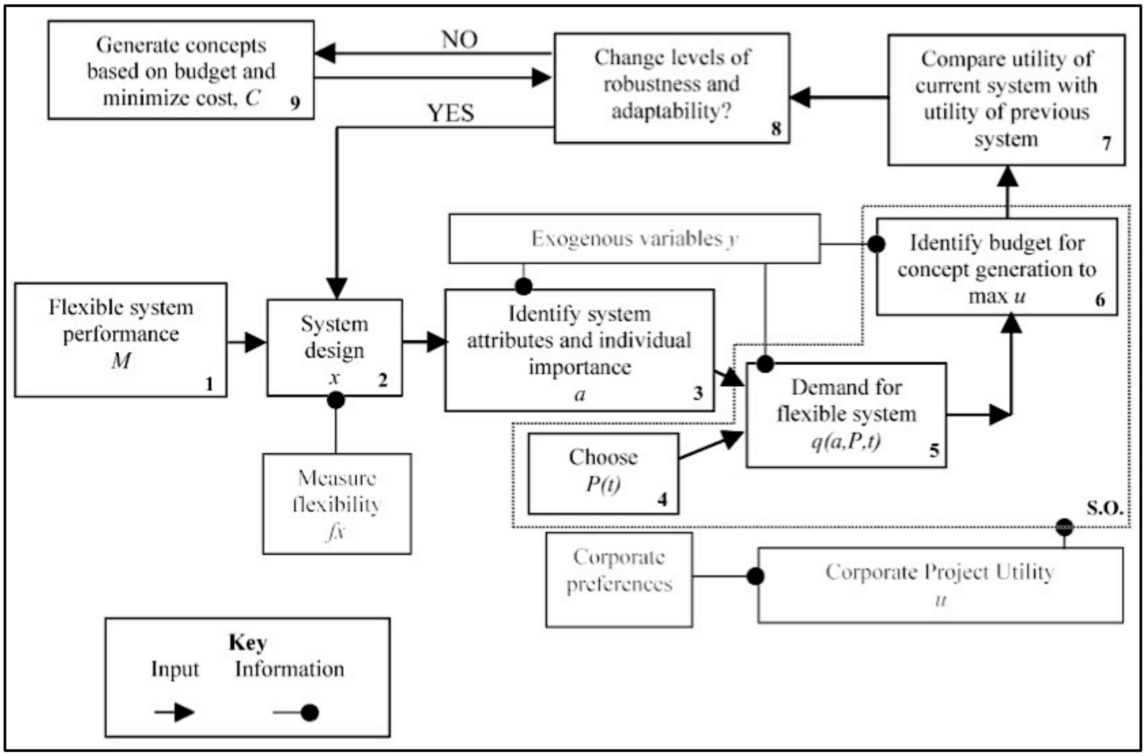


Figure 2.16: Visual representation of the DBD method to explore the design space of flexible systems (Olewnik & Lewis, 2006).

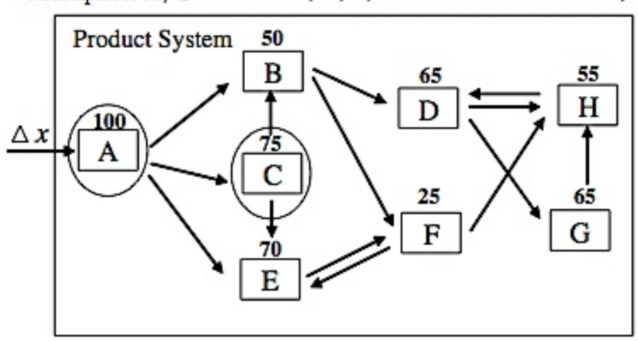
The framework has a step focusing on design concept generation (step 9), which borrows from standard brainstorming approach, but no particular attention is given as to how ones generate flexible design concepts. Ross (2006) also incorporated principles from utility theory to formulate his Multi-Attribute Tradespace Exploration (MATE) framework. The framework enables designers to explore the design space based on the configurations providing highest value, based on decision-makers perceived attributes of interest.

Design Structure Matrix

The DSM framework was studied extensively to support flexibility in design. Four methods – or algorithms – were developed for processing design information to identify and manage existing sources of flexibility within an engineering system: Change Propagation Analysis (CPA), sensitivity DSM (sDSM), the Engineering Systems Matrix (ESM), and the Coupled-DSM (C-DSM).

CPA combines the DSM with product platform strategies – see review by Simpson (2003). The technique looks at change multipliers as potential areas to insert flexibility (Figure 2.17). These are design elements creating more change in other design variables then they absorb when a design or functional requirement is changed (E. S. Suh, 2005). Making such variables more flexible reduces the amount of change created elsewhere in the design. It also reduces switching cost between possible product variants. CPA was used to identify opportunities for flexibility in a car body-in-white platform (E. S. Suh, de Weck et al., 2007), in a complex sensor system design (Giffin, de Weck et al., 2009), and in unmanned aero-vehicle design (Wilds, 2008).

Multiplier: A, C Carrier: B, D, F, G Absorber/Constant: E, H



Propagating Change / Receiving Change

	A	B	C	D	E	F	G	H	ΔE_{in}
A									0
B	1		1						2
C	1								1
D		1						1	2
E	1		1			1			3
F		1			1				2
G				1					1
H				1		1	1		3
ΔE_{out}	3	2	2	2	1	2	1	1	
CPI	**3**	**0**	**1**	**0**	**-1**	**0**	**0**	**-2**	
Class	M	C	M	C	A	C	C	A	

Figure 2.17: Representation of the CPA algorithm (E. S. Suh et al., 2007). ΔE_{in} and ΔE_{out} represent input and output changes respectively. CPI means Change Propagation Index.

sDSM (Figure 2.18) looks at design variables that are most sensitive to changes in design and functional requirements (Kalligeros, 2006) as potential areas to insert flexibility. It provides a high-level view of the design representation, “zooming out” from details to focus on important design elements to insert flexibility. Kalligeros (2006) demonstrated application of this approach for offshore oil platform design. The method helped identifying interesting design variables for other oil platform variants.

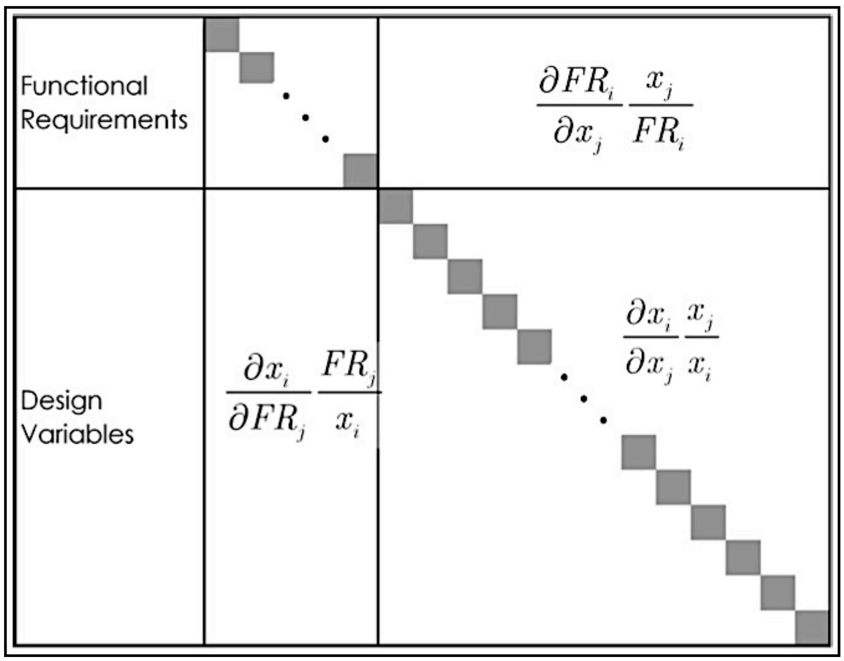


Figure 2.18: Representation of the sDSM algorithm for processing and identifying interesting sources of flexibility (Kalligeros, 2006). Design variables are represented by x_i, functional requirements by FR_j.

ESM emphasizes inclusion of social components in the DSM model, such as human stakeholders (e.g. managers, customers) and system drivers (e.g. purpose or mission) as Figure 2.19 shows. Bartolomei (2007) suggested incorporating CPA and sDSM within the ESM framework to identify sources of flexibility. The ESM framework also promotes use of good qualitative research methods to extend the model boundaries to socio-technical considerations.

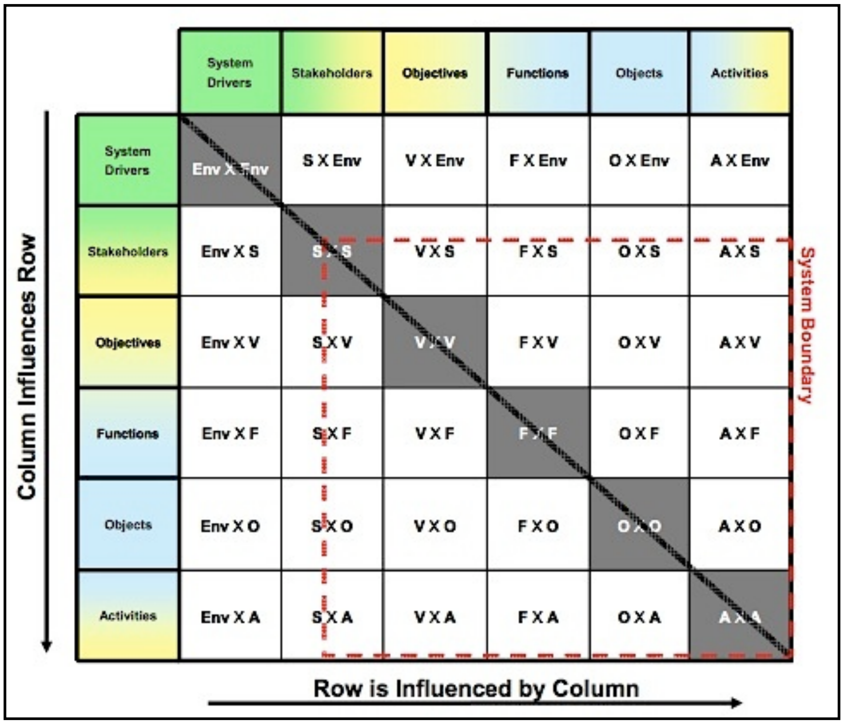


Figure 2.19: Visual representation of the ESM approach (Bartolomei, 2007).

Mikaelian (2009) suggested extending the DSM framework to an enterprise view (Figure 2.20) through the C-DSM. This allows consideration of a wider range of opportunities to insert flexibility, for instance at different levels of the enterprise structure. Flexibility can be investigated in terms of enterprise strategy, policies, organizational structure, process, product, service, and expert knowledge. The author presented a new characterization of flexibility as a tuple consisting of a flexibility mechanism and type. This characterization was used in the novel algorithm developed to process DSM information in search for valuable flexibility sources. The framework was applied to identify and value opportunities for flexibility in an unmanned air vehicle system and uncertainty management in surveillance missions.

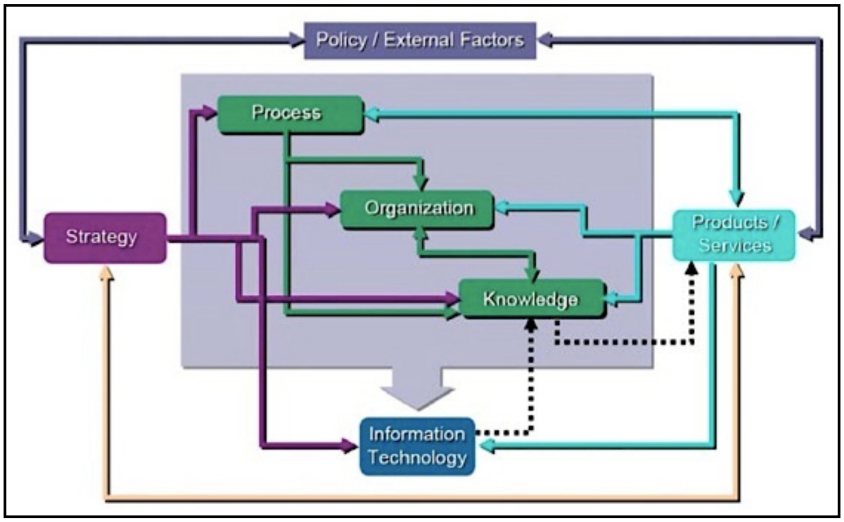


Figure 2.20: Example of dependencies among an enterprise view used in the C-DSM framework (Nightingale & Rhodes, 2007).

Industry Guidelines

Many authors have published guidelines based on past lessons in industry to support flexible design concept generation. Fricke and Schulz (2005) provided principles to enable changeability in systems engineering based on lessons in the automotive industry. They suggested basic principles for flexibility, similar to the approach employed in TRIZ. Example principles are:

- Ideality/Simplicity, aiming at reducing system complexity;
- Independence of design parameters, similar to N. P. Suh's definition (1990);
- Modularity/Encapsulation, to ease exchanging and adapting modules over time, and relying on the DSM approach described above.

Rajan et al. (2005) provided guidelines from experience with the product development sector:

- Design for modularity, because as the design becomes more integrated, it becomes more inflexible for redesign;
- Design modules as external attachments to each other;
- Design with more standard components and interfaces to facilitate integration over time;
- Partition a design into a greater number of elements (manifested through higher numbers of components and functions);

- Reduce the number of parts within each module.

Skiles et al. (2006) listed underlying principles from transformer theory. The principles are:

- Expand/Collapse, in the sense of changing physical dimensions of the product or object, similar to the capacity expansion strategy mentioned by Trigeorgis (1996);
- Expose/Cover the surface of the object to alter functionality, similar to the modularity principle providing adaptable interfaces;
- Fuse/Divide by enabling a single device to become two or more devices, which naturally arises from the independence principle.

Slack (2005) provided guidelines to insert flexibility in the manufacturing based on interviews with ten company managers. The author identified four general areas to insert flexibility, similar to the ones suggested by Trigeorgis (1996):

- Product changes, in the sense of enabling flexibility to switch between different products;
- Making different mixes of products;
- Adjusting the volume output;
- Changing delivery dates.

Screening Models

Many authors have taken a screening approach to explore the design space for flexible design concepts. As explained in Section 2.2.1, the screening model enables designers to explore the design space efficiently, and to identify valuable opportunities for flexibility. Different methods can be used to model uncertainty (e.g. lattice, Monte Carlo simulations), to evaluate design alternatives (e.g. ROA), and to structure the search process (e.g. optimization, DBD, DOE, MATE). Figure 2.21 depicts the screening framework developed by Lin (2009).

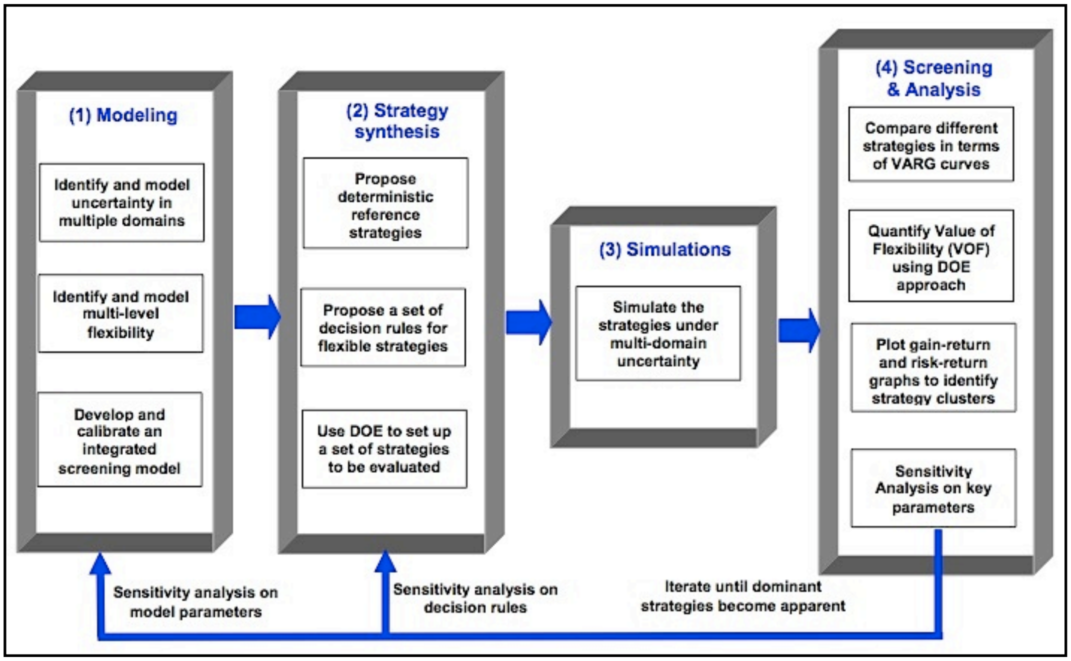


Figure 2.21: Suggested screening framework suited for flexibility analysis (Lin, 2009).

Wang (2005) was first to use screening models in the context of flexibility. He relied on mixed-integer programming to screen different flexible hydroelectric dam strategies in China. Hassan and de Neufville (2006) used a genetic algorithm to structure the exploration process for oil platform design. Lin (2009) developed a screening approach for offshore oil platform design. Y. Yang (2009) incorporated an efficient DOE algorithm called adaptive One Factor At a Time (aOFAT) (Figure 2.22) to screen flexible car manufacturing plant alternatives.

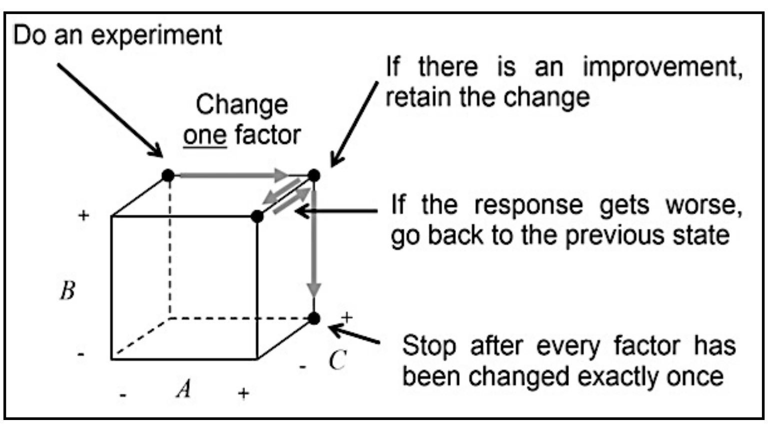


Figure 2.22: Representation of aOFAT to explore the design space (Frey & Wang, 2006).

2.3 Design Procedure Evaluation Methodologies

A methodology that is both effective and efficient is needed to evaluate a design procedure for flexible design concept generation. In this context effective means that the methodology can successfully produce a desired or intended result. In this case, the design procedure aims at improving anticipated performance of design concepts – because one needs to compare to the cost of acquiring the flexibility. Therefore one needs a methodology that can measure objectively and quantitatively anticipated performance improvements compared to a benchmark design. The methodology should also be efficient, in that it helps achieve maximum productivity with minimum wasted effort or expense.

This section surveys the literature in cognitive sciences, collaboration engineering, and engineering design to determine the state of the art in terms of design procedure evaluation methodologies. The focus is in particular on studies involving notions of design concept generation – under different forms. The thesis builds upon this body of research to suggest a complementary approach to evaluate the design procedure for flexibility based on anticipated performance of design concepts.

2.3.1 Categories of Methodologies

Methodologies for design procedure evaluation can be classified in general under three categories (Shah et al., 2000; 2002; Summers, Anandan et al., 2009):

1. Case studies;
2. Controlled user studies;
3. Protocol studies.

Case studies involve interviewing and/or working closely with a company to understand better or test a design procedure of interest (Summers et al., 2009). Controlled user studies – the approach pursued in this thesis – involve controlled experiments on short design problems that can be repeated efficiently, quickly, and rigorously with many groups of participants – typically undergraduate and graduate students, but sometimes practitioners. Protocol studies involve smaller number of designers "thinking out loud" while pursuing design activities. Design activities are videotaped, transcribed, and design data are analyzed to understand underlying cognitive processes (Finger & Dixon, 1989a). These studies may focus on design concept generation, and also other components of the design cycle.

2.3.2 Evaluation Metrics

Metrics used to evaluate design procedures for concept generation focus either on the process – how design activities are performed – or outcomes generated during design activities (Shah et al., 2000; 2002). In the former case, researchers are interested in determining whether a procedure helps designers perform a certain task better, or feel more comfortable with it than with another procedure. Studies focus on the cognitive processes in designers' minds, the subjective impressions of satisfaction with a procedure, or the time taken to perform the procedure. In the latter case, researchers are interested in the effect of a procedure on the design concepts generated or evaluated. Such studies often measure the quality and quantity of concepts generated, but there are other metrics in use. The development of relevant metrics to evaluate design procedures is an emerging field in and of itself (Otto & Holtta, 2007).

Table 2.3 summarizes example studies making use of different evaluation metrics, in alphabetical order within the Case Studies, Controlled User Studies, and Protocol Studies categories. The top row determines whether the evaluation metric focuses on the outcomes or process. The second row specifies the evaluation metric of interest. The first column on the left determines the type of example study. The second column gives the reference to the study of interest. Check marks in the remaining columns represent the metrics used in a given study. Only widely used and more recent evaluation metrics are considered.

The first metric looks at certain attributes of the design concepts, like anticipated cost and weight, to discriminate between design procedures. Attributes are matched to designers/participants' preferences and utility to evaluate the effects of the procedures (Hazelrigg, 1998; Hevner, March et al., 2004). Completeness is used as metric to measure how much a given concept variant addresses a sub-function depicted in the function structure (Kurtoglu, Campbell et al., 2009). Linkography evaluates the degree to which ideas are connected to each other in the design process (Goldschmidt & Weil, 1998; van der Lugt, 2002).

Table 2.3: Example studies making use of different metrics for evaluation of procedures for design concept generation.

Evaluation Metrics Based on:		Outcomes									Process	
	Example Studies	Attributes	Completeness	Linkography	Technical Feasibility	Novelty	Quality	Quantity	User Satisfaction	Variety	Generation Time	User Satisfaction
Case Studies	Avnet, 2009	✓					✓				✓	
	Stagney, 2003	✓					✓				✓	✓
	Ward et al., 1995						✓					
Controlled User Studies	Boehm et al., 2001								✓			✓
	de Vreede et al., 2002											✓
	Frey et al., 2009	✓					✓					
	Kolfschoten et al., 2009								✓			✓
	Kurtoglu et al., 2009		✓			✓				✓		
	Linsey et al., 2010							✓				
	Linsey et al., 2005				✓	✓		✓				
	Ostergaard et al., 2005							✓	✓		✓	
	Reinig et al., 2007						✓	✓				
	Santanen et al., 2004						✓	✓				
	Shah et al., 2001					✓	✓	✓	✓	✓		✓
	van der Lugt, 2002			✓								
	Wetmore et al., 2010							✓				
	Yang, 2009						✓	✓				
Protocol Studies	Goldschmidt and Weil, 1998			✓								
	Ullman et al., 1989							✓	✓			✓

The more a design procedure connects ideas, the better. Technical feasibility measures the degree to which a design concept can be implemented practically. It may rely on a scoring mechanism based on expert judgment to discriminate between design procedures.

Novelty, quality, quantity, and variety are defined by the foundational work of Shah et al. (2000; 2002). They also all rely on subjective expert judgment, and represent the metrics most widely used in design studies focusing on concept generation. Novelty is a measure of how unusual or unexpected an idea is as compared to other ideas. Quality measures the feasibility of an idea, and how close it comes to meet design specifications. Several metrics rely on ideation quality, like good idea count (number of ideas above a given quality threshold), sum of quality (summing the total quality of all ideas generated), and average quality among a set of ideas (Reinig et al., 2007). Quantity is the total number of ideas generated by a group under a certain idea generation method. Quantity can be features added or removed during design activities (e.g. Linsey, Tseng et al., 2010; Ostergaard, Wetmore et al., 2005; Shah, Vargas-Hernandez et al., 2001; Wetmore, Summers et al., 2010). Variety is a measure of the explored solution space during the idea generation process. It is similar to the range of ideas metric defined by Guilford (1959), or to the quantity of non-redundant ideas and quality ratings defined by Mullen, Johnson et al. (1991).

User satisfaction with the results and process are defined as affective positive arousal towards the results and process respectively (Briggs, Reinig et al., 2006). Generation time is defined as the time required going through a given design procedure.

The generic description of evaluation metrics above is complemented in the following subsections with example studies in the three categories of Case Studies, Controlled User Studies, or Protocol Studies. This intends to give the reader a better grasp at the kind of metrics used in different studies to evaluate different design procedures.

2.3.3 Example Case Studies

In this case study, Avnet (2009) measured how changes in shared knowledge between the beginning and end of concurrent design sessions at NASA correlated with design attributes of design concepts like cost, time to launch (i.e. design development time), and weight. He also evaluated the effects on a quality metric called Mission Technology Readiness Level – a scoring mechanism measuring holistically mission concept maturity. Stagney (2003) collaborated with a lead U.S. aerospace firm to study how development time and cost attribute of design concepts are affected when using an integrated concurrent design process, versus a more traditional design approach. The author also collected information about anticipated quality of outcomes, as judged by designers working at the firm. Ward, Liker et al. (1995) studied qualitatively the effect on economic performance of set-based concurrent engineering – generating a set of design alternatives and eliminating many alternatives, rather than iterating from an initial proposal – in Japanese and U.S. automakers in the 1980s and 1990s. This is closer to a quality evaluation metric, although it incorporates some notion of economic performance evaluated qualitatively. The authors interviewed managers, engineers, and suppliers at U.S. and Japanese firms to determine to what extent set-based concurrent engineering is in use at different firms, and how it relates to economic performance.

2.3.4 Example Controlled User Studies

The literature on collaboration engineering provides insights into user satisfaction with process and outcomes when different ideation and moderation conditions are used. Collaboration engineering "studies ways of designing recurring collaboration processes that can be transferred to groups that can be self-sustaining in these processes using collaboration techniques and technology" (de Vreede & Briggs, 2005). The goal of ideation sessions is to collectively suggest solutions to problems in management and engineering, and in various industrial sectors.

In terms of example controlled user studies in the collaboration engineering literature, Boehm, Gruenbacher et al. (2001) discussed satisfaction with a proposed design requirement negotiation Group Support System (GSS) and results by using qualitative inputs and comments from over thirty industrial partners. de Vreede, Boonstra et al. (2002) used qualitative interviews to study user satisfaction with meeting facilitation process supported with GSS. GSS can be defined broadly as "socio-technical systems consisting of software, hardware, meeting procedures, facilitation support, and a group of meeting participants engaged in intellectual collaborative work" (2003). Kolfschoten, Briggs et al. (2009) compared user satisfaction with the outcomes/results and process under different ideation moderation and GSS techniques by asking undergraduate and practitioner participants for inputs like "Very satisfied with the procedure/results" or "Not satisfied with the procedure/results". Reinig et al. (2007) studied the effect of invocation of social comparison on the ability to generate creative solutions to a problem by inserting comments such as "You are doing great!" when a team performed above a hypothetical group average, or "Have you fallen asleep?" when they performed lower. They counted the number of proposed solutions to two hypothetical management problems in business schools, the number of good solutions, the average quality, and sum of quality through expert ratings based on 5-point Likert scales. Santanen, Briggs et al. (2004) used similar mechanisms to study the effect of providing prompting stimuli at different rates on creativity and outcome quality.

In the engineering design research literature, Frey et al. (2009) evaluated the PuCC mechanism by simulating designers' collective quality assessments of desired attributes for different design alternatives. Kurtoglu et al. (2009) and Kurtoglu and Campbell (2009) studied the effects of a computer-aided design tool on completeness, novelty, and variety of design concepts generated through artificial intelligence principles. Linsey et al. (2010) determined the quantity of features reproduced from an initial design sketch to study design fixation in a group of engineering design faculty. Linsey, Green et al. (2005) also studied the effect of a functional modeling technique on concept novelty, quantity, and technical feasibility. Ostergaard et al. (2005) measured the effect of different communication mechanisms on the ability to identify the number of design flaws in a design review process. They measured satisfaction with results to some extent as they asked whether participants thought they had completely identified all design flaws. They also measured the time to get started on the collaborative design task, depending on the communication mechanisms allowed. Shah et al. (2001) compared the C-sketch procedure to the Gallery and 6-3-5 methods based on the number of design features added and removed through design cycles, as well as novelty, quality, and variety of design concepts. They also provided a post-experimental survey inquiring about user satisfaction with the process and overall results. van der Lugt (2002) used linkography to evaluate the effects of a brainstorming technique on concept generation, versus brainsketching – which relies on sketching as the primary means of recording ideas. Wetmore et al. (2010) used a similar "design flaw quantity" metric as Ostergaard et al. (2005) to study the effect of group familiarity and the level of information sharing in the design review process. M. Yang (2009) studied correlations between the quantity and quality of design outcomes in concept generation, relying on project course grades and ranking in a design contest.

2.3.5 Example Protocol Studies

In terms of protocol studies, Goldschmidt and Weil (1998) studied the effect of different "design moves" – bringing the design from one state to a more advanced state, by suggesting a critical idea for instance – on the linkography, or the connectedness of concepts generated during design activities. Ullman, Wood et al. (1989) demonstrated the importance of drawing and sketching for designers as part of design activities. They counted the number of design features generated, and analyzed marks and annotations to understand the effects of drawing activities on creativity and user satisfaction.

2.4 Research Opportunities

2.4.1 Related to Established Design Procedures

There is a clear opportunity to enhance established design procedures by having designers recognize and incorporate more systematically uncertainty and the flexibility to adapt to changing circumstances early on in the design cycle. Although well known design procedures from Table 2.1 have been very successful at assisting designers and engineers in creating impressive and well-functioning engineering systems, not all procedures clearly and systematically do this (de Neufville & Scholtes, 2011). This may leave aside significant opportunities for value and performance improvements. Rather, typical approaches to design assume that (de Neufville, 2010):

1. Customers know their needs, *but new ones may emerge*;
2. Design requirements are known, *but these may change with needs and new regulations*;
3. The system can be designed as a coherent whole, built and deployed in one step, *whereas this is often not possible*;
4. Only one system is being designed, *but families are most likely considered*;
5. The system will operate in a stable environment as far as regulations, technologies, demographics, and usage patterns are concerned, *we wish…*

Under these assumptions, a widespread approach in design is to simplify considerations of uncertainty by forecasting the future for cost, demand, prices, resource quantities, margins, and/or regulations, to determine the design requirements to satisfy these scenarios, and accordingly find the optimal system design configuration. For example in Axiomatic Design (Figure 2.4), a more detailed description of the design process starts from the assumption that customer demands and preferences are known, which triggers a definition of relevant design requirements. A "zigzagging" process is used to navigate, and go back and forth between the different domains. Although N. P. Suh (1998) recognizes the need to design flexible systems, the procedure does not clearly indicate where to consider uncertainty and flexibility as part of the design process.

A similar observation is made about the popular procedure developed by Pahl and Beitz. As seen in Figure 2.5, conceptual design and planning involve first determining market preferences, followed by a definition of design requirements. There is no explicit step recognizing variability in design variables, parameters, and operating conditions.

TRIZ on the other hand does provide some mechanisms to explicitly consider uncertainty and flexibility (Straker & Rawlinson, 2003). The principle of Prior Action can be viewed as creating some alternative flexibility, and being ready for a potential future event in case of occurrence. Similarly the Dynamicity principle means creating systems that are able to cope with change and intrusions from outside.

2.4.2 Related to Design Evaluation and Selection

An important opportunity is to develop a design procedure explicitly and systematically recognizing uncertainty and flexibility early in the design cycle. Design procedures relying on deterministic forecasts may lead to incorrect requirement definitions, design evaluations, and ultimately design selection, for the following three reasons.

First, research has shown that expert forecasts can be biased and incorrect for a number of reasons (Morgan & Henrion, 1990). Hence, it is most likely that exogenous uncertainties like market costs and prices, or endogenous ones like technology development will not turn out as planned for the entire project lifecycle.

Second, even if forecasts are correct (which is highly unlikely), Savage's "Flaw of Averages" (2000) shows that any decision based on the "average" or "most likely" scenario may lead to incorrect results, and bad investment decisions. This is a consequence of Jensen's inequality for systems with a non-linear response:

$$E[f(x)] \neq f(E[x])$$

This means that the expectation of a system response $E[f(x)]$ to input x is not the same as the response $f(E[x])$ of the expected input x. The net result is that the anticipated – or expected – performance of a design concept is different from the anticipated performance when only one central or most likely scenario is used for valuation. One reason for systems limited by capacity is that the benefits generated by upside scenarios (e.g. high price or demand) are limited, such that on average, the effect of low demand, loss-generating scenarios cannot be exactly counterbalanced.

Third, typical DCF valuation methods used to evaluate projects do not account for the fact that uncertain factors will inevitably change over the long lifecycle of an engineering system. Traditional valuation methods assume full commitment at $t = 0$ to a particular deployment path or strategy over the entire lifecycle. The reality is that things will change along the way, and managers will adapt to keep operating the system in the best available conditions. This reality is not captured in traditional valuation methods (Dixit & Pindyck, 1994; Trigeorgis, 1996). This can significantly affect investment decisions on large-scale technology deployment, as many case studies demonstrate[1].

2.4.3 Related to Concept Generation Procedures

There is an opportunity to develop a novel concept generation technique dedicated to the issue of helping designers generate and identify opportunities for flexibility early in the design cycle. This opportunity arises because existing IG techniques are more generic, and do not necessarily focus on the specific issue of design for flexibility. A more refined survey of IG techniques by Shah et al. (2000; 2002) as well as Knoll and Horton (2010) shows that many techniques can however be modified and developed with such focus.

Among existing design procedures for flexibility in Table 2.2, there appears to be no analytical, systematic, and creativity-based method to support explicitly flexible design concept generation. The closest techniques are based on Industry Guidelines and ROA, and are more akin to History-Based methods in the taxonomy by Shah et al. (2000; 2002), or to Analogies in the taxonomy by Knoll and Horton (2010). These techniques are more suggestive, not necessarily systematic, and not directive. This provides an opportunity to contribute to the existing portfolio of procedures for flexible design concept generation, which also appears to fit partially within the analytical and transformational categories defined in Figure 2.8.

There is a community interested in developing computer-aided techniques to support concept generation (Kurtoglu & Campbell, 2009; Kurtoglu et al., 2009). For instance Ward and Seering (1993) present a computer program proposing optimal component selection given a designer's utility function for hydraulic power transmission system. This topic is however out of scope of the current thesis because it focuses on automatic generation of design concepts. This thesis focuses on techniques to support design concept generation by humans – as opposed to artificially intelligent machines.

2.4.4 Related to Design for Flexibility

There is an opportunity to enhance current design procedures focusing on the issue of flexibility by providing a simple and user-friendly approach to generate flexible design concepts quickly and early in the design cycle. The goal in this case is not to be flexible at all costs, but to find ways to improve expected performance of the engineering system by means of flexibility, as demonstrated in many case studies.

However valuable design for flexibility may be, there are many reasons why it can also be a challenging process. Enabling flexibility in design requires careful analytical considerations in the early phases of design. Novel analytical tools developed with the purpose of helping designers enable flexibility should consider these challenges. The specific structure of the design procedure proposed in this thesis hopes to address some of the challenges described below.

One reason why flexibility in design is challenging is because there is no “one fits all” solution suiting all engineering systems. Each system is different, and is subject to different uncertainty sources. An infinite number of uncertainty sources can affect the performance of systems. It is difficult to identify important ones to focus the design effort. Equally, a considerable number of flexible strategies can be explored, depending on the system (e.g. phase capacity deployment, alter operating scale, switch product input/output, abandon or temporarily shut down activities, delay investment, etc.). Designers need to identify valuable opportunities, and engineer relevant design variables and parameters to enable flexibility. Furthermore, they may need to negotiate legal and/or financial disposition to enable flexibility.

There is a need to provide guidance and support to generate flexible design concepts. Industry guidelines and canonical strategies from the real options literature are useful, but they are suggestive, and do not provide a clear structured mechanism to guide designers through the appropriate thought process. Techniques relying on DSM help designers identify opportunities for flexibility already embedded within the boundaries of the system defined by the DSM through a detailed analysis of the system interconnections. Similarly to industry guidelines, DSM techniques do not provide a clear mechanism to explore opportunities for flexibility lying outside the set boundaries. In addition, collecting data and interviewing relevant actors can be a demanding task to construct a DSM. Such detailed analysis may require a very long time before algorithms like CPA, C-DSM, ESM, and sDSM can be applied to identify opportunities for flexibility. Design decision-making methods and screening models typically assume that design concepts are already generated. They provide an efficient and useful way to explore the design space for flexibility, but not explicitly to generate flexible design concepts. Therefore, there is a need for a simple and quick approach to identify opportunities for flexibility, and generate rapidly flexible design concepts. This technique may enable designers to capture “low hanging fruits” bringing significant and considerable value improvements.

Another reason is that designers operate within institutional, possibly cultural, engineering "silos" and do not consider how other system components might affect the overall economic value of the system. Dong (2002) showed that system-level knowledge required to think about flexibility is not well documented across different systems disciplines in the car manufacturing industry. To reinforce this assertion, it took Lin (2009) about a year of close collaboration with oil platform engineers to exploit the design of sub-sea tiebacks as a valuable opportunity for flexibility. This is not because designers did not know or did not think the flexibility would be valuable, rather they were not actively engaged in discussions with sub-surface engineers to consider this opportunity.

Designers may think they adequately consider uncertainty and risk when they subject a design to a range of scenarios through sensitivity analysis after an initial design is crafted. This approach, however, does not consider uncertainties in the early conceptual phase prior to more detailed design analysis. It does not recognize the power of adapting pro-actively to changing future conditions, and the potential to increase economic value by doing so.

It might be as well that engineering practice focuses too much on the use of detailed (exact or high-fidelity) models. Such models are often computationally expensive and cannot be used to explore many design configurations quickly, including flexibility and managerial decision rules, under a wide range of uncertain scenarios.

Finally and most importantly, designing extra contingencies for flexibility may require additional upfront costs. Therefore, designers must be prepared to justify the extra cost objectively and quantitatively, as there are cases where flexibility may cost too much, and is not worth the extra investment. In reverse there are also cases where flexibility comes for free, or lowers the initial capital expenditures, which should also be recognized explicitly (de Neufville et al., 2006). Many analytical tools can be used to measure quantitatively the expected value of flexibility based on ROA (Dixit & Pindyck, 1994; Trigeorgis, 1996), decision analysis, lattice analysis, and Monte Carlo simulations (e.g. Copeland & Antikarov, 2003; 2010; de Neufville et al., 2006). As explained above, these tools enable clear comparisons between the expected value of flexibility and its cost of acquisition, but their use is not yet widespread in design practice.

2.4.5 Related to Design Procedure Evaluation Methodologies

There is an opportunity to enhance and complement existing methodologies used in engineering design research to evaluate design procedures. This can be done by developing an experimental methodology incorporating existing techniques to study the effects on creativity, find out about user impressions of satisfaction with process and results, and quality assessments of results. The main addition of such experimental methodology would be to rely on objective and quantitative evaluation of the procedure based on the observed effects on anticipated performance of design concepts. In addition, this methodology should be rigorous and thorough, easily replicable, and should enable efficient data collection and analysis.

One reason creating this research opportunity is that existing evaluation methodologies based on case studies and protocol studies offer potential for detailed studies, but they cannot be reproduced efficiently and rigorously enough to evaluate effects of design procedures through statistical means. Even though they are absolutely essential to design studies, these methods can be difficult to apply, are time consuming, may suffer from poor or non-existent documentation of ideas generated during a project, or have limited access to important information for proprietary or security reasons (Shah et al., 2000). Different systems operate in different environments and contexts, so it is difficult to make general conclusions, even from meta-study analyses. Each study provides only one sample data, and acquiring a significant number of data can be daunting.

Another reason is that existing evaluation metrics make it difficult to assess objectively and quantitatively anticipated performance of design concepts generated early in the design cycle. Existing methods cannot answer objectively and quantitatively research questions like "does the design procedure help improve the anticipated economic performance of the system over its intended lifecycle?" or "can the design procedure improve the number of design concepts generated that have an expected value of more than $10 millions?" Measuring attributes like anticipated cost and weight of design concepts offers no measure of their future performance in operations. Generating design concepts with high levels of completeness, connectedness (as in linkography), technical feasibility, novelty, and variety will not necessarily result in good performance on the field. Having a high degree of satisfied users with the process and results offers no guarantee the system will perform well in the future. Also, generating many ideas does not necessarily imply good overall quality of outcomes (Taylor, Berry et al., 1958). All of these may provide better performing design concepts, but correlations yet have to be demonstrated.

Of all the metrics above, quality is the closest conceptually to enable assessment of anticipated performance of design concepts, although it does not satisfactorily does that. Quality is typically estimated based on subjective expert assessments, relying on some form of scoring mechanism. Experts rate proposed solutions from a holistic standpoint using, for instance, a 10-point Likert scale where 0 is a bad solution, and 10 is an excellent solution.

Based on the overviews by Reinig et al. (2007) and Shah et al. (2000; 2002), design procedures can be evaluated using a quality metric based on:

- The total quality score from unique solutions (i.e. sum of quality);
- The weighted sum of quality determined by different weights assigned to the functions that each design concept should fulfill (i.e. weighted sum of quality);
- The average quality determined from many unique solutions (i.e. average quality);
- The number of unique solutions exceeding a threshold for "good" quality (i.e. good idea count).

These quality-based criteria have several limitations. For instance, generating more bad ideas can increase the total sum or weighted sum of quality scores, and also degrade the average quality score even if superior ideas exist in the set. Reinig et al. (2007) suggested that the "good idea count" metric is the most internally valid of the above quality-based metrics. It is not clear, however, how one justifies the choice of a particular threshold for good quality (e.g. 9 out of 10 versus 10 out of 10). Conclusions of the study may change depending on this threshold. Also, it is not obvious how to weigh assessments from different experts, based for instance on experience, or the weight of different functions. If some experts have more experience with the topic, should one give the same weight to all expert ratings? How should one determine the appropriate weight that each function the system must fulfill? Balthazard, Ferrell et al. (1998) investigate some avenues to resolve these issues.

Another issue with existing quality metrics is that a high quality score offers no guarantee the system will actually perform well in the future. Anticipated performance should be assessed more directly than through subjective means. Experts can hardly include all the complexities of design activities when using a finite scale scoring mechanism. This is because too many decision rules, design variables, operating conditions, and parameters need to be considered in typical design activities. Finer resolution than provided by Likert scales is needed to discriminate between two ostensibly appropriate design alternatives that could receive similar scores, and where one may be superior in reality.

2.5 Anticipated Contributions

2.5.1 A Procedure for Flexible Design Concept Generation

A novel procedure is described in Chapter 4 to help designers generate flexible design concepts based on a lecture on the topic of flexibility, and a prompting ideation mechanism geared towards flexibility. The procedure addresses the issue that established design procedures often optimize system configurations for deterministic forecasts and design requirements, assuming customer demands and preferences are known *a priori*. The systematic design procedure introduced here complements existing ones by focusing early design efforts on important sources of uncertainty affecting anticipated performance, and guiding the thought process to create valuable strategies to deal with these uncertainties flexibly, both in the design process and operations. The design procedure fits well within the taxonomy by Shah et al. (2000; 2002) as a mix of analytical and transformational IG procedures. It extends existing design concept generation methods by focusing on the specific issue of uncertainty and flexibility. It contributes to the work in collaboration engineering often focusing on identifying important sources of uncertainty: it helps designers craft flexible strategies in design and management to deal with uncertainty pro-actively.

The proposed design procedure addresses some of the concerns raised in Section 2.4.4 as to why designing for flexibility is not yet a widespread approach to design. Given the proposed approach is generic, it can potentially be applied to any engineering system – although this is not proven explicitly in this thesis, and left as future work. The approach relies on designers' expertise, and does not require a lengthy DSM construction before valuable opportunities can be identified. The fact that the design procedure relies on a prompting mechanism is conducive of collaborative design activities, which encourages designers from different engineering expertise to come out and work together on the problem of interest. It is suggested that the design procedure be used in conjunction with ROA tools and screening models described in Section 2.2.3 to assess the value of flexibility. This approach is fundamentally different from performing a simple sensitivity analysis once an optimal design is selected, as it recognizes the managerial ability to adapt as environmental, market, regulatory, and technological conditions change.

2.5.2 A Methodology for Quantitative Performance-Based Design Procedure Evaluation

The experimental methodology in this thesis proposes evaluating the effects of a design procedure on anticipated performance of design concepts both objectively and quantitatively. The proposed approach assumes that this can be best done through analytical means supported by computer modeling, as suggested by Kurtoglu and Campbell (2009). Computer modeling techniques complement existing quality-based metrics for more thorough holistic evaluation, relying for instance on Computer-Aided Design (CAD) software, financial spreadsheets, optimization algorithms, etc. These techniques alleviate some of the concerns raised previously. For example, explicit modeling provides an objective baseline for determining the threshold for good design concepts (i.e. by modeling the performance of a benchmark design, and good ideas are the ones improving value compared to the benchmark). There is no need to determine subjectively the threshold for good quality. Computer models can integrate a larger set of design variables, objective functions, parameters, and future operating conditions that would otherwise be difficult for a single expert to consider holistically. Computer models can provide finer output resolution than Likert scales to discriminate between technically feasible design alternatives. They can also enable discriminating between two ostensibly good design alternatives that could receive similar scores. They do not require a particular weight assignment on expert judgment or design function.

In order to develop a methodology that is both effective and efficient – so it can be applied quicker than typical case studies and protocol studies – it is suggested to rely on computer-aided techniques for collecting and analyzing data. In addition to computationally efficient models to analyze design concepts, GSS technology and content analysis software are suggested for efficient data collection and analysis.

It is suggested to conduct the design procedure evaluation in a controlled user study environment. Reich (2010) and Tomiyama (2006) outlined a general criticism of engineering design research focusing on design theory and methodology that the scientific method is inefficient in this field because of the inherent complexity of the field. Hence, it is desirable to incorporate a thorough and rigorous experimental platform to evaluate the proposed design procedure for flexibility. This experimental approach can be done prior to more extensive case study applications at firms, so it does not preclude using the case study approach to further validate the design procedure. Since a design procedure stimulating flexible design concept generation is suggested, providing a high level of user satisfaction with the process and results, and improving anticipated quality of results (assessed subjectively) is desired. One can build upon and integrate existing methodologies to evaluate the quantity of non-redundant and good flexible design concepts generated, user satisfaction with the process and results, as well as anticipated quality of results, as part of the experimental methodology proposed here. This explicitly demonstrates the complementarity of the methodology *vis-à-vis* existing counterparts.

Chapter 3 – Research Questions and Approaches

"Science is the great antidote to the poison of enthusiasm and superstition." – Adam Smith (1723 – 1790)

This chapter explains the research questions investigated in this thesis, and the overall research approach to answer them. Three research areas are defined to answer the research questions of interest. Broader questions are formulated in each area to define the general intellectual contributions this thesis hopes to make. The more specific research questions addressed here are then formulated. It is not the goal to fully answer the broader questions, but it is useful to structure the framework necessary for further investigations on a longer timescale.

The first research area is concerned with the development of a design procedure to support flexible design concept generation, discussed in Section 3.1. The second research area introduces an evaluation methodology based on anticipated performance measurements of design concepts. Thus Section 3.2 focuses on the experimental methodology used for evaluation of a design concept generation procedure. The third research area described in Section 3.3 frames preliminary thinking to understand better the effects of different treatment conditions on discussion content during ideation sessions.

3.1 Area 1: Procedure for Flexible Design Concept Generation

Motivated by the needs identified in Chapter 2, the first research area is concerned with developing a design procedure for flexibility stimulating creativity, improving anticipated performance of design concepts, user satisfaction with process and results, and quality assessments of results as applied to a particular design problem. This broader general interest is summarized with the following question:

"What is the best way to package a design procedure for flexible design concept generation that can demonstrably stimulate creativity, improve anticipated performance, user satisfaction with process and results, and quality assessments of results when applied to a particular design problem?"

It is postulated that a design procedure composed of a short lecture on the topic of flexibility in design, complemented with a prompting ideation mechanism focusing on flexibility can help achieve these goals. This procedure is described in Chapter 4. The more specific research question addressed in this thesis is:

"Can the postulated design procedure for flexibility help increase the number of complete and good flexible design concepts generated, improve anticipated performance of flexible design concepts compared to a benchmark design, improve subjective user impressions of satisfaction with the design procedure and results, and provide improvements in terms of quality assessment of results, as compared to a baseline control design procedure?"

The notions of "complete" and "good" design concepts are necessary to measure the effects of the design procedure on the quantity of design concepts generated. A design concept is complete if it contains all the information necessary for computer and/or mathematical modeling by a third person with the goal of measuring anticipated performance. This definition is different from the concept of completeness by Kurtoglu et al. (2009), defined as "measuring how much a given concept variant addresses a sub-function depicted in the function structure." A complete design concept is good if it improves the anticipated performance compared to a benchmark design.

Both deductive and inductive approaches are suggested to answer this research question. It is hypothesized that at least one of the factors of the suggested design procedure will have a significant effect on all attributes of design concept generation activities listed above. The experimental methodology from Chapter 5 is applied to test this hypothesis and measure these effects statistically – a deductive approach. A factorial design of experiments is used where participants are assigned randomly to different treatment groups. To reinforce internal validity, each experimental session is structured using some variation of the pretest-posttest quasi-experimental non-equivalent group design (Campbell & Stanley, 1966). An inductive coding procedure is also necessary to extract design concepts from ideation session transcripts for computer modeling and evaluation of design concepts – similar to the approach suggested by Strauss and Corbin (1990).

3.2 Area 2: Performance-Based Design Procedure Evaluation

To answer the research question above, one needs a methodology to evaluate the effect of the design procedure on anticipated performance of design concepts. This is because existing design procedure evaluation methodologies cannot satisfactorily do this. Current methodologies are excellent, however, to answer questions related to creativity (i.e. by measuring quantity of complete and good design concepts), subjective user impressions of satisfaction with process and results, and quality assessments of results (i.e. by means of qualitative surveys).

To address this issue, it is proposed to develop a novel experimental methodology for design procedure evaluation based on the anticipated performance of design concepts generated. It is hoped the methodology will complement existing ones described in Section 2.3.

To be useful for engineering design research, the experimental methodology should be effective, and efficient. Relying on the definitions of The New Oxford American Dictionary (2006), the methodology is effective if it successfully produces the desired or intended result. Here, the intended result is to enable objective and quantitative measurements of the effects of a design procedure on anticipated performance of design concepts. The methodology is efficient if it can achieve maximum productivity with minimum wasted effort or expense.

The second research question of broader interest focuses solely on the ability to evaluate procedures based on anticipated performance of concepts generated. This implicitly assumes that other evaluation methodologies satisfactorily measure the effects on creativity, user satisfaction with process and results, and anticipated quality of results:

"Can one develop an experimental methodology enabling objective and quantitative evaluation of the effects of a design procedure on anticipated performance of design concepts, that is both effective and efficient?"

This thesis asserts that an experimental methodology incorporating computer-modeling techniques will enable effective and efficient measurements of anticipated performance of design concepts. Furthermore, it is postulated that relying on GSS technology will enable efficient data collection and analysis in experiments – in addition to having baseline effects on creativity. The more specific research question addressed in this thesis is:

"Is the proposed experimental methodology efficient and effective to measure objectively and quantitatively the effects on anticipated performance of design concepts generated in a collaborative design setting?"

It is proposed to test the hypothesis that the experimental methodology will be effective and efficient at measuring such performance both objectively and quantitatively by evaluating the design procedure for flexibility described in Chapter 4. The hypothesis of effectiveness is tested by showing that the methodology can be used to compare the effects of different treatments based on anticipated performance of design concepts. Testing the hypothesis of efficiency is done by recording the time and analytical resources necessary to collect and analyze experimental data, as compared qualitatively with other well-known methodologies.

As suggested by Frey and Li (2010), this empirical approach is similar conceptually to *naturalistic epistemology*, where knowledge is created by gathering evidence through sensorial experience and real experiments (Audi, 1998). A natural extension of this thesis would be to validate further the methodology by demonstrating evaluation of other design procedures.

One limitation to this research is that the hypotheses of effectiveness and efficiency can only be tested partially, because only one design procedure is evaluated. Evaluation of more design procedures using this experimental platform will further demonstrate effectiveness and efficiency, especially if many design procedures are compared to one another. Another limitation is that performance cannot be measured directly, but rather indirectly by having an intermediary person measuring anticipated performance by means of a computer model. Therefore, *ex post* studies comparing the actual performance outcomes of a given design with anticipated performance outcomes measured using the methodology can further validate effectiveness. This may require longer field studies and close collaboration with industry on research projects. It is left as an opportunity for future work.

Even if the thesis does not fully answer the research question above, it is important to pose it now to pave the way for further application of the methodology, evaluation of other design procedures, and ultimately to seek validation. This thesis is part of a larger research effort demonstrating that the experimental methodology can be used to evaluate procedures for concept generation – although complete demonstration and validation is out of the intended scope.

3.3 Area 3: Design Procedure Influence on Discussion Content

The third research area provides a preliminary basis to investigate the cognitive processes underlying participants' discussions in experiments, subject to different treatment conditions. Although this area is not central to the thesis, there is an opportunity to use data extracted from ideation transcripts to study the effects on the content of discussions between participants. The research question of broader interest is:

"What are the effects of different treatment conditions on the content of participants' discussions during collaborative design experiments?"

Given the thesis focuses on the specific issue of design for uncertainty and flexibility, it is expected that discussions will revolve more around uncertainty and flexibility related concepts after using the design procedure for flexibility. One way to measure this effect is to determine the *influence* of uncertainty and flexibility related words on the overall discussion content. Influence measures how words channel flows of meaning in a text. Some words are more meaningful than others to convey the content of an excerpt, and therefore more influential (Corman, Kuhn et al., 2002). Word influence is used here as opposed to word frequency because it incorporates notional elements related to the overall structure of the message – which cannot be done through word frequency analysis alone. The specific research question therefore is:

"What is the influence of uncertainty and flexibility related words on the content of ideation transcripts when participants are subjected to the design procedure for flexibility, as compared to a baseline control design procedure?"

The hypothesis is that at least one of the factors comprising the design procedure for flexibility will affect participants' discussions by having them focus more on uncertainty and flexibility related concepts. The effect will be measured by noticing a higher influence of uncertainty and flexibility related words in transcripts from sessions making use of the procedure for flexibility.

To test the hypothesis, it is suggested performing content analysis on experimental transcripts using specialized software to extract word influences. Centering Resonance Analysis (CRA) is suggested to calculate word influence, and to link words as a network (Corman et al., 2002). Intuitively, this technique measures the influence of a word as the ratio of the number of paths connecting a word to other words, versus the total number of connections between all words in the network. A word that is highly connected to other words has high influence. In reverse, least influential words are not well connected to other words.

CRA is grounded in centering theory, assuming that good authors and speakers typically structure utterances on conversational "centers". Centers are words and noun phrases that are the subjects and objects of utterances. They are generally entities like objects, events or persons (Gordon, Grosz et al., 1993). CRA also builds upon network theory, and produces an undirected network whose nodes represent center-related words.

The approach suggested here is inductive by nature, similar to the coding analysis addressing the research question in area 1. It makes inferences directly from the data to gain more insights into the cognitive processes driving participants' discussions, and to complement observations and conclusions from the deductive and naturalistic approaches (Broniatowski, 2010). The approach is also deductive through explicit testing of the hypothesis above.

Chapter 4 – Design Procedure for Flexibility

"We recognize understanding through a flexible performance. Understanding shows its face when people can think and act flexibly around what they know. To understand means to be able to perform flexibly" – David Perkins, Harvard University

This chapter presents the procedure for flexible design concept generation evaluated in this thesis. The main factors comprising the procedure are described in Sections 4.1 and 4.2. Section 4.3 provides the rationale for devising the design procedure as done here, and provides tips on how to best use it in practice.

The proposed design procedure consists of a lecture on the topic of flexibility in design, and a prompting mechanism, helping designers generate flexible design concepts through a series of simple questions. It is anticipated that a lecture will help designers become more aware of the effects of uncertainty on performance, and open their mind to the potential of flexibility to deal with uncertainty pro-actively. The prompting mechanism is expected to have a "triggering" effect, stronger than if industry guidelines or suggestions were used. Because design is a social creative process (Warr & O'Neill, 2005), the procedure should be usable in collective design tasks. The idea is to craft an approach simulating creativity, reducing contribution barriers during the collaborative design process, opening designers' minds to more design strategies, and stimulating synergy between participants. This design procedure is evaluated against a control procedure relying on prior training in science and engineering only (i.e. no lecture on flexibility), and a free undirected ideation mechanism (i.e. no flexibility-related prompting).

To evaluate the proposed design procedure against the control, a 2 x 2 DOE is suggested (Table 4.1). Two important factors encompass treatment conditions: an education (E) and an ideation (I) factor. These factors are independent variables potentially affecting the responses of interest (i.e. concept quantity, anticipated performance, user impression of satisfaction with process and results, quality assessments of results, uncertainty and flexibility influence). The first factor E is important because it establishes the level of training assumed regarding flexibility in design. Level –1 captures control conditions where no particular training is assumed, and participants rely on prior background in science and engineering. Level +1 provides a lecture on flexibility. The second factor *I* is important because there are many ways to generate design concepts creativity in engineering design, as demonstrated in the Literature Review from Chapter 2. The approach chosen here is captured by level +1, and consists of a prompting ideation mechanism geared towards flexibility. Level –1 represents control conditions, and leaves teams generate design concepts without particular guidance (i.e. referred to as free undirected ideation) (Santanen & de Vreede, 2004). The 2 x 2 DOE is evaluated using the experimental methodology described generically in Chapter 5, and more specifically to suit this experiment in Chapter 6.

Table 4.1: 2 x 2 DOE setup for evaluating the design procedure for flexibility.

Education Mechanism (*E*)	**Ideation Mechanism (*I*)**	
	Free undirected (–1)	**Prompting (+1)**
Prior training only (–1)	Treatment 1	Treatment 2
Lecture on flexibility (+1)	Treatment 3	Treatment 4

All treatment conditions rely on simple design procedures already in use in the engineering disciplines. Naturally, other design procedures could serve here as control and experimental conditions. The proposal does not pretend to encompass all possible design procedures for flexibility that can be crafted and evaluated. It does represent however a plausible and potentially useful approach to think about uncertainty and flexibility in design, and to compare its effects with existing approaches in design practice.

4.1 Education Factor (*E*)

Control conditions for the education factor *E* rely on prior training in science and engineering only – Treatments 1 and 2 in Table 4.1. The underlying hypothesis is that design for uncertainty and flexibility is not widespread in current engineering education and practice. No particular emphasis on flexibility in engineering systems design is expected in participants' training. It is assumed that if designers had exposure to this design approach – and thought it was valuable – they would naturally incorporate these concepts in their idea generation process to solve the design problem. If it were the case, the proposed design procedure for flexibility would most likely have little to no effect on the dependent variables of interest.

Experimental conditions rely on a short fifteen-twenty minutes lecture on the topic of flexibility in infrastructure design – a particular class of engineering systems – as in treatments 3 and 4. This choice is motivated by the desire to provide a minimal training on the topic of flexibility in design. The lecture provides basic conceptual elements related to flexibility in design, as well as real-world case example applications. The material is inspired from conceptual lessons taught at MIT in the course *Engineering Systems Analysis for Design* (de Neufville, 2010). The detailed lecture slides can be found in Appendix A. In essence, the lecture on flexibility:

- Defines and describes the sources of uncertainty affecting performance of infrastructure systems (e.g. environment, markets, technology);
- Provides the rationale why adapting flexibly to changing circumstances improves expected performance (i.e. reducing exposure to downside scenarios, capitalizing on upside opportunities);
- Justifies why these ideas should be considered early in the design cycle;
- Explains the notion of a complete idea in the context of designing for flexibility;
- Supports these concepts with real-world examples applications in the aerospace and petroleum industries showing performance improvement due to flexibility.

4.2 Ideation Factor (*I*)

Control conditions for the ideation factor *I* rely on free undirected ideation to stimulate creativity – treatments 1 and 3. This leaves designers generate concepts without particular guidance or structure. The reason for selecting this approach is provided by Osborn (1957), who suggested that face-to-face brainstorming – similar to the suggested free undirected ideation mechanism – should stimulate creativity by removing any kind of barriers to creativity, expression, and idea generation. Also, brainstorming is widely used in U.S. engineering design practice (Yang, 2007). This control condition represents well what is currently done at practicing engineering firms.

In terms of experimental conditions, a prompting ideation mechanism is chosen to stimulate creativity – treatments 2 and 4. This choice is justified by the work of Santanen et al. (2004) demonstrating that a prompting ideation mechanism can be most effective to generate valuable solutions to a given problem. Such prompting mechanism is also in line with the provocation change of perspective defined by Knoll and Horton (2010), which encompasses tens of different approaches to stimulate creativity in design and other fields. The provocation change of perspective approach challenges the underlying assumptions of the design problem. It resembles the prompting mechanism geared towards flexibility, which also challenges the benchmark design, and underlying assumptions about the design process.

The ideation mechanism (*I*) uses a set of four simple prompts to help designers generate complete flexible design concepts, inspired from processes described in Babajide et al. (2009) and Walker, Rahman et al. (2001). It provides analogies from other systems to clarify each prompt, based on lessons learned over years of research in flexibility in engineering design. Detailed prompts are provided in Appendix B. The prompts ask participants essentially to think creatively about:

1. The major sources of uncertainty affecting future performance of the system. Examples of exogenous uncertainty sources are provided (e.g. market demand, natural catastrophes), and endogenous uncertainties (e.g. technology failure rates, etc.). The prompt also asks participants to consider scenarios where things go really bad (e.g. prices drop, economic crisis), and scenarios where things go really well (e.g. demand rises suddenly);
2. The flexible strategies enabling the system to adapt if the uncertainty scenarios discussed previously occur in the design process and operations. The prompt provides examples of flexible strategies from the ROA literature (Trigeorgis, 1996) useful on shorter (i.e. operational) or longer timeframes (i.e. strategic): deferring the initial capital investment until favorable market conditions, abandoning the project to get out of bad, negative market situations, investing in research and development to support growth and future opportunities, phasing capacity deployment over time instead of deploying it all at once, altering operating scale by expanding or reducing production capacity depending on

market conditions, and/or switching production output or input depending on observed demand;

3. How to prepare, engineer, and design the system to enable the flexibilities discussed previously. In particular, this prompt asks participants to consider the engineering of the system so it can react to negative or bad scenarios (e.g. start with a smaller 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 if needed), and/or completely unexpected scenario (e.g. plan ahead for emergency procedure in case of hurricane);
4. How to manage and decide when it is appropriate to use, or exercise, the flexibilities in the system. The prompt inquires about the appropriate "decision rules", or triggering mechanism that managers can use to decide when it is appropriate to exercise the flexible strategies. Examples of decision rules are: "if demand is lower than capacity for two years, operations are shutdown for six months", or "if market price gets above a certain threshold, production capacity is expanded to level *X*".

The specific prompts used in this thesis can be modified to suit different experimental contexts, design problems, and audiences. As long as the prompts cover the above four structural elements, they can be formulated in many ways to stimulate creativity related to uncertainty and flexibility. The author modified the original prompts to suit the needs of a field study at the international robotic design competition Robocon[2] in Shanghai, China, during summer 2010. Wording in questions was generally simplified because not all participants were native English speakers (e.g. exogenous and endogenous uncertainty sources in the original prompts were referred as "outside" and "inside" the design, respectively). Specific prompts were modified to suit better the realities of robotic design as opposed to infrastructure design. For instance, example uncertainty sources were changes in contest rules, a robot falling on the side, or a potential mechanical failure. Example flexibility strategies were modified such that a general deferral strategy was described as "deferring important choice to gather more information".

---

[2] The International Design Contest Robocon official website: http://www.idc-robocon.org/idc2010/e/index.html.

The fact that the prompting mechanism can be adapted to a different context outlines the need for careful crafting of the prompts before experimental work. Nothing specific to the design problem under study must influence the formulation of these prompts.

4.3 Using the Design Procedure for Flexibility in Engineering Practice

The proposed design procedure is crafted to have real-world impact on design practice, and for wide dissemination in industry. It aims for simplicity and ease of use, even though it is built on solid grounds in cognitive science, collaboration engineering, engineering design, as well as rigorous theories in economics and finance (e.g. ROA). In order to have real impact, the design procedure should stimulate flexible design concept generation, demonstrably improve overall anticipated performance of design concepts and quality of results, and provide users with subjective impressions of satisfaction with the process and results. It is the main goal of this thesis to demonstrate these attributes. If the design procedure does not have these attributes, it is fairly unlikely that it will ever be used in practice.

This section explains first why the proposed design procedure should fit well within existing practice. It then describes the kind of training required for a "flexibility expert" – or moderator – who can guide implementation of the design procedure at a firm. The section also explains the benefits of using GSS as a complementary tool to support ideation sessions, although other tools can be used – or no support tools at all. Also, given that flexible design concepts will most likely be compared to the benchmark design, or with one another, the section explains the benefits of using ROA as described in Section 2.2.3 to value design concepts quantitatively and objectively.

4.3.1 A Complementary Design Procedure for Industry

The proposed design procedure is complementary to existing methods in design practice. This should lower anticipated barriers to entry in industry. The procedure is crafted carefully to fit within the existing design process at firms – whatever this process might be – without adding too much overhead in terms of time and resources. This should alleviate potential resistance naturally arising when a novel design procedure is introduced at a firm. The lecture is relatively short, and the combination of lecture and prompting mechanism can be applied within a couple hours at the beginning of the design cycle. Hence, designers should not feel stretched out of their comfort zone when presented with the proposed procedure. Prompting is only a variation of brainstorming, which is a creative design activity already familiar to U.S. engineering practitioners (Yang, 2007). All that prompting adds to brainstorming is to structure ideation towards flexible design concepts, with the anticipated goal of improving system performance.

4.3.2 Required Training for Application

The design procedure can be packaged so that no particular training is necessary. For example, the lecture can be packaged as a short movie, and the prompting mechanism can be given to a design team as a set of clearly defined questions. Given the prompting mechanism borrows from the structure of *generate ThinkLets* (Briggs & de Vreede, 2001; Knoll & Horton, 2010), it is conceivable that an untrained moderator could dispense it – as Kolfschoten et al. (2009) demonstrate to be feasible. A ThinkLet is “a codified packet of facilitation skills that can be applied by practitioners to achieve predictable, repeatable patterns of collaboration” (Briggs, de Vreede et al., 2003). A generate ThinkLet focuses on the creative generation of solutions. Even though ThinkLets aim to ease moderated interventions and stimulate creativity, it is clear that moderation experience would be a plus.

In order to guide the design process and give maximum results, it would be best to have an expert in design for flexibility available to assist a firm or design team go through the process. This trained practitioner could be certified through a professional course similar to the approach used by the Six Sigma initiative (iSixSigma, 2010) to train in design for manufacturing reliability. In this case however, the goal would be to design for uncertainty and flexibility. To this author’s knowledge however, no such professional training course exists today.

It appears that expertise in designing for uncertainty and flexibility requires engineers to be proficient in at least the following categories:

1. Uncertainty recognition, characterization, and modeling;
2. Flexible design concept generation, identification, and modeling;
3. Design space and decision rules exploration;
4. Implementation and management of flexibility.

Many analytical tools described in Chapter 2, Section 2.2, are relevant to fulfill these criteria. One could package such tools to form the basis of a professional development training. This endeavor is beyond the scope of this thesis, and is left as an opportunity for future research. The design procedure for flexible design concept generation suggested here is however one of the building blocks of such envisioned professional program.

4.3.3 Complementary Tools

Group Support System Technology

It is strongly recommended to make use of GSS technology when applying the design procedure for flexibility. In addition to expected benefits provided by the lecture and prompting mechanisms, one expects GSS technology to stimulate creativity further during the collaborative tasks, and to ease interventions (Bostrom & Nagasundaram, 1998). As explained in Chapter 5, GSS is an inherent component of the experimental methodology to record ideation data efficiently. It can however do more than simply help with data recording; it can significantly enhance the quality, efficiency, and outcomes from collaboration interventions.

For example, GSS technology can be used to structure ideation sessions by posting different brainstorming topics in a specific order, going through a particular meeting agenda, etc. It is used in the collaboration engineering community to stimulate creativity by lessening the effects from natural barriers to creativity during collaborative design tasks (Bostrom & Nagasundaram, 1998; Dennis & Valacich, 1993; Warr & O'Neill, 2005). GSS alleviates for instance concerns about evaluation apprehension – the fear of being judged – because it may rely on direct typing to record ideas, rather than audio/video speech recording. This way, shy participants need not feel afraid to express their thoughts, and the discussion is less at risk of being dominated by a louder, stronger personality. GSS alleviates effects from production blocking – losing an idea because someone else is talking – for similar reasons. Many participants can type simultaneously, so all ideas are recorded as soon as they emerge. Merging ThinkLets and design pattern systems with GSS technology demonstrably help untrained people to moderate ideation sessions almost as efficiently as trained professionals (Kolfschoten et al., 2009).

Real Options Based Valuation

The design procedure for flexibility should be used in concert with the Real Options valuation tools described in Section 2.2.3. The main reason is that flexibility is an abstract concept that most designers recognize as beneficial, but that is difficult to quantify and justify if it requires additional costs – which may or may not always be the case. ROA provides a set of analytical tools to assess the value of flexible design concepts objectively and quantitatively. This value can be compared to that of a benchmark design, or to rank order different flexible design alternatives, so that better alternatives can be selected. Equally important, ROA techniques enable a clear assessment of the value of flexibility, so that the value can be compared directly to the cost of acquiring the flexibility. In a world where design activities are often driven by cost minimization, ROA techniques shift the debate from an abstract space where the benefits of flexibility are difficult to quantify, to a space where enabling flexibility in design becomes a simple investment decision (i.e. acquire the flexibility if the cost is lower than the expected value improvement).

Chapter 5 – Experimental Methodology

"The strongest arguments prove nothing so long as the conclusions are not verified by experience. Experimental science is the queen of sciences and the goal of all speculations."
– Roger Bacon (c. 1220 – 1292)

This chapter describes generically the experimental methodology used to evaluate the design procedure for flexibility. The methodology extends existing design procedure evaluation methodologies described in Section 2.3 through use of computer-aided techniques to enable efficient data collection and analysis. The methodology also incorporates modeling techniques to assess objectively and quantitatively the anticipated performance of design concepts. All the techniques integrated in this approach aim at effective and efficient evaluation of a design procedure in a controlled experimental setting.

Steps 1 to 5 in Sections 5.1 to 5.5 represent the novel elements of the proposed methodology, summarized in Figure 5.1. They complement typical preliminary steps like setting up the DOE, choosing the design procedure of interest, breaking down the design procedure into treatment and control factors and levels, and selecting a pool of qualified participants, as described by Shah et al. (2000). After an initial implementation of the computer model in step 2, the design problem description can be refined if more details are needed for participants. Similarly after step 5 if results are inconclusive or not statistically significant, researchers may consider performing more experiments and collect more data for further analysis.

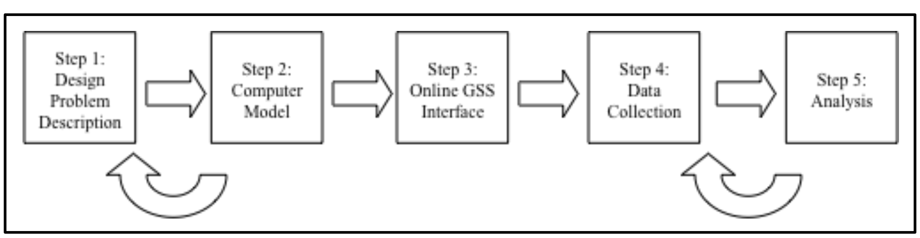


Figure 5.1: Flow chart summarizing the experimental methodology.

5.1 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 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. Although this is not an absolute requirement, there is considerable anecdotal evidence suggesting that mental imagery plays an important role in the creative process. Examples include Kekule's dream about a snake seizing its own tail, leading to the discovery of the Benzene structure (Findlay, 1948), or Watson and Crick's use of imagery to establish the helical structure of DNA (Miller, 1984). A quantitative metric is defined and explained in the problem description to assess performance of design concepts. An initial benchmark design solution is provided as starting point for ideation. The task is defined that participants should generate alternatives improving anticipated performance.

In an experimental setting, 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 creative concept generation. The problem should be chosen carefully so it is accessible to all participants. Engineering, management, and/or design background should be all that is necessary for participants to contribute positively to the discussion. If the design problem is too specialized, results may be biased in favor of participants having specialized knowledge. It may also be difficult for other participants to contribute anything useful. If the study is conducted in industry, the design problem can be a simplified or baseline version of an existing system at the company.

5.2 Step 2: Computer Model

The second step is to develop a computer model to measure the anticipated performance of design concepts objectively and quantitatively. The main characteristics of the model should be that it can 1) be developed relatively quickly (e.g. within a few weeks), and 2) run quickly (e.g. within minutes). A mid-fidelity screening model of the system fits well within this description (Lin, 2009; Steel, 2008; Wang, 2005; Y. Yang, 2009) in contrast to high-fidelity models typically developed in industry. These detailed models may take months of development, and hours and days of computational time to evaluate one or a few design alternatives.

To evaluate early design concepts, there is no need for high-level complexity and accuracy. What matters is to be able to rank-order different design alternatives objectively and quantitatively. Hence complexity reduction can be achieved through simplified versions of detailed descriptions, for instance by reducing the number of stochastic parameters in the model, and "fixing" their value. Complexity reduction can also be achieved through regression and response surface techniques, simplified development from first principles, and/or systems dynamics, giving a higher-level view of the system (de Neufville & Scholtes, 2011).

A quantitative performance metric is chosen to assess anticipated performance of design concepts objectively and quantitatively. It can measure an explicit attribute of performance (e.g. service rate emerging from a particular supply chain 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 from agent-based economic principles (e.g. Axtell, 2005), financial economic principles (e.g. de Neufville et al., 2006), Little's law and queuing theory (e.g. Berman & Larson, 2004), multi-disciplinary systems design optimization (e.g. Hassan & de Neufville, 2006), real-time computer-assisted sketching tools (e.g. iCampus MIT-Microsoft Alliance, 2010), and/or real-time CAD software (e.g. Fumarola, Seck et al., 2010; Shen, Ong et al., 2010). Performance measurements complement typical measurements in design studies (see Section 2.3.2).

5.3 Step 3: Online Group-Support System Interface

The third step sets up an online GSS interface to structure the collaborative design process, improve efficiency in recording ideation data, help in moderating ideation sessions, and stimulate creativity (Nunamaker, Briggs et al., 1997). The online GSS interface should provide 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 feature). The GSS interface should ideally not require any special facility (other than internet access) so ideation sessions can be conducted at any site. This approach is more efficient 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 design experiments anywhere, and without the burden of having all designers in the same physical location. Also as mentioned in Section 4.3.3, GSS technology has many advantages to stimulate creativity and ease collaborative interventions.

5.4 Step 4: Data Collection

The fourth step consists of structuring each experiment to enhance the signal-to-noise ratio, by collecting data before-and-after applying the design procedure under evaluation. Each experimental session is structured based on the pretest-posttest quasi-experimental non-equivalent group design suggested by Campbell and Stanley (1966). 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. 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. 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).

Borrowing conceptually from the structure of the *F*-statistics, if within-group variability is high, the variability between groups undergoing different treatments (i.e. the signal, or explained variability) must be at least as large to measure any meaningful effect. Therefore, if between-group variability is too weak, chances are it will be washed out by within-group variability, and no effect will be measured. Focusing on the differential performance of the same design groups reduces mean within-group variability, and restricts attention to improvements between different treatments to measure the best possible signal-to-noise ratio from experimental sessions. This approach adds another layer of control over an inherent creativity nuisance variable, as defined by Shah et al. (2000).

Step 4 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 within-group variability by measuring an *improvement* compared to an initial response set by each team. Therefore if a team is inherently creative, or has better knowledge than other teams on average, one can still measure a signal due to a particular treatment relative to the baseline response set by the team. For example, suppose an experiment measures the response y. The "Δ" is measured by initially determining response $y1$ from session 1 without providing any guidance, design procedure, or treatment (i.e. the baseline control procedure). A particular treatment is then applied in session 2, which leads to the response $y2$. The response that is subject to statistical analysis is $\Delta y = y2 - y1$.

This framework organizes each data collection as in Figure 5.2. That is, the design problem is described, and participants are asked in session 1 to generate design concepts that improve anticipated performance under the baseline control procedure. They are asked to vote on each design concept generated to discriminate between ostensibly contradictory concepts in the transcript analysis phase.

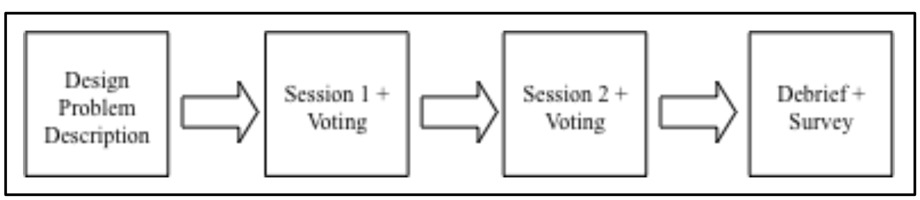


Figure 5.2: Suggested pretest-posttest experimental structure to control for inherent creativity levels, and for possible prior experience with the design procedure of interest.

In session 2, the same task is repeated, with the only difference that a treatment of choice is applied. 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 purpose of the experiment after session 2. A survey is passed to collect demographics information, as well as subjective impressions relevant to the study (e.g. satisfaction with process and results, anticipated quality of results, etc.). The survey is structured to compare impressions in sessions 1 and 2.

5.5 Step 5: Analysis

5.5.1 Coding Analysis and Response Measurements

This step consists of analyzing ideation transcript – produced at the end of each session by the GSS software – to extract complete and good design concepts to be evaluated quantitatively using the computer model. This is typically done using a coding procedure (Strauss & Corbin, 1990; Trauth & Jessup, 2000). A design concept is complete if it contains all the information necessary for computer implementation by a third party. This definition may change depending on what design procedure is evaluated. A complete design idea can be one that fulfills criteria set by the TRIZ approach, or others. A design concept is good if it improves the anticipated performance compared to a benchmark design.

By counting the number of complete and good design concepts generated in sessions 1 and 2, one can measure responses like the improvement in quantity of complete design concepts (ΔC), and the improvement in quantity of good design concepts (ΔG) generated from session 1 to session 2. Each good design concept is implemented using the computer model to extract the overall anticipated performance improvement response (ΔP).

Each response $\Delta y \geq 0$ is explained as follows. ΔC measures the improvement in the number of complete ideas. For example, if one complete idea is generated in session 1, and two *new* complete ideas are generated in session 2, $\Delta C = C2 - C1 = 3 - 1 = 2$. This measurement is in line with the typical idea quantity metric found in many design studies (Reinig et al., 2007; Shah et al., 2000; 2002). It can be used to assess the creativity level pertaining to a particular treatment. ΔG measures the improvement in the quantity of good ideas. This measurement is in line with the "good idea count" metric (Reinig et al., 2007). If a complete idea improves performance compared to the benchmark design solution (or the threshold for good quality), it is considered good. ΔP measures the improvement in anticipated performance by implementing only good ideas. This is measured in units of the quantitative metric. For example, assume a benchmark design offers anticipated financial performance of 9.3 millions (e.g. in dollars). If a good idea is generated in session 1 of \$9.5 millions ($P1$ = \$0.2 millions) and another good idea is generated is session 2 of \$10.0 millions ($P2$ = \$0.7 millions), $\Delta P = P2 - \text{P1} = 0.7 - 0.2 = \$0.5$ millions.

5.5.2 Survey Analysis

Survey responses are analyzed to measure improvements in other responses of interest Δy that do not rely on the quantitative computer model. For example, this study measures improvements in process satisfaction (ΔPS), results satisfaction (ΔRS), and anticipated result quality assessments (ΔQA) as defined in the validated survey by Briggs et al. (2006). These are the differences in user impressions of satisfaction with the process and results recorded between sessions 1 and 2, using for instance a Likert scale mechanism. Participants are also asked to assess quality of results, using a similar scoring mechanism.

Cronbach's α is a standard measure to determine inter-item reliability for the constructs used to measure the survey responses of interest (e.g. constructs ΔPS, ΔRS, and ΔQA). This metric was developed to get a sense of how consistent participants are in their answers. It helps determine whether participants understand the different constructs well enough, and whether the items used in the survey are representative of the constructs under study (Trochim, 2006). The higher the value of α – with maximum value of 1 – the more reliable the survey instrument is. The following equation can be used to calculate Cronbach's α:

$$\alpha = \frac{K}{K-1}\left(1 - \frac{\sum_{i=1}^{K}\sigma_{Y_i}^2}{\sigma_X^2}\right)$$

The variable K is the number of survey questions/items used to study a construct, σ_{Yi}^2 is the variance of survey scores obtained for item i among all participant responses, and σ_X^2 is the variance of observed total scores.

5.5.3 Content Analysis

Content analysis is performed to gain more insights about the cognitive processes involved under different treatment conditions by analyzing discussion contents explicitly. Specialized software can be used to extract word frequencies, create undirected networks of influential words, and measure the influence of each word on overall content.

This study focuses on the influence of uncertainty and flexibility related words on the content of each transcript – as explained in Section 3.3. The equation below shows how the influence I_i^T of word i on text T is calculated (Corman et al., 2002):

$$I_i^T = \frac{\sum_{j<k} g_{jk}(i) / g_{jk}}{\left[(N-1)(N-2)/2\right]}$$

The variable g_{jk} is the number of shortest paths connecting the j^{th} and k^{th} words, $g_{jk}(i)$ is the number of those paths containing word i, and N is the number of words in the network. The denominator avoids double counting.

Each transcript produces an influence score related to uncertainty and flexibility. This influence score is measured as the sum of influences of conceptually related words. The equation below exemplifies how the uncertainty influence score UI is calculated for one session of a particular experiment. It assumes a transcript text T with N words, where only uncertainty-related words with $I_i^T > 0$ are considered:

$$UI = \sum_{i=1}^{N} I_i^T$$

The difference between the influence scores for sessions 1 and 2 provides the measurements of interest for uncertainty influence (ΔUI). Similar reasoning is applied to calculate flexibility influence (ΔFI).

5.5.4 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 can be calculated using standard least-square minimization regression. β_0 approximates the 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. ε is assumed to have a normal distribution with mean 0 and variance σ^2. The null hypothesis is that H_0: $\beta_i = \beta_{ij} = 0, \forall\ i, j, j > i$.

The p-values for the main effect coefficients (β_i) and interaction effect coefficients (β_{ij}) are calculated using a non-parametric permutations test (also called randomization or exact test) (Fisher, 1935; Pitman, 1937; Welch, 1990). This approximates the probability of incorrectly rejecting the null hypothesis for each coefficient when in fact it is true (Type I error). This approach is recommended because under the pretest-posttest "Δ" framework, $\Delta y \geq 0$ for dependent variables like ΔC, ΔG, and ΔP, which truncates sample distributions about zero by excluding negative values. Therefore, it is impossible to determine whether the underlying distributions of sample data are normally distributed. Since coefficients β_i and β_{ij} are linear combinations of the underlying probability distributions (see p. 424 in Milton and Arnold, 1990), one cannot assume that coefficients are normally distributed. This may violate the normality assumption behind using the t-distribution to calculate the p-value of each coefficient, which is in fact required for most parametric significance tests. Also, it is not clear whether the underlying probability distributions from different treatments have the same shape, hence ruling out other non-parametric tests. The permutation test can also be used for other responses like ΔPS or ΔRS that do not necessarily exclude negative values, although it may be less precise than parametric tests.

The main assumptions satisfying a permutations test are that treatment groups are equivalent, and that members are sampled from the same population. There is no need to assume normality of underlying distribution functions, and similarity in shape and/or variances, because the distribution of the statistics of interest (e.g. β_i) is generated from random permutations of the original dataset. If treatments have no significant effect on observable variables under the null hypothesis, one can interchange randomly the data and assign them to different treatment groups, without regards to what treatment generated what data. One can measure the likelihood of observing the statistics of interest calculated from the original experimental dataset, as compared to all possible values of this statistics created from random permutations of the dataset. The probability of observing a given value for the statistics of interest from the original dataset provides an estimate of its p-value, or its significance level – as shown in Figure 5.3.

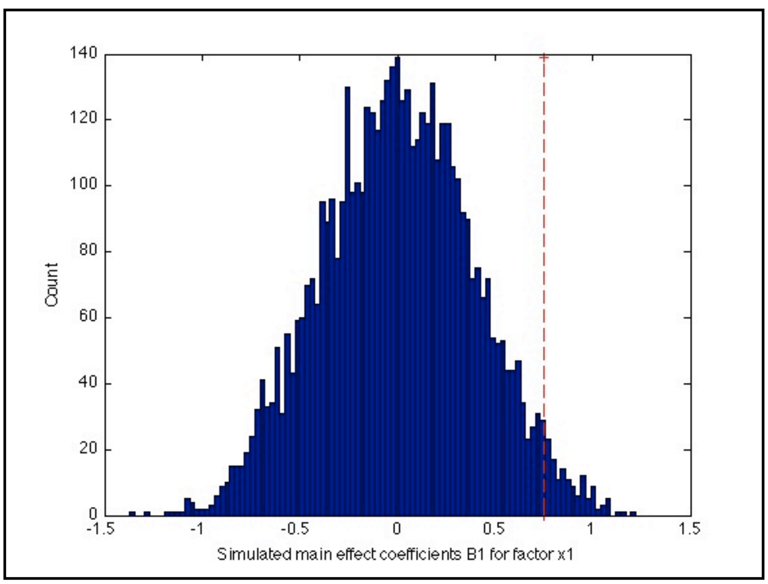


Figure 5.3: Example distribution of simulated coefficients β_1 for factor x_1 main effects, obtained from five thousand random permutations of the original dataset of a hypothetical response Δy. The location of the test statistic $\beta_1 = 0.75$ is shown as the vertical dashed line.

Figure 5.3 shows the distribution of main effect coefficient β_1 for factor x_1, obtained by randomly permuting five thousand times the original dataset for a hypothetical response Δy, and calculating the main effect coefficients through linear regression each time. The main effect coefficient value for factor x_1 calculated from the experimental dataset is $\beta_1 = 0.75$. To test for the null hypothesis that $\beta_1 = 0$, one must count the number of β_1 values falling beyond $\beta_1 = 0.75$, and below $\beta_1 = -0.75$. This corresponds to the right and left tails of the distribution. In this example, two hundred and fifty-five random β_1 values fall within this range. Hence, the two-tail p-value is $p = 0.05$, providing evidence that the null hypothesis of no effect (i.e. $\beta_1 = 0$) can be safely rejected.

Chapter 6 – Specific Experimental Implementation

"It doesn't matter how beautiful your theory is, it doesn't matter how smart you are. If it doesn't agree with experiment, it's wrong." – Richard P. Feynman (1918 – 1988)

This chapter describes the specific implementation of the experimental methodology to evaluate the design for procedure for flexibility. It describes explicitly the analysis performed in each of the five steps described generically in Chapter 5. The preliminary setup is described in Section 6.1. Sections 6.2 to 6.6 describe the specific experimental implementation.

6.1 Preliminary Setup

Table 6.1 summarizes the 2 x 2 DOE. Teams of participants were randomly assigned to one of the four treatments. As explained in Chapter 4, two mechanisms accounted for the education mechanism on flexibility (factor *E*). One factor level relies on prior education in engineering and applied sciences only, assuming no specific training on flexibility. It is denoted by level –1. The other level (+1) assumes that participants received a fifteen to twenty minutes lecture on flexibility described in Section 4.1. Also, two ideation mechanisms were evaluated for factor *I*. One is free undirected ideation (Santanen & de Vreede, 2004), where no guidance is provided (level –1). The other uses a prompting mechanism geared towards flexibility (level +1).

Seventy-one participants divided into twenty-six teams participated in these experiments. Teams of three students were formed, although a few last minute cancellations forced teams of two students. Eight teams participated in treatments 1 and 4, five teams in treatments 2 and 3.

Table 6.1: 2 x 2 DOE setup for evaluating the design procedure for flexibility.

Education Mechanism (*E*)	**Ideation Mechanism (*I*)**	
	Free undirected (–1)	**Prompting (+1)**
Prior training only (–1)	Treatment 1	Treatment 2
Lecture on flexibility (+1)	Treatment 3	Treatment 4

As Table 6.2 shows, many participants were mature graduate students with several years of work experience. Half had masters and doctoral degrees, and all participants had at least a bachelor's degree in engineering, science, and/or management. The learning objectives behind the experiment were explained in a debrief session to ensure educational benefits.

Table 6.2: Participant demographics.

Group Characteristics	Category	Percent (%)
Age	< 25	16
	25-34	69
	> 35	15
Gender	Female	16
	Male	84
Highest Education Level	Bachelor	50
	Master	49
	PhD	1
Work Experience (years)	< 5 years	39
	5-9 years	38
	> 10 years	23

6.2 Step 1: Design Problem Description

A simplified real estate development design problem was given to teams of participants at the beginning of each experiment, with slides provided in Appendix C. Computer generated images of a real case study were provided to conceptualize the design problem mentally (Figure 6.1). The moderator explained that the team is an internal consulting firm at a renowned design and development firm specializing in multi-family residential real estate. Their expertise is in developing buildings with units sold as condominium (condo) and/or apartments.

The moderator also explained the difference between the two unit types in terms of marketing, design, development plan, and engineering. For example, a condo unit may cost more to develop because it targets an upper-scale market segment, but it also sells at a higher price. An apartment building is better suited for middle-class families, workers, and students. A condo building may have fewer units than an apartment building, because each unit is more spacious. Materials used in condo units may be more luxurious, hence costing more. Common infrastructures like electrical, heating, ventilation, and water systems may be arranged differently because condo units are configured differently than apartments. Different development plans are possible, like deploying all units at once in one building, or deploying them in phases with fewer units for each phase. In this case, it is not clear whether phasing should be horizontal or vertical, how to make best use of land, etc. In both cases, there are many design and engineering issues to consider.

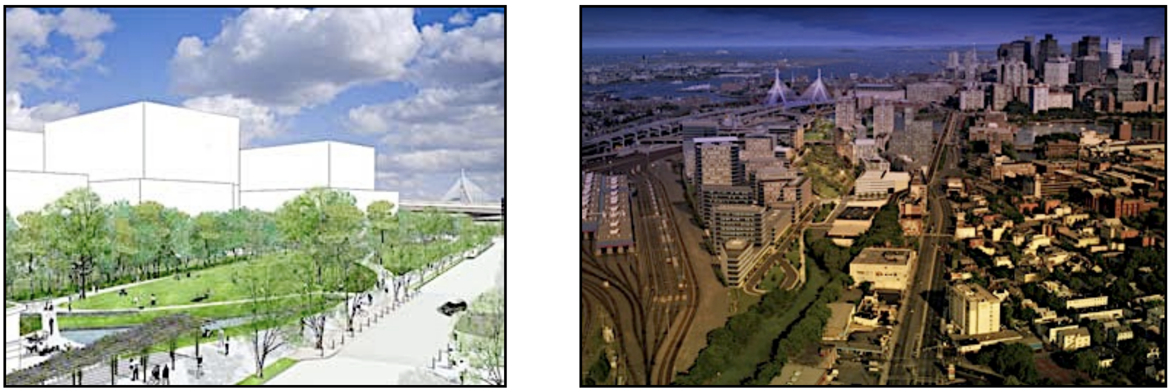


Figure 6.1: Example figures provided for mental conceptualization of the real estate development problem assigned to participants (http://www.northpointcambridge.com).

Participants were told that the firm's objective is to find a design and development plan enabling selling the building at the highest profit possible over time. This justified using NPV as the criterion for decision-making. In this context the land was already bought, and the contractual agreement was for three hundred and nine units developed and sold over a period of three years. Existing market conditions were explained for price, demand, and construction costs. Based on the NPV performance criterion, the firm was currently considering a condo-only design, with one building of three hundred and nine units capacity, developed in one block. This was the "benchmark design", which was a realistic approach to undertaking a real estate development and making best use of land.

It was explained that the consulting team was hired because management wanted a better design in terms of anticipated – or expected – performance over a range of future scenarios. The management team realized that market demand, construction costs, and prices could change. They were not convinced the benchmark design was best suited to deal with a range of fluctuating scenarios, being too inflexible. They asked the consulting team to exert their creativity freely to come up with a design and development plan they anticipated would provide better NPV performance.

6.3 Step 2: Computer-Model

ENPV – mean NPV or E[NPV] – was the metric chosen to measure anticipated performance over a range of market construction cost, demand, and price scenarios. Other metrics could have been used to suit different needs and preference utilities – e.g. standard deviation, as in the Taguchi method (1987). The equations below show how NPV and ENPV were calculated from the cash flows streams generated by different design concepts. Variable R_t represents the revenue generated at time $t = 0, 1, \ldots, T$, C_t represents the total construction and sales cost at time t, and parameter r represents the canonical Opportunity Cost of Capital (OCC), or the discount rate used to account for the time value of money. ENPV is the mean over M sampled NPV outcomes from the simulation model described below.

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

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

A DCF model was developed in Excel® to measure NPV and ENPV objectively for each flexible design alternative suggested in experiments based on cash flows, inspired from the model developed by Geltner and Cardin (2008). Figure 6.2 gives a graphical example of the DCF model. All model assumptions and equations are detailed in Appendix D.

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 6.2: DCF model measuring NPV for the real estate development design problem.

The DCF model developed incorporates assumptions about the design, development plan, engineering, and market, as stated in the design problem description. Most of the design and engineering tradeoffs are in terms of the number and timing of development phases, unit capacity in each phase, and the type of unit (i.e. condo or apartment). Such design and development decisions ultimately affect the cash flows (i.e. revenues – costs) generated and the NPV (i.e. sum of discounted cash flows), which enable discriminating between different design alternatives. For example, the design decision to select condo versus apartments affects the sales prices and construction cost, as they are both higher for condo units than apartment units. This decision has specific engineering and cost implications affecting later phases of the design process. Also, given a non-zero discount rate is used for discounting cash flows, timing and unit capacity of each phase are important from a managerial standpoint. Later cash flows generated by later deployment are more heavily discounted, and weigh less in the NPV calculation. The economic model incorporates these design and development trade-offs, and provides ways of measuring objectively the anticipated performance of different design concepts.

The ENPV of a proposed design concept is obtained by simulating stochastically a range of market demand, price, and construction cost scenarios. For each combination of scenarios, the model simulates the associated flexible strategy based on the design and decision rule implemented. One NPV is calculated automatically for each combination of stochastic scenarios, leading to one ENPV measurement from the simulation of two thousands combined scenarios. The fourth row from the top of Figure 6.2 saying "Next Phase Developed As:" shows an example of flexible adjustment to one scenario combination. This row shows the decision to switch development flexibly from condo to apartments after phase 1 if demand is higher for apartments. This flexible strategy requires developing the project in phases, and designing each unit as empty shells. The sequence of development decisions may change from one combination of stochastic scenarios to another.

Monte Carlo simulations of the inflexible and flexible DCF models lead to distributions of NPV outcomes shown in Figure 6.3 for each flexible design concept implementation. The CDFs or "target curves" show the cumulative probability of having NPV outcomes less than a certain amount. Dark target curves show outcomes for the inflexible condo-only and apartment-only designs, which are very similar and difficult to differentiate. The light curve shows outcomes from the flexible "switching" solution described above. It dominates the two other curves stochastically, showing better NPV outcomes overall than the two inflexible designs. This CDF also shows a 7% chance that the flexible design suggested will lead to negative NPV outcomes. This probability is about 18% for the inflexible condo-only and apartment-only designs.

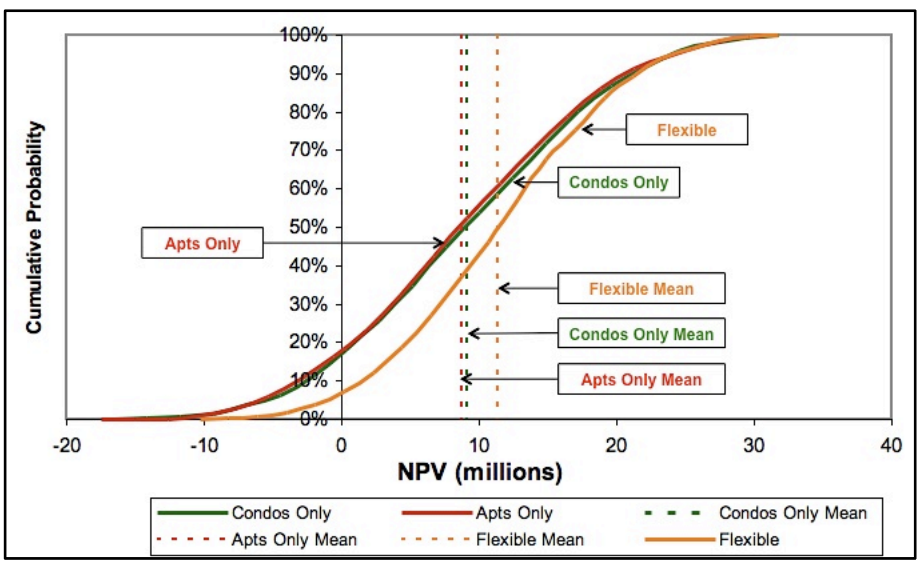


Figure 6.3: CDFs or "target curves" for the condo-only, apartment (apts)- only, and flexible designs. The curves show NPV outcomes from simulations of the DCF model in Excel®, as well as ENPV (vertical dashed lines). The dark curves for the inflexible designs are very similar, thus almost indistinguishable. The light curve represents the flexible case.

Dashed vertical lines represent the ENPV over all simulations of different design alternatives. The ENPV of the apartment-only design is approximately 8.7 millions (e.g. in dollars), 9.3 millions for the condo-only design, and 11.3 millions for the "switching" flexible design concept introduced above – showing the overall mean improvement brought by flexibility. The ENPV of the flexible design changes every time a new flexible design alternative and decision rule is analyzed. ENPV = 9.3 millions represents the expected performance of the benchmark design used in this study, as it is the highest of the two suggested inflexible design alternatives in the initial setup. It represents the quality threshold for a "good" flexible design concept.

6.4 Step 3: Online Group-Support System Interface

ThinkTank® by GroupSystems® was used as the online GSS interface (see Figure 6.4). The interface was prepared beforehand by structuring the experiment according to Figure 5.2. For example, the sequence of topics on the left shows the sequence of ideation topics, voting, and post-experimental surveys used in experiments for treatment 4. Session 1 was name coded "No CoP", while session 2 was name coded "Provocation CoP", following the lecture on flexibility. These names came from the original intent of calling the control procedure in session 1 "No Change of Perspective" (No CoP), and in session 2 "Provocation CoP" for the design procedure for flexibility – inspired by the categories developed by Knoll and Horton (2010).

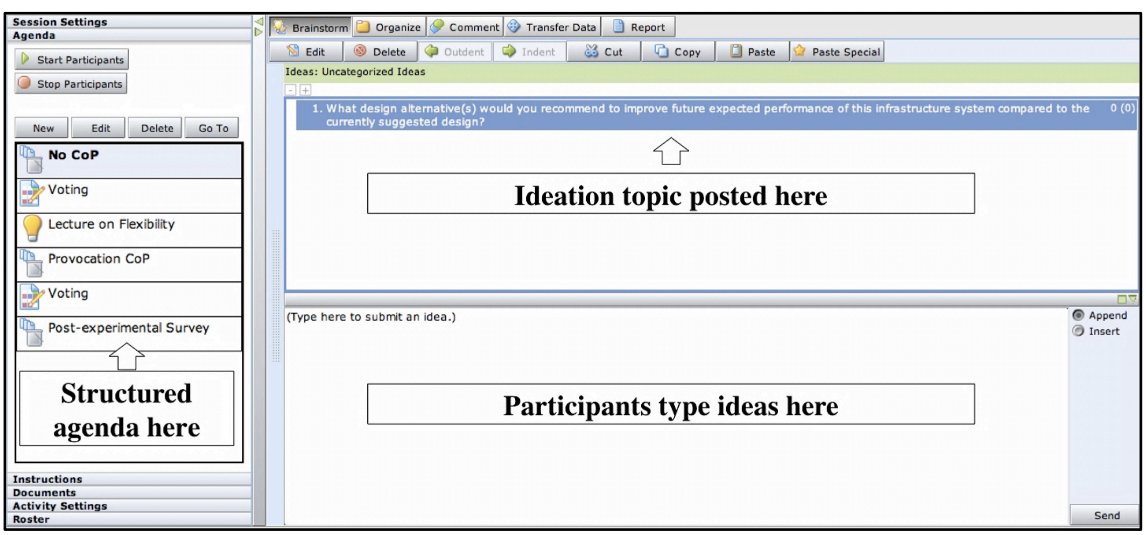


Figure 6.4: Example online GSS interface by GroupSystems® used in this study.

The online interface enabled participants to type in real-time creative design solutions to the design problem of interest in the bottom section. The interface is similar to that of chatting software: the moderator posted the ideation topic in the top section, each team member wrote an idea addressing this topic, and ideas were displayed to all team members to stimulate creativity in the top section. Each team member could reply, comment, and append new ideas to any idea in the thread. They could vote on each idea based on overall quality at the end of each session. The software program produced a transcript at the end of each session summarizing discussions and voting scores.

Using direct typing, the software efficiently recorded ideas, and succinctly accounted for all ideas at the end of each session. Data was therefore immediately ready for further analysis at the end of each experiment – without need for additional transcription and processing. The software was used anonymously and silently through typing – even though silence was not enforced in these experiments, and verbal collaboration was actually encouraged. These features alleviated both evaluation apprehension for shy participants who may have feared being judged by more imposing and loud participants, and production blocking, because many participants could type design ideas simultaneously without fear of losing an idea while someone else is talking. Given the software was available online, participants did not need to be physically collocated, or require any special facility apart from individual laptops. Some experiments were in fact conducted with participants as far as Germany, and in other U.S. cities than Cambridge (MA, United States). The platform also offered in at least three occasions the flexibility to have several teams simultaneously performing a particular treatment, which enabled efficient replication.

6.5 Step 4: Data Collection

Each experiment was structured as in Figure 5.2 to enable measurements of the "Δ" improvement response. For each experiment, the real estate design problem was described first, and a benchmark design solution was presented. Teams were then asked to recommend alternative designs they thought would improve economic performance compared to this benchmark, without mentioning the concept of flexibility. Session 1 began under this request, thus relying on prior training only and a free undirected ideation mechanism – as in treatment 1,

Table 6.1. Each team collectively suggested design solutions for twenty-five minutes, which they recorded in writing using the online GSS interface.

After session 1, each team member voted independently for five minutes on the quality of proposed solutions, using a 10-point Likert scale (1 for a bad design concept, 10 for excellent). Session 2 then began, and used one of the treatments from Table 4.1 for another twenty-five minutes. For example, treatment 1 repeated exactly the same setup in session 2 as in session 1. It was used as between-group control treatment. Treatment 2 introduced the prompting mechanism to structure ideation towards flexibility. Treatment 3 provided the lecture on flexibility, followed by free undirected ideation. Treatment 4 provided the lecture on flexibility, and used the prompting mechanism. Voting followed session 2 as well.

An online debrief session was provided to explain the purpose of the experiment. This was followed by an online survey evaluating the constructs of Process Satisfaction (*PS*), Results Satisfaction (*RS*), and Quality Assessment (*QA*) of results in each session. The purpose was to measure improvement responses between sessions 1 and 2 for ΔPS, ΔRS, and ΔQA. Five to six questions per construct were evaluated using a seven-point Likert scale. Questions were inspired from the questionnaire validated experimentally by Briggs et al. (2006). The online survey was used as well to collect demographics information shown in Table 6.2. Examples debrief and survey questions for *PS*, *RS*, and *QA* measurements in session 1 are provided in Appendix E. The same questions were used for session 2. All activities in each session were performed within ninety minutes and two hours. Resulting data consisted of written ideation transcripts describing conceptual design solutions in plain text, with voting scores, and survey data.

6.6 Step 5: Analysis

The goal was to evaluate the effects of the design procedure by measuring improvements on the responses of interest Δy between sessions 1 and 2. Coding analysis was used to extract relevant design concepts from ideation transcripts. Survey analysis was used to analyze user impressions of satisfaction with the process and results, and for quality assessments of results. Content analysis was performed to determine the effect of different treatment conditions on discussion content. The main and interaction effects on the different Δy responses were evaluated through statistical analysis.

6.6.1 Coding Analysis and Response Measurements

Two independent coders reviewed each ideation transcript to identify the relevant elements forming complete ideas, as defined below. If content was found for each element in a coherent design idea, the idea – or design concept – was recorded as complete. Many complete ideas were recorded using this approach. The average inter-rater agreement reached 95% over all ideation transcripts. Concepts retained for implementation, evaluation, and statistical analysis were the ones agreed upon by both reviewers after meeting and discussions. An example coding analysis on an original transcript is presented in Appendix F, showing how some of these elements were extracted from raw ideation data.

This analysis looked for complete design ideas enabling flexibility in engineering systems. Based on the procedures described in Babajide et al. (2009) and Walker et al. (2001), a flexible design idea was considered complete if there was:

1. A clearly identified uncertainty source affecting anticipated performance;
2. A clearly identified flexible strategy to deal with the above uncertainty in design and operations, to adapt as the uncertainty is resolved over time;
3. A clear conceptual description of how the flexibility is enabled concretely, considering engineering design, legal, management, and/or financial aspects;
4. A clear decision rule, or "triggering mechanism" based on observations of the uncertainty source, determining when it is appropriate to exercise the flexibility in operations.

One team came up with the complete flexibility idea to switch between condo and apartment during development phases – similar to the example used above to describe the computer model. Switching can be done by designing units as empty shells to be finished and sold later either as condo or apartment. This strategy captures whichever of the two markets has the highest demand. This approach contrasts with the benchmark inflexible plan where all units are developed at once as condos. It is clear the flexible approach has different design implications than the inflexible one at later phases of the design cycle. It is not clear however at the conceptual stage which design is most profitable, hence the need for explicit modeling.

This switching example flexibility idea was considered complete because it:

1. Clearly identified a source of uncertainty in market demand for either unit types;
2. Suggested switching to one unit type or another as a flexible strategy to deal with demand uncertainty;
3. Suggested a concrete design modification to enable the flexibility (i.e. the unit shell);
4. Proposed a clear decision rule to trigger the switch (i.e. when demand in any one phase is higher for one unit type, the shell is finished for this type of unit and sold).

Other suitable examples identified by participant teams – and summarized in Appendix G – were the ability to phase the development planning and deploy capacity over time only when needed (e.g. if unit construction cost is lower than unit sales price, deploy another phase), expand or reduce unit capacity in each phase whenever appropriate (e.g. expand/reduce unit capacity within a given phase if demand is higher/lower than planned capacity), temporarily abandon the project if market conditions are not suitable (e.g. if unit construction cost is higher than unit sales price), and not develop a phase at all if market conditions are too unfavorable (e.g. develop if unit sales price is above a minimum threshold).

Each complete idea generated in sessions 1 and 2 was counted and implemented using the computer model to determine the subset of good ideas belonging to each sessions, giving rise to response measurements ΔC and ΔG. Different combinations of good ideas were evaluated to determine the best performance achieved in each session compared to the benchmark design. It is important to evaluate different combinations of good ideas, as some may interact positively, and others negatively. This measures the performance improvement response ΔP in each experiment. A summary table is presented in Appendix H with all complete and good ideas identified, as well as measured performance improvements.

Table 6.3 summarizes example ΔC responses obtained for all treatment groups, with factors $x_1 = E$ and $x_2 = I \in \{-1, +1\}$. There were eight replicates for treatments 1 and 4, and five replicates for treatment 2 and 3. For example, the first team undergoing treatment 1 generated one more complete idea in session 2 that was not generated in session 1. All remaining teams for treatment 1 did not identify any new complete design concept in session 2, with the exception of team R_5. Similar data were compiled for responses ΔG and $\Delta ENPV$, as provided in Appendix I.

Table 6.3: Complete dataset of ΔC measurements for all treatment groups with $x_1 = E$ and $x_2 = I \in \{-1, +1\}$. R_k is the k^{th} replicate of a treatment. Expressed in units of complete ideas.

Treatment	E	I	R_1	R_2	R_3	R_4	R_5	R_6	R_7	R_8
1	–1	–1	1	0	0	0	1	0	0	0
2	–1	+1	2	2	1	3	1			
3	+1	–1	1	1	1	0	1			
4	+1	+1	2	3	1	4	1	2	2	3

6.6.2 Survey Analysis

Participant responses from the online survey were analyzed to evaluate the constructs ΔPS, ΔRS, and ΔQA. For each participant, a total score for PS, RS, and QA was obtained in each session by summing Likert scores for each individual construct. For example, an individual may have scored 27/35 if five questions were asked about process satisfaction in session 1 (i.e. implying a maximum of seven points for each question). The score difference between the two sessions measured ΔPS, ΔRS, and ΔQA for each participant, with examples results for ΔPS in Table 6.4. For instance under treatment 1, the first participant recorded an improvement of 8 points between sessions 1 and 2, the second participant 24 points, etc. This means the two first participants were more satisfied with the treatment in session 2 as opposed to session 1. Negative values mean participants were satisfied better with the treatment in session 1 than in session 2.

Cronbach's α values were measured to determine inter-item reliability for all constructs in sessions 1 and 2, with results in Table 6.5. Given that most values are well above the smallest value of 0.89 indicates that participants' responses were consistent across different items measuring a similar construct (Trochim, 2006). Although this does not aim to demonstrated full survey validation – already done by Briggs et al. (2006) – this suggests that the different items used to measure each construct indeed helped measure the concept reliably within and across participants (Cortina, 1993; Nunnally, 1978). A summary dataset for all transcript analyses and response measurements for ΔRS and ΔQA is provided in Appendix I.

Table 6.4: Sample dataset of eight ΔPS measurements for each treatment group with $x_1 = E$ and $x_2 = I \in \{-1, +1\}$. R_k is the k^{th} participant for a treatment. Expressed in PS points.

Treatment	E	I	R_1	R_2	R_3	R_4	R_5	R_6	R_7	R_8
1	–1	–1	8	24	0	-5	4	-6	2	2
2	–1	+1	-3	0	1	-10	8	2	5	0
3	+1	–1	-18	3	10	0	2	0	10	4
4	+1	+1	9	20	5	5	10	-6	5	15

Table 6.5: Cronbach's α measured for constructs ΔPS, ΔRS, and ΔQA in sessions 1 and 2.

Treatment	E	I	Session 1			Session 2		
			ΔPS	ΔRS	ΔQA	ΔPS	ΔRS	ΔQA
1	–1	–1	0.97	0.98	0.94	0.96	0.98	0.96
2	–1	+1	0.97	0.97	0.95	0.98	0.99	0.96
3	+1	–1	0.89	0.95	0.96	0.98	0.97	0.94
4	+1	+1	0.94	0.97	0.91	0.97	0.97	0.95

6.6.3 Content Analysis

Content analysis was performed on ideation transcripts separately for sessions 1 and 2 of all four treatments using software Crawdad®[3] to measure the influence of each word on overall content. Word influence was used to measure the responses of interest ΔUI and ΔFI.

---

[3] Available at http://www.crawdadtech.com/.

Data preparation involved saving documents in plain ASCII text format (i.e. extension .txt) for processing in Crawdad®. Each document was then reviewed for spelling mistakes – which occurred frequently as participants were typing quickly during sessions – and replacing words with similar meaning with the same transcription (e.g. replacing "apt" with "apartment"). This review ensured the same word was used to characterize the same concept, and counted properly in the influence measures. Finally, each document was reviewed to complete sentence punctuation, as the CRA algorithm developed by Corman et al. (2002) expects complete sentences – while participants may sometimes submit incomplete sentences.

A list of thirty-one uncertainty related words and twenty-three flexibility related words was extracted from ideation transcripts, as shown in Appendix J. These words were selected because they were mentioned during the introduction and/or lecture on flexibility. It is assumed that participants would most likely use similar terms in their discussion – if the treatments had any effect at all. Example uncertainty related words are cost, demand, fluctuation, uncertainty, and value. Example flexibility related words are conversion, convertibility, differentiation, flexibility, and upgrade. For each session, an influence score was measured as the sum of influences for uncertainty and flexibility related words separately. Therefore, each session produced an uncertainty influence (*UI*) score, and a flexibility influence (*FI*) score. The score difference between the two sessions measures ΔUI and ΔFI for each experiment, with example results for ΔUI presented in Table 6.6 (results for ΔFI are provided in Appendix I).

Table 6.6: Complete dataset of ΔUI measurements for all treatment groups with $x_1 = E$ and $x_2 = I \in \{-1, +1\}$. R_k is the k^{th} replicate of a treatment. Expressed in uncertainty influence points.

Treatment	E	I	R_1	R_2	R_3	R_4	R_5	R_6	R_7	R_8
1	–1	–1	0.0	0.2	-0.1	0.9	-0.1	0.3	0.3	0.2
2	–1	+1	0.0	-0.1	0.6	-0.5	0.2			
3	+1	–1	0.0	0.5	0.0	-0.2	-0.2			
4	+1	+1	0.2	0.0	0.2	-0.1	-0.3	-0.2	0.1	0.7

For instance under treatment 1, the first team recorded an improvement of 0.02 uncertainty influence points between sessions 1 and 2. This means uncertainty related words were slightly more influential in session 2 than in session 1. Negative values mean uncertainty related words were less influential in session 2 as opposed to session 1.

6.6.4 Statistical Analysis

The GLM above was modified to suit the response of interest Δy:

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

The response Δy takes values ΔC, ΔG, $\Delta ENPV$, ΔPS, ΔRS, ΔQA, ΔUI, or ΔFI. The coefficient β_0 approximates the total mean, β_E and β_I model the main effects of factors E and I respectively, and β_{EI} models the two-way interaction effect between the two factors.

The main and interaction effects were calculated using regression analysis on each dataset in Matlab®. Tests for the null hypotheses H_0: $\beta_E = \beta_I = \beta_{EI} = 0$ were performed. The p-values for each coefficient were obtained using the permutation technique described in Section 5.5.4. Example code inspired by Frey (2008) is provided in Appendix K to calculate main and interaction effects, as well as corresponding p-values.

Chapter 7 – Results

"All truths are easy to understand once they are discovered; the point is to discover them."
– Galileo Galilei (1564 – 1642)

This chapter presents results obtained to quantify the main and interaction effects of the design procedure on the ability to generate valuable flexible design concepts (ΔC, ΔG), improve anticipated performance (ΔP), improve impressions of satisfaction with the process (ΔPS) and results (ΔRS), and anticipated quality assessments of results (ΔQA). Preliminary results are also shown regarding the effects of the procedure on the influence of uncertainty related words (ΔUI), and flexibility related words (ΔFI).

The more specific hypotheses are tested that the lecture on flexibility and/or prompting ideation mechanism focusing on flexibility will have significant main effects on responses of interest Δy (i.e. the null hypothesis is of no main and interaction effects of the factors). Sections 7.1 to 7.8 below present in turn the results for each of evaluation metric.

7.1 Improvement in Complete Ideas (ΔC)

Table 7.1 shows the mean ΔC improvements brought by the design procedure under all four treatment conditions. For example, prior training only ($E = -1$) and free undirected ideation ($I = -1$) led to a mean ΔC improvement of 0.25 complete ideas from session 1 to session 2. The marginal means on the rightmost column represent the mean values for a given factor level, confounding the means of the two other factor levels. For instance, the marginal mean ΔC improvement when factor E is set at level –1 is 0.85 complete ideas. This is the average obtained along all treatment conditions involving $E = -1$, including both conditions $I \in \{-1, +1\}$. Similarly, the marginal mean for factor $E = +1$ is 1.69. The total mean for all experimental data is shown on the bottom right corner of the table (1.27 complete ideas).

There are three ways to study the main and interaction effects of the factors. First, one can observe whether the response of interest changes significantly when a factor level is toggled from level –1 to +1, relative to the total mean. For example, given the marginal mean ΔC

response changes from 0.85 to 1.69 when toggling factor *E* between levels –1 and +1 may suggest that factor *E* has a significant main effect because it differs significantly on each side of the total mean 1.27. A regression analysis enables one to confirm this by calculating the main effect explicitly, and the associated *p*-value – as explained below.

Table 7.1: Mean values for Δ*C* for all four treatments, including marginal means for each factor, and the total mean on the bottom-right corner. Expressed in units of complete ideas.

Education Mechanism (*E*)	**Ideation Mechanism (*I*)**		**Marginal means (*E*)**
	Free undirected (–1)	**Prompting (+1)**	
Prior training only (–1)	0.25	1.80	0.85
Lecture on flexibility (+1)	0.80	2.25	1.69
Marginal means (*I*)	0.46	2.08	1.27

The second approach is to plot each mean value as done in Figure 7.1. Since there are two independent variables (*E* and *I*), the response surface should be seen in 3-dimensions. Thus, the representation in Figure 7.1 is often used to compress this information to a 2-dimensional graph. The lower light curve connects the two mean values corresponding to treatment conditions when $E = -1$ (i.e. $\Delta C = 0.25$ when $I = -1$, and $\Delta C = 1.80$ when $I = +1$). The upper dark curve connects the two mean values when $E = +1$ (i.e. $\Delta C = 0.80$ when $I = -1$, and $\Delta C = 2.25$).

If both factors had no effect on the response Δ*C*, one would most likely observe two horizontal curves aligned with one another, at about the mean Δ*C* corresponding to the total mean. If *E* only had an effect, one would notice two horizontal parallel curves at different Δ*C* values quite apart from the total mean. If *I* only had an effect, one would notice two curves superimposed and diagonal (showing a change between levels –1 and +1 for factor *I* only, not for factor *E*). Finally, an interaction effect would occur if the two curves did not display consistent behaviors as different combinations of factor levels occurred. For instance, if the curves were crossing, one would expect an interaction effect between the two factors. This interpretation would lead to the same conclusions if the curves connected the mean values when $I = -1$ and $I = +1$ respectively, instead of connecting the mean values when $E = -1$ and $E = +1$ as done here.

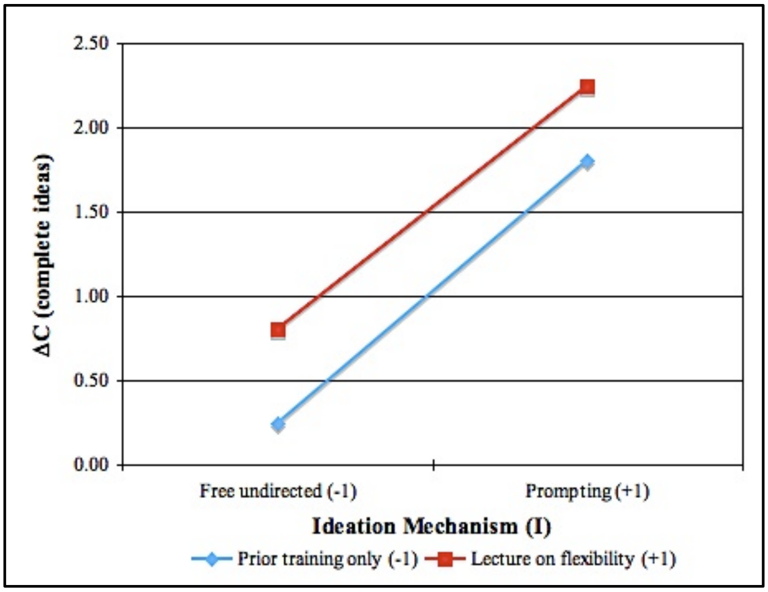


Figure 7.1: Mean plots for ΔC, for all treatments. The lower light curve depicts the means with prior training only (factor E, level –1). The upper dark curve depicts the means when a lecture on flexibility is provided (factor E, level +1).

The third approach to evaluate main and interaction effects is to replicate the mean ΔC response using linear regression, based on the GLM equation presented in Section 5.5.4. In this case, a main effect is quantified by the numerical value of the coefficients β_i, while an interaction effect is quantified by the value of β_{ij}. If those values depart significantly from the approximation of the total mean β_0, one can expect significant main and/or interaction effects. The p-value associated to each coefficient further confirms this finding, as measured using the randomization technique in Section 5.5.4. From the linear regression outputs, the mean values were modeled using the GLM equation below:

$$\Delta C = 1.28 + 0.25E + 0.75I - 0.03EI$$

The following example illustrates how the regression equation replicates the mean ΔC responses obtained under the four different treatment conditions. For treatment condition 1, $E = -1$ and $I =$

–1, which gives the experimental mean value $\Delta C = 0.25$ complete ideas as noted in Table 7.1. Using these inputs for E and I in the equation above replicates the experimental response:

$$\Delta C = 1.28 + 0.25(-1) + 0.75(-1) - 0.03(-1)(-1) = 0.25$$

Based on these three interpretations, experimental results showed that only the ideation mechanism (I) produced a significant main effect on the number of new complete ideas generated after session 1 (ΔC). This is seen first by the fact the marginal means for factor I change more significantly around the total mean than for factor E. As seen on Figure 7.1, ΔC values increase much more from left to right (toggling the ideation factor I from –1 to +1) than from bottom to top (toggling the education factor E from –1 to +1). Also, given both curves are parallel shows no significant interaction effect. Calculating explicitly the main and interaction effects using regression – as done in the equation above – led to similar conclusions. Only the main effect of I was significant ($\beta_I = 0.75$, $p = 0.00$), while other main and interaction effects were not.

7.2 Improvement in Good Ideas (ΔG)

Table 7.2 shows the mean values for ΔG with graphical representation in Figure 7.2. Similar conclusions are drawn as for ΔC. The main effect of I was significant ($\beta_I = 0.61$, $p = 0.00$), while other main and interaction effects were not. The mean values are typically lower than for ΔC. This is because a fraction of complete design concepts did not improve the ENPV compared to the benchmark. This explains why the main effect of I was not as strong here. Eleven percent of complete concepts generated under treatment 2 could not be considered as good, while seventeen percent were rejected under treatment 4. Some concepts were rejected because the decision rule was not appropriate; the team may have felt rushed in the process and did not think thoroughly about the proposal. The response for good flexible design concepts was modeled as:

$$\Delta G = 1.13 + 0.21E + 0.61I - 0.07EI$$

Table 7.2: Mean values for ΔG for all four treatments, including marginal means for each factor, and the total mean on the bottom-right corner. Expressed in units of good ideas.

Education Mechanism (*E*)	Ideation Mechanism (*I*)		Marginal means (*E*)
	Free undirected (–1)	Prompting (+1)	
Prior training only (–1)	0.25	1.60	0.77
Lecture on flexibility (+1)	0.80	1.88	1.46
Marginal means (*I*)	0.46	1.77	1.12

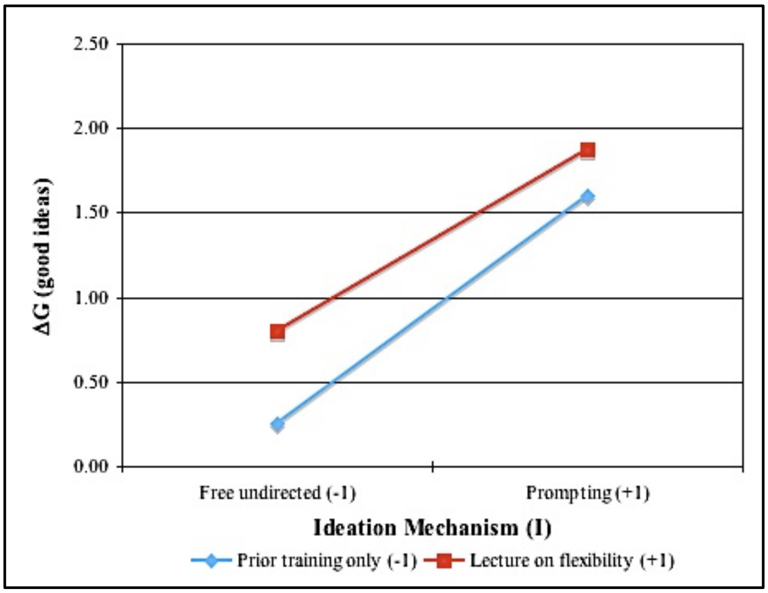


Figure 7.2: Mean plots for ΔG, for all four treatments. The lower light curve depicts the means with prior training only (factor *E*, level –1). The upper dark curve depicts the means when a lecture on flexibility is provided (factor *E*, level +1).

7.3 Improvement in ENPV ($\Delta ENPV$)

Table 7.3 shows the mean values for anticipate performance improvements $\Delta ENPV$ with graphical representation in Figure 7.3. In this case the main effect of *E* was significant (β_E = 0.72, p = 0.06), as well as the main effect of *I* (β_I = 0.74, p = 0.06). These effects clearly show that providing a lecture on flexibility in addition to using a prompting mechanism improved the overall anticipated performance of design concepts. From the linear regression outputs, the GLM equation was rewritten as:

$$\Delta ENPV = 2.08 + 0.72E + 0.74I - 0.20EI$$

Table 7.3: Mean values for $\Delta ENPV$ for all four treatments, including marginal means for each factor, and the total mean on the bottom-right corner. Expressed in millions.

Education Mechanism (E)	Ideation Mechanism (I)		Marginal means (E)
	Free undirected (–1)	Prompting (+1)	
Prior training only (–1)	0.41	2.30	1.14
Lecture on flexibility (+1)	2.26	3.34	2.92
Marginal means (I)	1.12	2.94	2.03

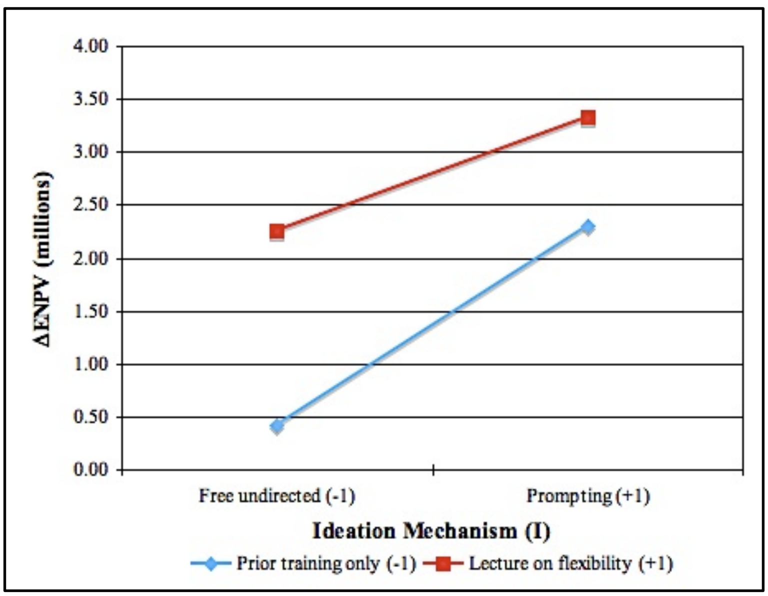


Figure 7.3: Mean plots for $\Delta ENPV$, for all treatments. The lower light curve depicts the means with prior training only (factor E, level –1). The upper dark curve depicts the means when a lecture on flexibility is provided (factor E, level +1).

7.4 Improvement in Process Satisfaction (ΔPS)

The mean value results for ΔPS are shown in Table 7.4, with mean plots in Figure 7.4. There was a significant main effect for the education factor (β_E = 1.45, p = 0.08), and a considerable interaction effect (β_{EI} = 1.36, p = 0.11). The GLM response obtained was:

$$\Delta PS = 2.41 + 1.45E + 0.19I + 1.36EI$$

Table 7.4: Mean values for Δ*PS* for all four treatments, including marginal means for each factor, and the total mean on the bottom-right corner. Expressed in points of satisfaction.

Education Mechanism (*E*)	Ideation Mechanism (*I*)		Marginal means (*E*)
	Free undirected (–1)	Prompting (+1)	
Prior training only (–1)	2.14	-0.21	1.22
Lecture on flexibility (+1)	2.31	5.41	4.26
Marginal means (*I*)	2.20	3.22	2.72

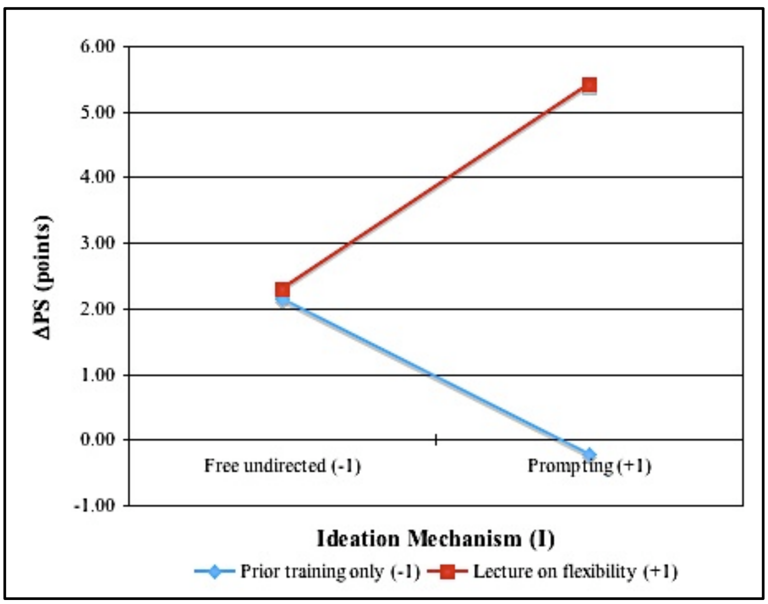


Figure 7.4: Mean plots for Δ*PS*, for all four treatments. The lower light curve depicts the means with prior training only (factor *E*, level –1). The upper dark curve depicts the means when a lecture on flexibility is provided (factor *E*, level +1).

This means on average that participants who received the lecture on flexibility were significantly more satisfied with the process than participants who did not receive the lecture. Also, results show that using the prompting mechanism without the lecture led to lower Δ*PS*. It seems reasonable cognitively to think that someone following a prompting mechanism without knowing the purpose will not be as satisfied as someone who does.

One notices that the mean value for participants using prompting jumps from –0.21 satisfaction points without the lecture (treatment 2, factor levels $E = -1$, $I = +1$) to 5.41 points with the lecture (treatment 4, factor levels $E = +1$, $I = +1$). Also, participants not using the prompting mechanism showed the same level of satisfaction improvement with and without the lecture (see treatments 1 and 3, factor levels $E = -1$, $I = -1$ and $E = +1$, $I = -1$ respectively on Figure 7.4). These two observations reinforce the idea that when using the prompting mechanism, participants were more satisfied when they learned about the concepts of flexibility than when they did not.

7.5 Improvement in Results Satisfaction (Δ*RS*)

Mean values for satisfaction with results are shown in Table 7.5 and Figure 7.5. The only significant main effect was from the education factor E ($\beta_E = 2.68$, $p = 0.00$). Participants were much more satisfied with the results when provided with the lecture on flexibility. This observation concurs with the fact that ΔPS was also enhanced when the lecture was provided. The GLM response is:

$$\Delta RS = 4.74 + 2.68E - 0.73I + 0.54EI$$

Table 7.5: Mean values for Δ*RS* for all four treatments, including marginal means for each factor, and the total mean on the bottom-right corner. Expressed in points of satisfaction.

Education Mechanism (*E*)	Ideation Mechanism (*I*)		Marginal means (*E*)
	Free undirected (–1)	Prompting (+1)	
Prior training only (–1)	3.32	0.79	2.33
Lecture on flexibility (+1)	7.62	7.23	7.37
Marginal means (*I*)	4.91	4.72	4.82

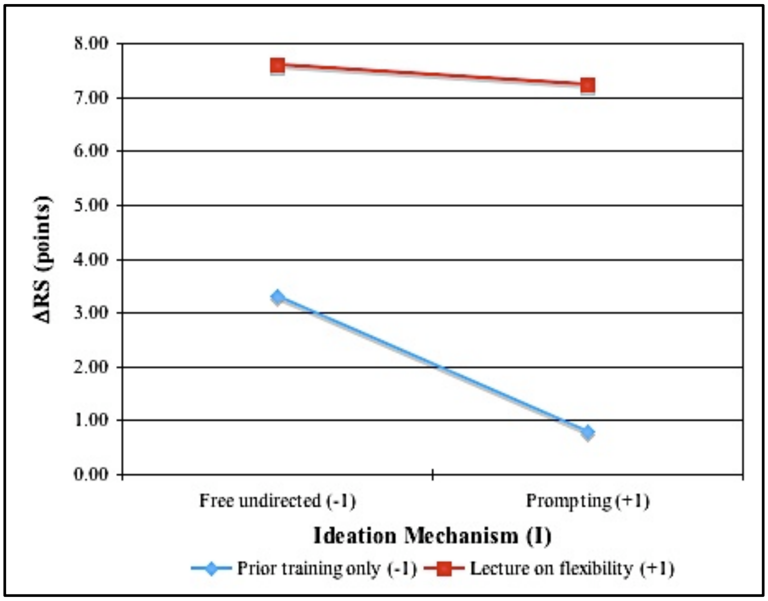


Figure 7.5: Mean plots for ΔRS, for all four treatments. The lower light curve depicts the means with prior training only (factor E, level –1). The upper dark curve depicts the means when a lecture on flexibility is provided (factor E, level +1).

7.6 Improvement in Quality Assessment (ΔQA)

Table 7.6 and Figure 7.6 show results for quality assessments of results. The education factor E had a significant main effect ($\beta_E = 3.42$, $p = 0.00$). Again, the lecture on flexibility made a considerable difference. Participants anticipated better quality of results when they received such training. The GLM response is:

$$\Delta QA = 5.80 + 3.42E - 0.22I + 0.45EI$$

The ideation mechanism (I), although creating a significant main effect in $\Delta ENPV$, did not improve assessments of results quality as judged subjectively by participants. This is an example where a particular design procedure would be discarded if evaluated solely based on a subjective basis, because useless in generating anticipated quality improvements. This approach would not lead to correct results, and conclusions. The fact that measuring anticipated performance of design concepts objectively and quantitatively yielded significant, measurable value improvements shows that it is worth developing more rigorous and thorough methodologies to evaluate design procedures, based on objective and quantitative performance-based arguments.

Table 7.6: Mean values for ΔQA for all four treatments, including marginal means for each factor, and the total mean on the bottom-right corner. Expressed in quality points.

Education Mechanism (E)	Ideation Mechanism (I)		Marginal means (E)
	Free undirected (–1)	Prompting (+1)	
Prior training only (–1)	3.05	1.71	2.53
Lecture on flexibility (+1)	9.00	9.45	9.29
Marginal means (I)	5.26	6.44	5.86

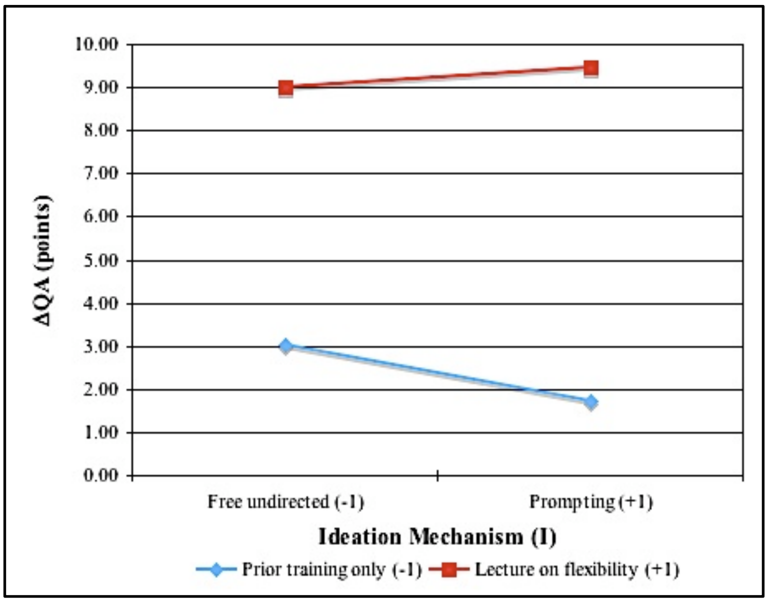


Figure 7.6: Mean plots for ΔQA, for all four treatments. The lower light curve depicts the means with prior training only (factor E, level –1). The upper dark curve depicts the means when a lecture on flexibility is provided (factor E, level +1).

7.7 Improvement in Uncertainty Influence (ΔUI)

Table 7.7 and Figure 7.7 show preliminary results for improvement in uncertainty influence responses ΔUI. None of the factors E and I had a significant effect on the mean responses. The GLM equation is:

$$\Delta UI = 0.09 - 0.04E - 0.02I + 0.06EI$$

Table 7.7: Mean values for Δ*UI* for all four treatments, including marginal means for each factor, and the total mean on the bottom-right corner. Expressed in influence points.

Education Mechanism (*E*)	**Ideation Mechanism (*I*)**		**Marginal means (*E*)**
	Free undirected (–1)	**Prompting (+1)**	
Prior training only (–1)	0.20	0.05	0.14
Lecture on flexibility (+1)	0.01	0.09	0.06
Marginal means (*I*)	0.13	0.07	0.10

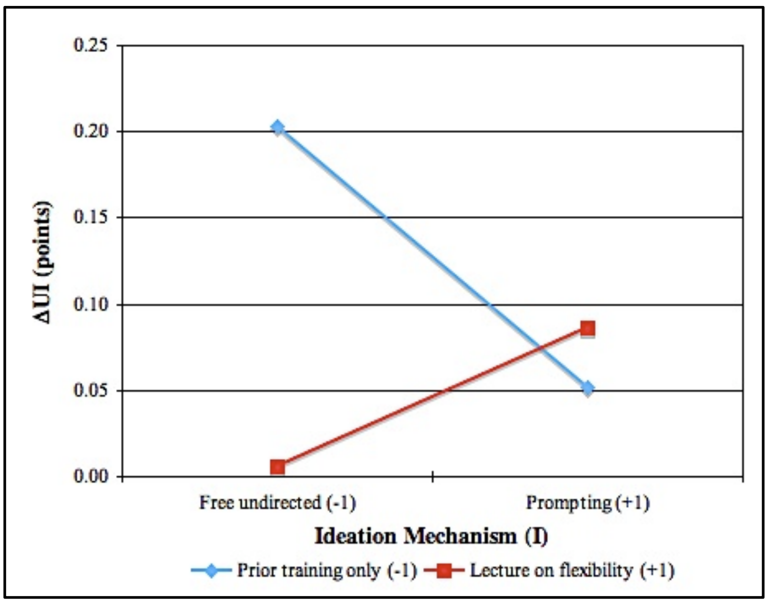


Figure 7.7: Mean plots for Δ*UI*, for all four treatments. The light curve depicts the means with prior training only (factor *E*, level –1). The dark curve depicts the means when a lecture on flexibility is provided (factor *E*, level +1).

The preliminary results imply that uncertainty related words did not have more influence after using the proposed procedure for flexibility in session 2 as compared to session 1. This result is surprising, as one would expect uncertainty related words to have more influence on discussion content once the design procedure for flexibility was applied. Explanations are provided in

Section 8.3 as to why these observations may arise, despite being counter to the hypothesis formulated in Section 3.3.

7.8 Improvement in Flexibility Influence (ΔFI)

Table 7.8 and Figure 7.8 show preliminary results for flexibility influence response ΔFI. The education factor E had a significant main effect ($\beta_E = 0.09$, $p = 0.01$), as well as the ideation factor I ($\beta_I = -0.06$, $p = 0.08$). There was also a significant interaction effect between the two factors ($\beta_{EI} = -0.06$, $p = 0.09$).

Table 7.8: Mean values for ΔFI for all four treatments, including marginal means for each factor, and the total mean on the bottom-right corner. Expressed in influence points.

Education Mechanism (*E*)	**Ideation Mechanism (*I*)**		**Marginal means (*E*)**
	Free undirected (–1)	**Prompting (+1)**	
Prior training only (–1)	-0.01	-0.01	-0.01
Lecture on flexibility (+1)	0.30	0.05	0.15
Marginal means (*I*)	0.11	0.03	0.07

The GLM response obtained is:

$$\Delta FI = 0.08 + 0.09E - 0.06I - 0.06EI$$

These results imply that flexibility related words had more influence in session 2 after the lecture was provided as compared to session 1. In contrast, the flexibility influence response decreased when prompting was used. The interaction shows that the effect of the lecture was decreased significantly when the prompting mechanism was used. This may be because the prompting mechanism diluted discussions over at least four topics – the main criteria forming a complete idea – which may not all rely directly on flexibility related words. For instance under the prompting mechanism, participants discussed in turn uncertainty sources, flexibility strategies, flexibility enabler in design, management, and finance, as well as decision rules.

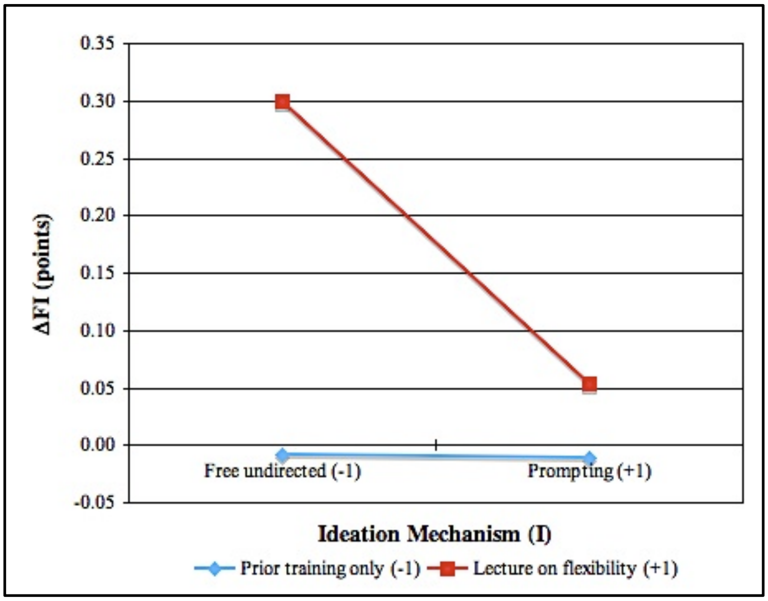


Figure 7.8: Mean plots for ΔFI, for all four treatments. The lower light curve depicts the means with prior training only (factor E, level –1). The upper dark curve depicts the means when a lecture on flexibility is provided (factor E, level +1).

In summary, this chapter demonstrated that both the lecture and prompting mechanism significantly improved anticipated lifecycle performance compared to the benchmark design. The prompting mechanism only significantly improved the generation of good flexible design concepts, while lecturing only improved user satisfaction with the process and results, as well as result quality assessments. Even though prompting demonstrably improved anticipated performance and flexible design concept generation, it did not have an effect on participants' satisfaction with the process and results. Also, prompting did not lead participants to expect better result quality. It is only when combined with the lecture that prompting improved user satisfaction. A preliminary analysis suggested that the design procedure had no effect on the influence of uncertainty-related words. The lecture had a main effect on the influence of flexibility-related words. The prompting had a main effect as well, although it diminished the influence of flexibility-related words – contrary to the stated hypothesis. The following chapter provides a more detailed interpretation of these results.

Chapter 8 – Findings and Discussion

“The search for truth is more precious than its possession.” – Albert Einstein (1879 – 1955)

This chapter summarizes and discusses the main findings obtained from the design procedure evaluation in Chapter 7. Sections 8.1 to 8.3 reiterate the research questions driving the overall approach from the research areas described in Chapter 3, and determine whether these questions are answered satisfactorily. Results limitations and validity are discussed in Section 8.4.

8.1 Area 1: Procedure for Flexible Design Concept Generation

“Can the postulated design procedure for flexibility help increase the number of complete and good flexible design concepts generated, improve anticipated performance of flexible design concepts compared to a benchmark design, improve subjective user impressions of satisfaction with the design procedure and results, and provide improvements in terms of quality assessment of results, as compared to a baseline control design procedure?”

Results in Chapter 7 indicate that the postulated design procedure for flexible design concept generation comprising a short lecture and prompting ideation mechanism focusing on flexibility had significant main effects on participants’ ability to generate complete and good flexible design concepts, and to improve overall anticipated performance compared to the benchmark design. This procedure also improved user satisfaction with process and results, and anticipated quality of results. The following paragraphs explain what element(s) of the design procedure were most influential in producing these results.

Results above demonstrate that the suggested design procedure for flexibility increased the number of complete and good ideas generated after session 1. Prompting was the main contributing factor to the creative generation of valuable flexible design concepts. Providing education on a conceptual topic like flexibility helped improve the number of complete and good flexible design concepts generated, but not in a significant manner. There are two possible explanations for this. The first explanation relies on the Cognitive Network of Model of creativity by Santanen et al. (2004). When attending the lecture on flexibility, participants stored concepts in long-term memory, which had to be activated by the working memory to create new ideas. Because cognitive resources were still devoted to storage shortly after the lecture, education might in effect have hindered the creative process. The prompting mechanism on the other hand stimulated frame activation from long-term memory to short-term memory, which may be why it was more effective at generating more flexible design concepts.

The second explanation is that participants were learning a new skill during experiments. If presented with the lecture, the skills may not have been as sharp to enable participants structuring the thought process well enough to generate new flexible design concepts. The prompting mechanism on the other hand helped scaffold the thought process, and guided participants throughout. In this view, it seems natural that a prompting mechanism helped generate more valuable flexible design concepts than simply lecturing.

This conclusion does not imply however that only this particular ideation mechanism could have been productive. Other ones could have been suggested and possibly evaluated using the experimental methodology. For example, future studies could compare the outcomes generated by different design procedures mentioned in the Literature Review from Chapter 2 to generate flexible design concepts, based on Industry Guidelines and ROA canonical strategies.

The design procedure exploited the concept of flexibility to improve overall anticipated performance of design concepts significantly. Both the lecture on flexibility and prompting mechanism contributed towards this finding. From a cognitive standpoint, this can be explained by the fact that even though prompting was useful to generate more good flexible design concepts, the lecture helped enhance the average anticipated performance of those design concepts. Providing designers with general strategies of flexibility, past case studies, and the reasons why flexibility helps generate more value contributed towards producing better quality results. Scaffolding the thought process through prompting also helped participants structure their thoughts appropriately to improve anticipated performance.

The demonstration that flexibility can improve anticipated performance and expected value is not new. Value improvements demonstrated in this study confirm the ideas put forward in the ROA literature (Dixit & Pindyck, 1994; Myers, 1977; Trigeorgis, 1996). Teams who used the procedure for flexibility generated an average improvement of thirty-six percent compared to the benchmark design (i.e. 3.34 millions/9.30 millions). Improvements ranging between ten and thirty percent compared to initial designs were shown in a significant number of case studies.[1]

In general the sources of value for flexibility come from reducing the negative impacts from downside scenarios (e.g. price or demand lower than expected), while enabling contingencies to capture upside opportunities (e.g. more copper reserves than anticipated). Given that the world is froth with uncertainty – market prices fluctuate, customer demand and preferences vary, quantities of available resources are unknown (e.g. oil, ore), regulations change, and technology inevitably evolves – anticipated performance of design concepts can only be known probabilistically. Flexibility acts explicitly on such distribution by selecting designs that improve the performance of the worst possible scenario, while aiming at extending the value of the best possible scenario. The net effect is to improve expected value and performance. Flexibility must however be considered early in the design cycle because at the detailed design phase, many decisions are made, and some of the design components might be locked. Also, it may be more expensive to modify designs in later design phases.

Providing education on the topic of flexibility played a favorable role in the subjective impressions that participants had about the procedure, the results, and quality of results. This finding confirms explicitly an intuition often outlined in GSS experimental research – off the record – that if participants are not told what an experiment or procedure is about, they will not necessarily appreciate the intervention. This is especially true when interventions are supported by technology they may sometimes find cumbersome. Similarly here, participants did not appreciate as much being steered in a particular conceptual direction during the design process if they did not know the purpose and potential benefits of doing so. Also, results show that participants who received the lecture anticipated better result quality than those who did not. This may be because they felt more committed, or believed more in the quality of results once exposed to the ideas of flexibility in design.

It is interesting to outline that even if the prompting mechanism had main effects on flexible design concept generation and anticipated performance, it had no significant effect on user satisfaction with the process and results, and quality assessment of results. This observation is reinforced by the interaction effect in Figure 7.4 showing that prompting alone decreased satisfaction significantly without the lecture on flexibility as compared to with the lecture. In line with the thought mentioned previously, this effect may occur because participants were acquiring a new skill, and did not necessarily see the benefit immediately in terms of process and results satisfaction, and quality of results. This may be as well because participants could not measure in real-time the NPV impact of their ideas (i.e. all concepts were evaluated *after* ideation sessions). This aspect may be an interesting avenue for future research.

This observation shows that relying solely on subjective user impressions may not necessarily highlight all the potential contributions that a design procedure can bring to design activities, in particular to stimulate creativity and improve anticipated performance of design concepts. Similarly relying only on objective quantitative measurements may not be enough; a procedure may very well improve performance, but not satisfy users in terms of process and results. There is a need to evaluate design procedures using both qualitative subjective measurements, as well as objective quantitative ones. This conclusion supports the need outlined by Frey and Dym (2006), Reich (2010), and Tomiyama et al. (2009) to develop methodologies enabling rigorous and thorough evaluation of design procedures based on objective and quantitative metrics, both in controlled laboratory and real-world settings.

8.2 Area 2: Performance-Based Design Procedure Evaluation

"Is the proposed experimental methodology efficient and effective to measure objectively and quantitatively the effects on anticipated performance of design concepts generated in a collaborative design setting?"

Application of the experimental methodology to evaluate the design procedure for flexibility demonstrated by example application that it could be used to measure objectively and quantitatively the effects on anticipated performance of design concepts. This property was used to evaluate the design procedure for flexibility compared to control conditions. Whether the methodology can be used to evaluate *any* design procedure based on anticipated performance remains to be shown, but the demonstration suggests such endeavor is feasible. The methodology described in Chapter 5 is general enough to evaluate other design procedures, and there is nothing specific to designing for flexibility in the overall approach.

The experimental methodology was effective in this case because it enabled discrimination between different treatment conditions based on anticipated performance. This does not, however, validate absolutely the effectiveness of the methodology, but suggests it is indeed effective. More studies and evaluation of other design procedures should further validate this point. These studies could ideally be complemented by *ex post* assessments comparing the anticipated performance outcomes with actual outcomes in industry.

The methodology enabled efficient evaluation of the procedure for flexibility, based on three observations. First, it took about a month to set up the experiments, recruit participants, and collect data. This was contingent on having participants available, which a university setting favors greatly. Given that participants – who are mostly graduate students – had, however, other priorities than contributing to experimental research, this still shows that the data gathering process was done efficiently. This compares favorably to typical case studies requiring extended fieldwork over several months at a company or organization. This is in general an advantage of controlled user studies as compared to individual case studies (Summers et al., 2009). Second, developing the case study and computer model was done over a period of two weeks. This was also efficient compared to other research where the modeling part is more time consuming, and may require months. Third, coding analysis took about an hour per ideation session transcript to process per reviewer, while implementing flexible design concepts took about another hour using the Excel® simulation model. The analysis thus took a total of about seventy-five hours, or two weeks of dedicated work. Content analysis with Crawdad® took an additional five hours of data processing and analysis of ideation transcripts. This computer-aided analysis improved efficiency dramatically compared to the intensive word for word transcription required prior to coding and content analyses, if audio/video material is used.

One may argue that the design procedure for flexibility itself was efficient. The time taken to bring significant quantitative performance improvement was extremely short. For teams participating in treatment 4 for instance, only ninety minutes to two hours were necessary to bring about average performance improvements of thirty-six percent compared to the benchmark design. For large-scale complex systems requiring investments in the order of millions and billions, this amount can be quite significant.

8.3 Area 3: Design Procedure Influence on Discussion Content

"What is the influence of uncertainty and flexibility related words on the content of ideation transcripts when participants are subjected to the design procedure for flexibility, as compared to a baseline control design procedure?"

As observed in section 7.7, preliminary results showed that factors *E* and *I* had no significant effect on uncertainty influence scores. This result is counter to the hypothesis formulated in Section 3.3, suggesting that uncertainty related words should have more influence on the content when the design procedure for flexibility is used.

There are two possible explanations for this. First, it is possible that participants used uncertainty related words in session 1 because they were getting familiar with the problem, and naturally recognized some of the uncertainties inherent to the problem based on the introduction session. This is realistic, as designers may very well recognize uncertainties affecting a system, although they may not necessarily devise strategies to deal with them effectively by means of flexibility. In session 2, teams under treatment 1 continued analyzing the problem under similar cognitive conditions, and did not necessarily use more uncertainty related words. This may be because they did not have the proper tools and framework to move from thinking about the uncertainties to solutions dealing with these uncertainties. Regarding teams under treatments 2, 3, and 4, the discussion content may have been diluted due to the lecture and prompting mechanisms. Instead of focusing solely on uncertainty, participants introduced thinking about flexibility, as well as design elements enabling flexibility, and decision rules to manage such flexibilities. These topics may not be fully captured by uncertainty related words, and therefore may not have created much difference as compared to session 1.

The second explanation may be that the experimental platform introduced in this thesis is not fully suited for content analysis of ideation transcripts. The author noticed significant changes in responses ΔUI (and even ΔFI) when uncertainty and flexibility related words were changed. Changing the list of words changed results and conclusions significantly. This may be because written transcriptions do not reflect well enough the overall content of discussions, because participants only record their thoughts partially. Too much conversational information may be lost in transcriptions without capturing the full content. While the information provided is enough to extract complete design concepts – because coders look for specific pieces of information – more information may be needed to understand underlying cognitive processes and discussion content. This is a consequence of the tradeoff exploited by this experimental methodology, which gives up accuracy in transcription reports for more analytical efficiency. This suggests it may be more appropriate to complement transcription with audio/video recordings of the sessions to study the content of discussions more appropriately, as often done in protocol studies.

On the other hand, results show that the lecture on flexibility did produce a significant effect on the influence of flexibility related words – in line with the stated hypothesis. This may be because participants felt stimulated by the content covered during the lecture, and made use of more flexibility words on average after the presentation. One reason why the prompting mechanism had a counter effect may be that – similar to what is mentioned above – the discussion was diluted between different aspects of uncertainty and flexibility using the prompting mechanism. This is because the prompting mechanism inherently divided discussions into four different topics.

8.4 Results Validity and Limitations

8.4.1 Internal Validity

A major difficulty in experimental work is to control for as many exogenous factors that may affect the dependent variable responses as possible. Without appropriate identification and control of such exogenous variables, it is difficult to conclude anything on perceived cause-effect relationships, which may affect the internal validity of results.

Many strategies were explored to control for undesired factors potentially affecting the responses, and to enhance internal validity. For example, a second independent coder – who is also an expert in the design problem – reviewed ideation transcripts to enhance interpretive validity. This strategy also alleviated concerns about inherent researchers' bias of trying to measure positive treatment effects. Each experiment was structured with two sessions to control for inherent creativity level and knowledge of flexibility – based on the structure of a pretest-posttest quasi-experimental design (Campbell & Stanley, 1966). The same ideation time was allocated to all teams. Having the same number of participants in each team – as much as possible – ensured that the number of participants in a team would not constitute a major factor affecting responses. Participants were selected to have as untainted knowledge of flexibility as possible, so responses could not be biased by expert knowledge of flexibility. Repeating the exact same procedure in all experiments ensured that different sequences of actions would not bias results. Providing the same content in all lectures on flexibility, and assigning the exact same task to each team controlled for variability in terms of information passed along to participants. Using the same location for all experiments as much as possible removed the possibility that the environment might be a factor affecting creativity and/or subjective impressions with the design procedure. Modifying and extending a survey that was already validated experimentally by Briggs et al. (2006) favored reliable survey responses. Measuring Cronbach's α for each construct demonstrated such inter-item reliability. Finally, participants did not know what treatment group they were assigned until the very end of the experiment, in the debrief session – randomized assignment. This was done to ensure that participants would not try harder in order to make the study successful in demonstrating that only treatments related to flexibility could have an effect. Researchers did not know either which of the four treatment conditions would lead to better results. It could have very well been observed that free undirected ideation and prior training only are the necessary conditions to produce improved responses, or that no difference could be observed between the treatment conditions.

Allocated time may also have been an important factor in generating flexible design concepts. Although this study does not establish a direct correlation, allocated time may very well explain the baseline *ΔENPV* improvement of 2.08 millions (see Section 7.3) observed among all teams – since it was a confounded variable in all treatment conditions. It is left as future work to measure and determine more explicitly the effect of time on the overall performance response.

Another internal validity issue regards the prompting mechanism, which nudges participants to think explicitly about uncertainty and flexibility. A legitimate concern is whether this is a valid component of the design procedure, as it may seem like answers were given away to participants. Logically, if the design procedure itself is flawed, the experimental evaluation results may as well be flawed. There are four answers addressing this concern.

The first answer is to outline the subtlety that the author knew *of* flexible strategies that could improve anticipated performance, but did not know *of all* the possible strategies participants could come up with. The design problem left enough freedom for creativity so that novel ideas could emerge. This was observed indeed, as participants identified strategies during sessions that the author did not think of. For instance, one team suggested developing units "just in time", implying they could follow demand exactly in each phase. Although the author thought about the concept of capacity expansion, the possibility of adjusting downwards was not equally considered! Other teams proposed solutions that did not make sense conceptually. For example, one team suggested to expand capacity when construction cost increases. Although this may have made sense during conversations, it did not seem appropriate as read from the ideation transcript. Other design solutions extracted from ideation transcripts are summarized in Appendix G. These design solutions demonstrate clearly that participants could come up with their own design concepts, without taking answers out of the prompting mechanism. Even if it looks as though answers were given away as part of the design procedure – which was not the case – all teams would have benefitted from it equally, and so no difference would be noticeable in the responses. It was necessary, however, to create and implement some flexible design solutions *a priori* to test and validate the computer model before analysis. This does not imply that all possible flexible design concepts were identified so that they could be packaged as part of the prompting mechanism. The author did not know ahead of time how participants would formulate their strategies, which ones they would formulate, and under what decision rules and proposed design implementations. The preliminary implementation step was necessary to have sufficient confidence that the model would be appropriate to implement most ideas suggested in experiments.

The second answer addressing the concern lies in the fact that design concepts were evaluated both objectively and quantitatively by modeling their future cash flows. Concepts were not evaluated solely based on the fact that they satisfied the four criteria of complete design concepts – which may seem arbitrary to neophytes. Even if the structure of the prompting mechanism was similar to the evaluation scheme to identify flexible design concepts – the prompting mechanism had the same four-step structure – satisfying the four criteria did not ensure anticipated performance improvement compared to the benchmark. The objective and quantitative cash flows of a proposed concept where modeled, and the resulting ENPV was compared to that of the benchmark. If the cash flows resulting from a flexible design concept did not improve the ENPV of the benchmark, the concept was not counted as a good, and not accounted for as a strategy to improve ENPV (i.e. it would not be selected by any rationale decision-maker). Thus what mattered really was whether participants could think of flexible design concepts, *and* whether these concepts could improve anticipated performance. In effect, the overall goal of this design procedure is to improve anticipated performance by means of flexibility, and not necessarily to be more flexible above all. If this were the goal, and only the quantity of complete flexible design concepts was accounted for, this would seem like a self-fulfilling prophecy to use a prompting mechanism based on the same four criteria that are used to evaluate design concepts. This was not the case here however, because cash flows emerging from design concepts were explicitly modeled and evaluated, and improvements in anticipated performance were clearly demonstrated.

The third answer can be understood by reading directly the prompts used in experiments in Appendix B. One notices that the prompting mechanism was constructed to be general, so it was not necessarily suited to the particular design problem in this study. The prompting mechanism could ostensibly be used to evaluate other design concepts related to a different engineering system, without any modification. This may in fact be the topic of a future study, to determine how ones needs to tailor the prompting mechanism to the problem of interest, or provide different levels of direction – as discussed next.

The fourth answer is to outline that the prompting mechanism was crafted carefully to stimulate flexible design concept generation by providing some level of direction, without being completely directive as to give away answers. In this regard, Figure 8.1 depicts conceptually a spectrum of "amount of direction" used in prompting mechanisms in general to stimulate creativity, ranging from no direction at all, to complete direction where answers are effectively given away. Santanen et al. (2004) studied the effects of prompting on solution quality and quantity when no direction at all is provided, locating this prompting mechanism on the left-hand side of the spectrum. Santanen and de Vreede (2004) studied the effect of different levels of direction in prompting mechanisms. In contrast, a prompting mechanism geared towards giving away answers to a particular problem would provide one hundred percent complete direction, on the right-hand side of the spectrum. The prompting mechanism suggested here errs somewhere between the two extremes, closer to the tail involving complete direction. It aims at providing some direction, although it does not completely give away answers as if it were tailored to the design problem at hand. This mechanism may be closer to the provocation change of perspective defined by Knoll and Horton (2010), because it challenges the underlying assumptions of the design problem.

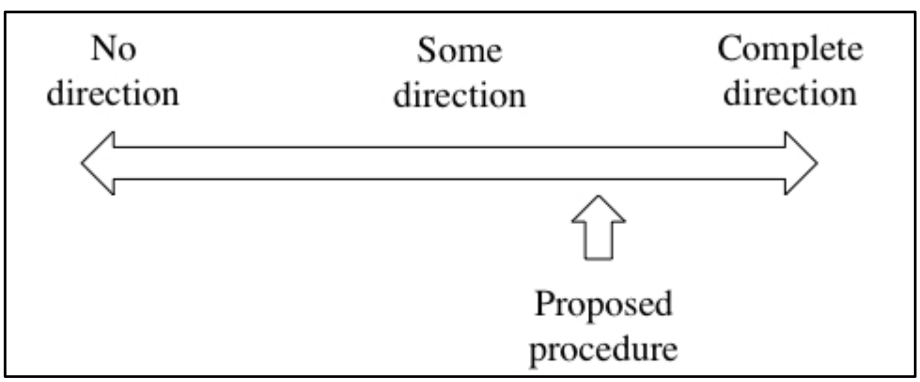


Figure 8.1: Conceptualization of the spectrum of "amount of direction" provided in prompting mechanisms to stimulate creativity. The extremes range from "no direction" to "complete direction" where answers are effectively given away.

In general, researchers espousing a directive prompting mechanism should be careful to select the appropriate amount of direction to stimulate creativity. The level of direction should consider the possibility that some designers may be more creative by modifying existing solutions, while others are better at creating entirely new solutions from scratch (Thompson & Lordan, 1999).

8.4.2 External Validity

An important criticism to external validity of results – the ability to generalize conclusions to wider populations – was the fact that experiments were done with graduate students, in a controlled setting that may not represent the reality of practicing engineers, and using a simplified design problem. Although this study demonstrates that the proposed design procedure for flexibility had significant effects in such context, this does not imply that the conclusions necessarily apply to the entire population of practicing engineers, and to all design problems in a real-world setting. Rather, these results suggest the design procedure may be productive and useful, and it is left as an opportunity for future research to study it in real-world practice.

Experimental conditions and the sample population were nonetheless representative of some of the realities of engineering practice. Participants were chosen carefully to represent the wider population of practicing engineers in many different industries. Graduate students in the selected programs, as opposed to undergraduate students, typically had many years of experience in industry, management, consulting, or elsewhere. Participants had different educational backgrounds, as is often the case in real engineering firms. To reflect this, cultural and personality differences arising in real-world practice were indeed noticed in experiments through inter-personal relations. The design problem was modeled through close interactions with experts at the MIT Center for Real Estate. The problem was simplified so it could be tackled within a short experimental session, but nonetheless represented some of the best practices in this field. One aspect that may have biased the population sample was the willingness to participate in these experiments (i.e. self-selection bias). This suggests that participants were already open-minded to trying new approaches, which may not reflect well the resistance that may naturally occur in industry when introducing a new design procedure.

Related to content analysis, a general conclusion is that the experimental platform introduced here seems only partially suitable for this kind of analysis. Preliminary results showed it is indeed feasible to perform content analysis, although the experimental platform may not be fully suited because too much conversational information is lost in the data recording process. The analysis did suggest the lecture on flexibility had an influence on discussion content related to flexibility. In contrast, the demonstration was not fully satisfactory because both factors *E* and *I* did not appear to affect content related to uncertainty, and even worsened influenced of flexibility-related content – counter to the stated and intuitive hypothesis. The experimental platform may not be completely suited for content analysis because it focuses on efficiency rather than accuracy of transcriptions. Content analysis on the other hand heavily relies on accurate transcriptions of discussions. Hence, protocol studies may be more appropriate for this kind of analysis, relying on full audio/video recording of experiments with designers talking out loud their thought process.

The feasibility of the analysis nonetheless opens the door for further research opportunities, to extend the current experimental platform to enable full audio/video recording of ideation sessions – at the cost of more analytical requirements.

Chapter 9 – Conclusion

"Many of the problems the world faces today are the eventual result of short-term measures taken last century." – Jay W. Forrester, Massachusetts Institute of Technology

"For tomorrow belongs to the people who prepare for it today." – African proverb

This thesis demonstrated that the proposed experimental methodology can be used to evaluate a design procedure effectively and efficiently based on anticipated performance. Evaluation results demonstrated that both the lecture and prompting mechanism significantly improved anticipated lifecycle performance of the engineering system compared to the benchmark design, by nearly thirty-six percent on average. The prompting mechanism significantly improved the generation of valuable flexible design concepts, while lecturing improved user satisfaction with the process and results, as well as result quality assessments. Even though prompting demonstrably improved anticipated performance and flexible design concept generation, it did not have an effect on participants' satisfaction with the process and results. Also, prompting did not lead participants to expect better result quality. It is only when combined with the lecture that prompting improved user satisfaction. A preliminary analysis suggested that the experimental platform can be used to study the influence of uncertainty and flexibility related words on discussion content, although more work is necessary to validate the approach.

An important finding is that design procedures should be evaluated using both subjective user impressions, and objective quantitative measurements. Evaluating a procedure based on user impressions may miss important contributions based on anticipated performance of design concepts – as shown here since participants did not think the prompting mechanism would improve quality of results. Similarly, relying solely on anticipated performance may not allow measurements of user satisfaction with the process and results. Hence a procedure may be effective in improving anticipated performance, but will be rejected in practice because too cumbersome and not user-friendly. Although this was not the case here – participants showed improved impressions of satisfaction with the process and results – this is an important attribute to consider for any design procedure aiming at real impact in industry practice.

The experimental methodology integrated a set of computer-aided techniques to promote effectiveness and efficiency based on GSS technology, computer-based modeling, and content analysis software. The design procedure for flexibility involved a short lecture on flexibility, and a prompting ideation mechanism focusing on flexibility in engineering systems design. It complemented participants' inherent design approach by leading them to a better design outcome with overall improved anticipated performance.

Collaborative design experiments were performed where participants suggested alternative solutions to a design problem under different treatment conditions. Experimental conditions made use of the procedure for flexibility, while control conditions relied on prior training in science and engineering only (i.e. no lecture), and a free undirected ideation mechanism (i.e. no prompting). The design procedure for flexibility was evaluated in a controlled setting based on its cognitive effects on the quantity of flexible design concepts generated, anticipated economic performance improvements compared to a benchmark design, participants' subjective impressions of satisfaction with the process and results, and quality assessments of results. Seventy-one experienced designers divided among twenty-six collaborative teams performed the experiment involving a simplified real estate infrastructure design problem.

The methodology represents an important component of an experimental platform that aims at evaluating design procedures for concept generation thoroughly, rigorously, and efficiently in controlled laboratory and real-world settings. It complements the body of existing evaluation methodologies relying on subjective expert performance assessments, or other objective metrics not explicitly measuring anticipated performance (e.g. development time and other utility attributes). It is hoped that the experimental platform will be used to study other design procedures of interest, to determine their potential before beginning a deeper field study involving application in industry.

This work also provided another case example that specifically focusing the design effort on uncertainty and flexibility can improve the expected value and anticipated performance of an engineering system – in this case a simplified real estate development project. The fact that such value can be measured objectively and quantitatively enables designers to investigate more thoroughly the sets of worthwhile flexible strategies, and compare them to their acquisition costs. This was demonstrated in these experiments, as many experienced designers proposed flexible design alternatives that improved anticipated performance by nearly thirty-six percent compared to the benchmark. For projects requiring upfront capital investments in millions and billions of dollars, this is important to consider. This kind of approach brings more rigors in the design process, especially in light of the challenges inherent to designing for uncertainty and flexibility.

9.1 Extending Current Approaches to Design

Early conceptual design activities are of utmost importance as they affect the way our critical infrastructures work, which ultimately affect local and global economies. The early conceptual phase is where designers and decision-makers can have the most influence on anticipated performance, schedule, and costs of engineering systems. The U.S. National Research Council (1991) estimated that nearly seventy percent of the life cycle cost is determined during conceptual design.

Early conceptual design procedures are needed to help designers think at system-level, integrating as much as possible understanding of economic, social, and technological forces. This training is not widespread for a variety of reasons. One reason may be the current education paradigm – in science and engineering, but also in other domains – focusing on understanding the parts of a system, with the implicit assumption that this approach naturally leads to an understanding of the whole. It is not clear whether this assumption holds true. In fact, the finest institutions in the world are just getting started at crafting education programs integrating engineering, management, and social sciences to better prepare tomorrow's leaders and designers to deal with such system-level perspective. The development of Engineering Systems-like programs may result from the realization that more explicit training is needed to understand how complex systems really work, as a whole, rather then focusing deeply on each individual part.

To address these issues and extend current approaches to design, this thesis stressed the need to recognize today's complexities and uncertainties in engineering systems design activities, and to deal with them pro-actively in design. As demonstrated in Section 2.4.1, this is not something that existing design procedures taught at engineering institutions, and used in industry, do completely satisfactorily. Many design procedures assume that customer demands and preferences, market conditions, operating environments, and regulatory frameworks are known *a priori*, so that engineers can rely on deterministic projections, freeze design requirements early, and optimize the system accordingly (de Neufville & Scholtes, 2011; Eckert et al., 2009). Such approach to design can lead to incorrect evaluation and design choices. Design activities should naturally evolve to tackle the complexities of today's engineering systems. Hence the need to develop novel design procedures incorporating thinking and recognition of a wide range of uncertainty factors explicitly in the early design cycle.

This thesis contributed specifically towards a new approach to design by suggesting a procedure that is based on rigorous economic, engineering, and risk management principles, but presented and used in a way that is simple and accessible, for most impactful contributions to industry. The demonstration application of the design procedure done here is for an infrastructure system, but nothing prevents it to be applicable to other systems like financial systems, manufacturing systems, product development, etc. It is left as a future research opportunity to demonstrate the effects of the procedure on real-world engineering systems.

9.2 Pursuing Real Impact on Design Practice

Having real impact on design practice to improve anticipated performance of engineering systems was the spirit of this thesis. The goal was to help designers consider uncertainty and risk more explicitly in the early conceptual design through a series of simple questions, when there is still a wide margin for impactful decisions. This justifies the proposal of devising an efficient, quick, and simple procedure that can be incorporated quickly to existing design process at a firm or company. The aim was to stimulate designers' creativity to identify major uncertainty sources affecting future performance, and crafting relevant flexible strategies to deal with these uncertainties pro-actively. This approach supports generation of flexible design concepts without adding too much overhead to the existing design process at a firm. A design procedure that is intuitive and user friendly was desirable, without bringing designers out of their comfort zone.

This thesis contributed to research and design practice by providing an experimental methodology to evaluate thoroughly, rigorously, and efficiently a design procedure of interest. This may contribute to engineering design research, by encouraging structure to design studies, along the lines promoted Frey and Dym (2006), Reich (2010), and Tomiyama et al. (2009).

9.3 Future Research Opportunities

This thesis opens the door to many exciting research opportunities. Using the existing experimental methodology enables evaluation and comparison of other design procedures for concept generation, such as Axiomatic Design, Pahl and Beitz, and TRIZ. An interesting study could compare design procedures based on anticipated performance of the design concepts. Evaluating other design procedures and demonstrating applicability to more than one design procedure should contribute to further validating the effectiveness and efficiency of the experimental methodology. Validating this methodology may involve as well applying it in different contexts, for different engineering systems in the aerospace, energy, mining, product development and manufacturing, transportation, and real estate industries, to name a few.

While the experimental methodology suggested here might be most useful to evaluate design procedures for concept generation, more research can be devoted to develop other experimental platforms enabling thorough and rigorous evaluation of design procedures of interest. For example, the proposed platform can be modified to enable evaluation of design procedures for design space exploration, as well as management and representation of the design process (see Table 2.1). For example the platform could be used to test and evaluate procedures for product safety (e.g. FMEA described by McDermott et al., 1996), or reliability (e.g. Robust Design developed by Taguchi, 1987). The platform could rely on different mechanisms to record ideation outputs than ThinkTank®. For instance, a computer-based design sketching tools could be used (e.g. see iCampus MIT-Microsoft Alliance, 2010). The platform could be extended to enable more detailed analysis on conversation contents, for instance through audio/video recording of ideation sessions. Computer-based transcription using speech recognition software could speed up the data processing phase. The content analysis performed here suggests this kind of analysis may be feasible, but too much information may be lost in transcription. Thus the full content of discussions may not be fully captured through this efficient – but less accurate –

recording mechanism, which standard audio/video recording can complement. Content analysis could further be re-organized by selecting uncertainty and flexibility-related keywords based on the criteria defining a complete idea.

Another research avenue could be to use the experimental platform for quick and efficient preliminary evaluation of a design procedure in industry, before more lengthy and/or longitudinal field study – as opposed to conducting controlled experiments only in an academic setting. For instance, a study could be conducted first at a company on a relevant but simplified design problem to determine the potential of the procedure based on anticipated performance. If such potential is established, researchers could move on to more lengthy application, evaluation, and implementation of the design procedure – with tangible benefits to the real system implementation and operations. This approach would allow at a minimum filtering out for design procedures that may not be as impactful as theoretically thought. It would also enable comparisons between the anticipated performance of design concepts generated in experiments, and the actual performance outcomes – if the study is long enough to see the system being built.

Regarding the more specific issue of designing for flexibility, an interesting avenue could be to develop an integrated package or course to support such design thinking, similar to the approach pursued by the lean Six Sigma initiative (iSixSigma, 2010). The design procedure for flexibility studied here focuses on concept generation only, so one needs to evaluate and integrate other procedures for management and representation, as well as design space exploration. For example, one could compare the effectiveness of DBD, MATE, and screening models in terms of exploring the design space for valuable flexible design opportunities. One could compare DBD to the PuCC mechanism for concept convergence and selection. It would be interesting as well to compare the C-DSM, CPA, ESM, and sDSM to determine which one(s) give the best results in terms of management and representation of flexible engineering systems, and which ones enable better identification of existing opportunities for flexibility. Depending on the outcome in each category, the best tools could be packaged and taught as an integrated professional development course supporting design for flexibility. Experimental and case study evaluations would validate such course more thoroughly.

Bibliography

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

Arnold, T. A., & Crack, T. F. (2003). *Option Pricing in the Real World: A Generalized Binomial Model with Applications to Real Options*. Paper presented at the Real Options Conference, Washington, D. C., United States.

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

Avnet, M. S. (2009). *Socio-Cognitive Analysis of Engineering Systems Design: Shared Knowledge, Process, and Product*. Doctoral Dissertation in Engineering Systems, Massachusetts Institute of Technology, Cambridge, MA, United States.

Axtell, R. L. (2005). Three Distinct Kinds of Empirically-Relevant Agent-Based Models. Washington, D. C., United States: Brookings Institution, Center on Social and Economic Dynamics.

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

Balthazard, P., Ferrell, W. R., & Aguilar, D. L. (1998). Influence Allocation Methods in Group Decision Support Systems. *Group Decision and Negotiation, 7*(4), 347-362.

Bartolomei, J. (2007). *Qualitative Knowledge Construction for Engineering Systems: Extending the Design Structure Matrix Methodology in Scope and Perspective*. Doctoral Dissertation in Engineering Systems, Massachusetts Institute of Technology, Cambridge, MA, United States.

Berman, O., & Larson, R. C. (2004). A Queueing Control Model for Retail Services Having Backroom Operations and Cross-Trained Workers. *Computers and Operations Research, 31*, 201-222.

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

Boehm, B., Gruenbacher, P., & Briggs, R. (2001). Developing Groupware for Requirements Negotiation: Lessons Learned. *IEEE Computer, 18*(3), 46-55.

Bostrom, R. P., & Nagasundaram, M. (1998). *Research in Creativity and GSS*. Paper presented at the Proceedings of the 31st Annual Hawaii International Conference on System Sciences, Kohala Coast, HI, United States.

Boyne, W. (2001). Fifty Years of the B-52. *Air Force Magazine,* 50-57.

Briggs, R. O., & de Vreede, G.-J. (2001). ThinkLets, Building Blocks for Concerted Collaboration. Netherlands: Delft University of Technology.

Briggs, R. O., de Vreede, G.-J., & Nunamaker, J. F. (2003). Collaboration Engineering with ThinkLets to Pursue Sustained Success with Group Support Systems. *Journal of Management Information Systems, 19*(4), 31-63.

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

Broniatowski, D. (2010). *A Method for Analysis of Expert Committee Decision-Making Applied to FDA Medical Device Panels.* Doctoral Dissertation in Engineering Systems, Massachusetts Institute of Technology, Cambridge, MA, United States.

Browning, T. R. (2001). Applying the Design Structure Matrix to System Decomposition and Integration Problems: A Review and New Directions. *IEEE Transactions on Engineering Management, 48*(3), 292-306.

Campbell, D. T., & Stanley, J. C. (1966). *Experimental and Quasi-experimental Designs for Research*. United States: Houghton Mifflin.

Cardin, M.-A., de Neufville, R., & Kazakidis, V. (2008). A Process to Improve Expected Value in Mining Operations. *Mining Technology: IMM Transactions, Section A, 117*(2), 65-70.

Cardin, M.-A., Steer, S. J., Nuttall, W. J., Parks, G. T., Gonçalves, L. V. N., & de Neufville, R. (2010). Minimizing the Cost of Innovative Nuclear Technology Through Flexibility: The Case of an Accelerator-Driven Subcritical Reactor Park Demonstrator. *University of Cambridge Electricity Policy Research Group Working Paper Series*. Retrieved from http://www.eprg.group.cam.ac.uk

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

Corman, S. R., Kuhn, T., McPhee, R. D., & Dooley, K. J. (2002). Studying Complex Discursive Systems – Centering Resonance Analysis of Communication. *Human Communication Research, 28*(2), 157-206.

Cortina, J. M. (1993). What is Coefficient alpha? An Examination of Theory and Applications. *Journal of Applied Psychology, 78*, 98-104.

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

de Neufville, R. (2010). *ESD.71: Engineering Systems Analysis for Design* [Selected Lecture Notes and Exercises]. Cambridge, MA, United States: Massachusetts Institute of Technology.

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

de Neufville, R., Scholtes, S., & Wang, T. (2006). Valuing Options by Spreadsheet: Parking Garage Case Example. *ASCE Journal of Infrastructure Systems, 12*(2), 107-111.

de Toni, A., & Tonchia, S. (1998). Manufacturing Flexibility: A Literature Review. *International Journal of Production Research, 36*(6), 1587-1617.

de Vreede, G.-J., Boonstra, J., & Niederman, F. A. (2002). *What is Effective GSS Facilitation? A Qualitative Inquiry into Participants' Perceptions*. Paper presented at the Proceedings of the Annual Hawaii International Conference on System Sciences, Los Alamitos, CA, United States.

de Vreede, G.-J., & Briggs, R. O. (2005). *Collaboration Engineering: Designing Repeatable Processes for High-Value Collaborative Tasks*. Paper presented at the Proceedings of the 38th Annual Hawaii International Conference on System Sciences, Big Island, HI, United States.

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

de Weck, O. L., de Neufville, R., & Chaize, M. (2004). Staged Deployment of Communications Satellite Constellations in Low Earth Orbit. *Journal of Aerospace Computing, Information, and Communication, 1*, 119-136.

Dennis, A. R., & Valacich, J. S. (1993). Computer Brainstorms: More Heads Are Better Than One. *Journal of Applied Psychology, 78*(4), 531-536.

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

Dong, Q. (2002). *Predicting and Managing System Interactions at Early Phase of the Product Development Process*. Doctoral Dissertation in Mechanical Engineering, Massachusetts Institute of Technology, Cambridge, MA, United States.

Dorr, R. F., & Peacock, L. (1995). *Boeing's Cold War Warrior – B-52 Stratofortress*. Botley, United Kingdom: Osprey Publishing.

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

Eppinger, S. D., Whitney, D. E., Smith, R. P., & Gebala, D. A. (1994). A Model-Based Method for Organizing Tasks in Product Development. *Research in Engineering Design, 6*, 1-13.

Ferguson, S., Siddiqi, A., Lewis, K., & de Weck, O. L. (2007). *Flexible and Reconfigurable Systems: Nomenclature and Review*. Paper presented at the ASME International Design

Engineering Technical Conferences and Computers and Information in Engineering Conference, Las Vegas, NV, United States.

Findlay, A. (1948). *A Hundred Years of Chemistry* (2nd ed.). London, United Kingdom: Duckworth.

Finger, S., & Dixon, J. R. (1989a). A Review of Research in Mechanical Engineering Design. Part I: Descriptive, Prescriptive, and Computer-Based Models of Design Processes. *Research in Engineering Design, 1*(1), 51–67.

Finger, S., & Dixon, J. R. (1989b). A Review of Research in Mechanical Engineering Design. Part II: Representations, Analysis, and Design for the Life Cycle. *Research in Engineering Design, 1*(2), 121-137.

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

Frey, D. D. (2008). Hypothesis Testing. *ESD.86: Models, Data and Inference for Socio-Technical Systems* [Lecture Notes]. Cambridge, MA, United States: Massachusetts Institute of Technology.

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

Frey, D. D., Herder, P. M., Wijnia, Y., Subrahmanian, E., Katsikopoulos, K., & Clausing, D. P. (2009). The Pugh Controlled Convergence Method: Model-Based Evaluation and Implications for Design Theory. *Research in Engineering Design, 20*, 41-58.

Frey, D. D., & Li, X. (2010). *Model-Based Validation of Design Methods*. Book chapter in preparation. Massachusetts Institute of Technology. Cambridge, MA, United States.

Frey, D. D., & Wang, H. (2006). Adaptive One-Factor-at-a-Time Experimentation and Expected Value of Improvement. *Technometrics, 48*(3), 418-431.

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

Fumarola, M., Seck, M., & Verbraeck, A. (2010). *A DEVS Component Library for Simulation-based Design of Automated Container Terminals*. Working Paper. Systems Engineering Section. Delft University of Technology. Netherlands.

Geltner, D., & Cardin, M.-A. (2008). Engineering Economy Module. *11.431: Real Estate Finance and Investment* [Lecture Notes]. Cambridge, MA, United States: Massachusetts Institute of Technology.

Giffin, M., de Weck, O. L., Bounova, G., Keller, R., Eckert, C., & Clarkson, J. (2009). Change Propagation Analysis in Complex Technical Systems. *Journal of Mechanical Design, 131*(8), 1-14.

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

Gordon, P. C., Grosz, B. J., & Gilliom, L. A. (1993). Pronouns, Names, and the Centering of Attention in Discourse. *Cognitive Science, 17*(311-347).

Gu, P., Rao, H., & Tseng, M. M. (2001). Systematic Design of Manufacturing Systems Based on Axiomatic Design Approach. *Annals of CIRP, 50*(1), 299–304.

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

Guma, A. (2008). *A Real Options Analysis of a Vertically Expandable Real Estate Development*. Master of Science Thesis in Real Estate Development, Massachusetts Institute of Technology, Cambridge, MA, United States.

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

Hassan, R., & de Neufville, R. (2006). *Design of Engineering Systems Under Uncertainty via Real Options and Heuristic Optimization*. Paper presented at the 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.

Hesseldahl, A. (2001). The Return of Iridium. Retrieved November 1, 2010, from http://www.forbes.com/2001/11/30/1130tentech.html

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

Horváth, I. (2004). A Treatise on Order in Engineering Design Research. *Research in Engineering Design, 15*(3), 155–181.

iCampus MIT-Microsoft Alliance. (2010). Magic Paper: Introducing Natural Interaction. Retrieved June 25, 2010, from http://icampus.mit.edu/magicpaper/

iSixSigma. (2010). iSixSigma – Quality Resources for Achieving Six Sigma Results. Retrieved October 15, 2010, from http://www.isixsigma.com/

Jacoby, H., & Loucks, D. (1972). The Combined Use of Optimization and Simulation Models in River Basin Planning. *Water Resources Research, 8*(6), 1401-1414.

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

Kalligeros, K. (2006). *Platforms and Real Options in Large-Scale Engineering Systems*. Doctoral Dissertation in Engineering Systems, Massachusetts Institute of Technology, Cambridge, MA, United States.

Keeney, R. L. (2009). The Foundations of Collaborative Group Decision. *International Journal of Collaborative Engineering, 1*(1-2), 4-18.

Kepner, C. H., & Tregoe, B. B. (1981). *The New Rational Manager*. NJ, United States: Princeton Research Press.

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

Kolfschoten, G., Briggs, R. O., Duivenvoorde, G. P. J., & de Vreede, G.-J. (2009). *Practitioners vs. Facilitators a Comparison of Participant Perceptions on Success*. Paper presented at the 42nd Annual Hawaii International Conference on System Sciences, Big Island, HI, United States.

Kurtoglu, T., & Campbell, M. I. (2009). An Evaluation Scheme for Assessing the Worth of Automatically Generated Design Alternatives. *Research in Engineering Design, 20*, 59-76.

Kurtoglu, T., Campbell, M. I., & Linsey, J. S. (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.

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

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

Malmstrom, J., & Malmquist, J. (1998). Trade-Off Analysis of Product Decompositions. Retrieved from http://w3.ppd.chalmers.se/~joma/publications/nd98-paper.ps

McDermott, R. E., Mikulak, R. J., & Beauregard, M. R. (1996). *The Basics of FMEA*. New York, NY, United States: Productivity Press.

Mikaelian, T. (2009). *An Integrated Real Options Framework for Model-based Identification and Valuation of Options under Uncertainty*. Doctoral Dissertation in Aeronautics and Astronautics, Massachusetts Institute of Technology, Cambridge, MA, United States.

Miller, A. (1984). *Imagery in Scientific Thought: Creating 20th Century Physics*. Boston, MA, United States: Birkhäuser.

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

Mizuno, S., & Akao, Y. (1993). *QFD: The Customer-Driven Approach to Quality Planning & Deployment*. Tokyo, Japan: Asian Productivity Organization.

Montulli, L. T. (1986). Lessons Learned from the B-52 Program Evolution: Past, Present, and Future *AIAA/AHS/ASEE Aircraft Systems, Design and Technology Meeting*. Dayton, OH, United States.

Morgan, M. G., & Henrion, M. (1990). *Uncertainty: A Guide to Dealing with Uncertainty in Quantitative Risk and Policy Analysis*. United Kingdom: Cambridge University Press.

Mullen, B., Johnson, C., & Salas, E. (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.

National Research Council, U. S. (1991). *Improving Engineering Design: Designing for Competitive Advantage*. Washington, D. C., United States: National Academy Press.

Nightingale, D. J., & Rhodes, D. H. (2007). *ESD.38: Enterprise Architecting* [Lecture Notes]. Cambridge, MA, United States: Massachusetts Institute of Technology.

Nunamaker, J. F., Briggs, R. O., Mittleman, D. D., Vogel, D. R., & Balthazard, P. A. (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.

Nunnally, J. C. (1978). *Psychometric Theory*. New York, NY, United States: McGraw-Hill.

Olewnik, A., Brauen, T., Ferguson, S., & Lewis, K. (2001). *A Framework for Flexible Systems and Its Implementation in Multi-Attribute Decision-Making*. Paper presented at the ASME International Design Engineering Technical Conferences and Computers and Information in Engineering Conference, Pittsburgh, PA, United States.

Olewnik, A., & Lewis, K. (2006). A Decision Support Framework for Flexible System Design. *Journal of Engineering Design, 17*(1), 75-97.

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

Ostergaard, K., Wetmore, W., Divekar, A., Vitali, H., & Summers, J. D. (2005). An Experimental Methodology for Investigating Communication in Collaborative Design Review Meetings. *CoDesign, 1*(1), 1-17.

Otto, K., & Holtta, K. (2007). A Multi-Criteria Framework for Screening Preliminary Product Platform Concepts. *Journal of Intelligent Manufacturing, 18*, 59-75.

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

Pimmler, T. U., & Eppinger, S. D. (1994). *Integration Analysis of Product Decompositions*. Paper presented at the ASME Design Theory and Methodology Conference, Minneapolis, MN, United States.

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 pp. 225-132 (Part II).

Pugh, S. (1991). *Total Design: Integrated Methods for Successful Product Engineering*. New York, NY, United States: Prentice Hall.

Rajan, P. K., Wie, M. V., Campbell, M. I., Wood, K. L., & Otto, K. N. (2005). An Empirical Foundation for Product Flexibility. *Design Studies, 26*(4), 405-438.

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

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

Reuleaux, F. (1861). *Konstrukteur*. Braunschweig, Germany: Vieweg und Sohn.

Reuleaux, F. (1875). *Theorische Kinematic*. Braunschweig, Germany: Vieweg und Sohn.

Ross, A. (2006). *Managing Unarticulated Value: Changeability in Multi-Attribute Tradespace Exploration*. Doctoral Dissertation in Engineering Systems, Massachusetts Institute of Technology, Cambridge, MA, United States.

Saleh, J. H., & Hastings, D. E. (2000). *On Flexibility In Design: Analyzing Flexibility of Space Systems*. Paper presented at the AIAA Space 2000 Conference and Exposition, Long Beach, CA, United States.

Saleh, J. H., Mark, G., & Jordan, N. C. (2008). Flexibility: a Multi-Disciplinary Literature Review and a Research Agenda for Designing Flexible Engineering Systems. *Journal of Engineering Design, 1*(1-17).

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

Santanen, E. L., & de Vreede, G. J. (2004). *Creative Approaches to Measuring Creativity: Comparing the Effectiveness of Four Divergence Thinklets*. Paper presented at the Hawaii International Conference on System Sciences, Los Alamitos, CA, United States.

Savage, S. (2000, October 8). The Flaw of Averages, *San Jose Mercury News*.

Sethi, A. K., & Sethi, S. P. (1990). Flexibility in Manufacturing: A Survey. *The International Journal of Flexible Manufacturing Systems, 2*, 289-328.

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

Shah, J. J., Vargas-Hernandez, N., Summers, J. D., & Kulkarni, S. (2001). Collaborative Sketching (C-Sketch) - An Idea Generation Technique for Engineering Design. *Journal of Creative Behavior, 35*(3), 168-198.

Shen, Y., Ong, S. K., & Nee, A. Y. C. (2010). Augmented Reality for Collaborative Product Design and Development. *Design Studies, 31*(2), 118-145.

Silver, M. R., & de Weck, O. L. (2007). Time-Expanded Decision Networks: A Framework for Designing Evolvable Complex Systems. *Systems Engineering, 10*(2), 167-186.

Simon, H. A. (1960). *The New Science of Management Decision*. New York, NY, United States: Harper and Row.

Simpson, T. W. (2003). *Product Platform Design and Customization: Status and Promise*. Paper presented at the ASME International Design Engineering Technical Conferences and Computers and Information in Engineering Conference, Chicago, IL, United States.

Skiles, S. M., Singh, V., J., K., Wood, K., Jensen, D., & Szmerekovsky, A. (2006). *Adapted Concept Generation and Computation Techniques for the Application of a Transformer Design Theory*. Paper presented at the ASME International Design Engineering Technical Conferences and Computers and Information in Engineering Conference, Philadelphia, PA, United States.

Slack, N. (2005). The Flexibility of Manufacturing Systems. *International Journal of Operations and Production Management, 25*(12), 1190-1200.

Stagney, D. B. (2003). *The Integrated Concurrent Enterprise*. Master of Science Thesis in Aeronautics and Astronautics, Massachusetts Institute of Technology, Cambridge, MA, United States.

Steel, K. D. (2008). *Energy System Development in Africa: The Case of Grid and Off-Grid Power in Kenya*. Doctoral Dissertation in Engineering Systems, Massachusetts Institute of Technology, Cambridge, MA, United States.

Steward, D. V. (1981). The Design Structure System: A Method for Managing the Design of Complex Systems. *IEEE Transactions on Engineering Management, 28*, 71-74.

Straker, D., & Rawlinson, G. (2003). How to Invent (Almost) Anything. Retrieved from http://creatingminds.org/

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

Suh, E. S. (2005). *Flexible Product Platforms*. Doctoral Dissertation in Engineering Systems, Massachusetts Institute of Technology, Cambridge, MA, United States.

Suh, E. S., de Weck, O. L., & Chang, D. (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.

Suh, N. P. (1998). Axiomatic Design Theory for Systems. *Research in Engineering Design, 10*(4), 189-209.

Suh, N. P. (2001). *Axiomatic Design: Advances and Applications*. United Kingdom: Oxford University Press.

Summers, J. D., Anandan, S., & Teegavarapu, S. (2009). Introduction of Design Enabling Tools: Development, Validation, and Lessons Learned. In A. B. Markman & K. L. Wood (Eds.), *Tools for Innovation*. 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.

Taylor, D. W., Berry, P. C., & H., B. C. (1958). Does Group Participation When Using Brainstorming Facilitate or Inhibit Creative Thinking? *Administrative Science Quarterly, 3*(1), 23-47.

The New Oxford American Dictionary. (2006). (2nd ed.). United Kingdom: Oxford University Press.

Thompson, G., & Lordan, M. (1999). *A Review of Creativity Principles Applied to Engineering Design*. Paper presented at the Proceedings of the Institution of Mechanical Engineers.

Tichem, M. (1997). *A Design Coordination Approach to Design for X*. Netherlands: Delft University Press.

Tomiyama, T. (2006). *A Classification of Design Theories and Methodologies*. Paper presented at the ASME International Design Engineering Technical Conferences and Computers and Information in Engineering Conference, Philadelphia, PA, United States.

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

Tomiyama, T., & Yoshikawa, H. (1987). Extended General Design Theory. In Y. H. & E. A. Warman (Eds.), *Design Theory for CAD* (pp. 95–130). Amsterdam, North-Holland, Netherlands.

Trauth, E. M., & Jessup, L. M. (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.

Trochim, W. M. K. (2006). Research Methods Knowledge Base. Retrieved from http://www.socialresearchmethods.net/

Ullman, D. (1995). *The Mechanical Design Process*. New York, NY, United States: McGraw-Hill.

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

van der Lugt, R. (2002). Brainsketching and How it Differs from Brainstorming. *Creativity and Innovation Management, 11*(1), 43-54.

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

Wang, T. (2005). *Real Options in Projects and Systems Design – Identification of Options and Solutions for Path Dependency*. Doctoral Dissertation in Engineering Systems, Massachusetts Institute of Technology, Cambridge, MA, United States.

Wang, T., & de Neufville, R. (2005). *Real Options 'In' Projects*. Paper presented at the Real Options Conference, Paris, France.

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

Ward, A. C., & Seering, W. P. (1993). Quantitative Inference in a Mechanical Design Compiler (A. I. Laboratory, Trans.) *A. I. Memo No. 1062*. Cambridge, MA, United States: Massachusetts Institute of Technology.

Warr, A., & O'Neill, E. (2005). *Understanding Design as a Social Creative Process*. Paper presented at the 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.

Wetmore, W. R., Summers, J. D., & Greenstein, J. S. (2010). Experimental Study of Influence of Group Familiarity and Information Sharing on Design Review Effectiveness. *Journal of Engineering Design, 21*(1), 111-126.

Wilds, J. (2008). *A Methodology for Identifying Flexible Design Opportunities*. Master of Science Thesis in Technology and Policy, Massachusetts Institute of Technology, Cambridge, MA, United States.

Womack, J. P., Jones, D. T., & Roos, D. (1991). *The Machine that Changed the World: The Story of Lean Production*. New York, NY, United States: Free Press.

Wu, C. F. J., & Hamada, M. (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*. Paper presented at the ASME International Design Engineering Technical Conferences and Computers and Information in Engineering Conference, Las Vegas, NV, United States.

Yang, M. (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.

Yoshikawa, H. (1981). General Design Theory and a CAD System. In T. Sata & E. Warman (Eds.), *Man–Machine Communication in CAD/CAM* (pp. 35–58). Amsterdam, North-Holland, Netherlands.

Zuo, Q., Leonard, W., & MaloneBeach, E. E. (2010). Integrating Performance-Based Design in Beginning Interior Design Education: An Interactive Dialog Between the Built Environment and Its Context. *Design Studies, 31*(3), 268-287.

Appendix A – Slides for Lecture on Flexibility

Flexibility In Infrastructure Design and Management

Michel-Alexandre Cardin, PhD Candidate
Spring 2010

MIT ESD
Massachusetts Institute of Technology
Engineering Systems Division

TU Delft
Delft University of Technology

Definitions of Uncertainty

- "Anything that can impact promises or business objectives in the future" (Verbraeck, 2010)
- Uncertainty types (McManus and Hasting, 2005)
  - Exogenous: out of managerial control
  - Endogenous: managers/designers can control
  - Statistically characterized: price or demand
  - Known unknowns: future budget, system performance
  - Unknown unknowns: hard to quantify, not considered in design

MIT ESD Massachusetts Institute of Technology Engineering Systems Division
2
TU Delft Delft University of Technology

Human (In-)ability to Forecast

Heavier-than-air flying machine are impossible.

Lord Kelvin – British Mathematician, Physicist, and President of the British Royal Society, c. 1895

Everything that can be invented has been invented.

Charles H. Duell – Commissioner of the U.S. Patent Office, 1899

Reagan doesn't have the presidential look.

United Artists Executive – dismissing Ronald Reagan for the starring role in the movie of THE BEST MAN, 1964

Environmental Uncertainty

http://www.hurricanekatrinanews.org

http://www.jordoncooper.com

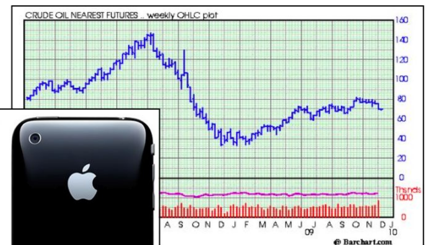
http://www.washingtonpost.com

Market, Technology, and Others…

http://www.growthstock.com

Stock Market

Crude Oil Price



http://images.angelpub.com

http://img.gsmarena.com

Technology

Uncertainty and Design

- Forecast can be (severely) wrong
  - No escape from this
  - … infrastructure are long-lived, undergo much variations
  - … analysis based on too many assumptions
  - … inevitable surprises
  - … typically over-optimistic about outcomes
  - … over-confident about prediction errors
- There are many design choices beyond obvious ones
  - Typically, combining different characteristics
  - … enabling different future designs
  - … and are thus more flexible

modified from de Neufville, 2008

Current Design Process

- Very successful, BUT…
- Based on deterministic forecasts, heuristics, point value estimates (e.g. $60/oil barrel, $1/pound of copper)
- Optimized for limited set of conditions
- Uncertainty considered *ex post* through sensitivity analysis
- "Compartmentalized" engin. and manag't
- Often focused on risk minimization
- OPPORTUNITY TO DO BETTER!

http://www.orchardscotts.com.sg/

MIT ESD Massachusetts Institute of Technology Engineering Systems Division

7

TU Delft Delft University of Technology

Issues with Current Process

- Uncertainty affects performance
  - Downside scenarios ⇒ risk of losses, lower performance
  - Upside scenarios ⇒ opportunities for additional gains
- "Fixed" design and management miss opportunities to recognize additional value!
  - Can be sub-optimal… when reality departs from forecast!
  - Cannot "reduce" exposure to risk easily
  - Cannot "seize" good opportunities easily
  - Adapting project can be more costly
  - Typical project valuation does not account for flexibility

MIT ESD Massachusetts Institute of Technology Engineering Systems Division

8

TU Delft Delft University of Technology

Better Approach

- ☐ Recognize uncertainty *a priori*
  - Range, distribution of possible outcomes
- ☐ Plan for **flexibility** in design and management
  - E.g. vertical building expansion in Chicago

Start smaller

Reduce exposure to downside market risk

Guma, 2008

Expand when needed

Extra gains on upside market opportunity

MIT ESD Massachusetts Institute of Technology Engineering Systems Division

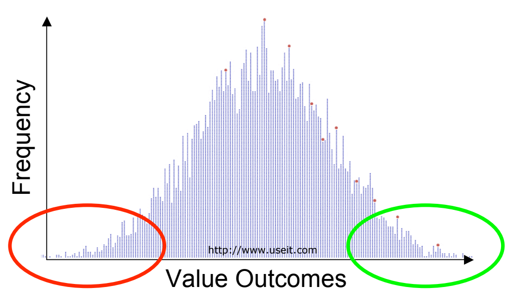
TU Delft Delft University of Technology

Better Approach (cont.)

- ☐ Find designs acting on outcome distribution, rather than optimizing point forecast

Aim at shifting whole distribution towards better value

Cut bad, downside outcomes



Improve on good, upside outcomes

MIT ESD Massachusetts Institute of Technology Engineering Systems Division

TU Delft Delft University of Technology

Why Better with Flexibility?

- Forces considering downsides explicitly
  - And prepare for it
  - Prevents over-optimism
- Forces considering upsides as well
  - Position to capture upside opportunities
- Both improve expected (or "average") performance compared to fixed design
- No absolute best guaranteed, but can do BETTER THAN FIXED (INFLEXIBLE) DESIGN!

Example Flexible Strategies

- Flexibility "on" system: managerial
  - Defer investment until favorable conditions
  - Growth through R&D investment
  - Abandon temporarily or permanently
- Flexibility "in" system: technical
  - Phase asset deployment over time
  - Alter operating scale (expand or reduce capacity)
  - Switch input/output

Criteria for "Good" Flexibility

1. Identify major uncertainty source(s) affecting anticipated performance
2. Suggest relevant flexible strategies to deal with uncertainties
3. Identify early on appropriate design variables and parameters enabling flexibility
   - E.g. stronger structure to support expansion, additional piece of land, legal/financial/contractual arrangements if necessary
4. Identify relevant management decision rule to "exercise" flexibility
   - "Trigger" based on some observations
   - E.g. if demand > capacity for 2 years, expand
   - Price threshold, regulatory change, etc

Examples from Industry

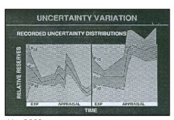
- Satellite communication network (de Weck et al., 2004)
  - '90s award-winning system, $4BN development cost
  - Wrong market forecasts: did not plan for land cell phones. Led to over-capacity design!
  - Sold for $25M in bankruptcy
- Expected cost saved <u>if used flexible phased deployment</u> strategy: $1.8BN (~30%)


Individual Globalstar Satellite


Individual Iridium Satellite

Examples from Industry (cont.)

- Offshore oil platform (Lin, 2009)
  - Multi-billion project off coast of Angola
  - Typical design to "most likely" oil reserves estimate
  - Production capacity expansion flex.: connect more sub sea tiebacks as more oil discovered
- Study shows 80% expected NPV improvement!



Lin, 2009


http://www.offshore-technology.com

Sub sea tiebacks


http://msnmoney.brand.edgar-online.com

http://www.w-industries.com

Take Aways

- Uncertainty has downsides… BUT ALSO PROVIDES UPSIDE OPPORTUNITIES!
- Flexibility in design and management = best approach to deal with uncertainty
- Flexibility helps harvest EXTRA value from uncertainty
- Not easy however to identify and value it…

Appendix B – Prompts for Ideation Mechanism

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

Appendix C – Slides for Design Problem Description

A Multi-Family Residential Development Project

Michel-Alexandre Cardin, PhD Candidate
Spring 2010

MIT ESD
Massachusetts Institute of Technology
Engineering Systems Division

TU Delft
Delft University of Technology

Introduction

- Thanks for being here!
- Personal background
- About design experiments
  - Test under different contexts
  - Might wonder: "why this procedure?"
  - Can't answer all questions… unfortunately!

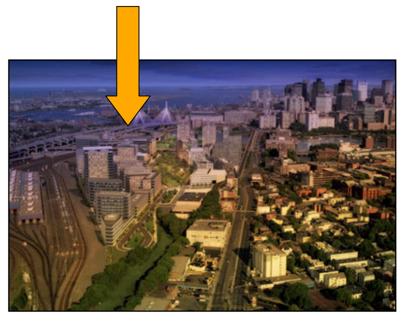
MIT ESD Massachusetts Institute of Technology Engineering Systems Division

2

TU Delft Delft University of Technology

Example Project: Cambridge, MA

http://www.northpointcambridge.com



http://www.northpointcambridge.com

MIT ESD Massachusetts Institute of Technology Engineering Systems Division

3

TU Delft Delft University of Technology

Setup

- You are lead design team at renown real estate development firm
- Specialize in multi-family residential real estate
  - Condo and apartment buildings
- Firm's objective: sell building at highest profit
  - Performance metric: Net Present Value (NPV)
- Land already bought
- Zoning obtained for either condos and/or apartments
- Building permit for 310 units maximum over 3 years

MIT ESD Massachusetts Institute of Technology Engineering Systems Division

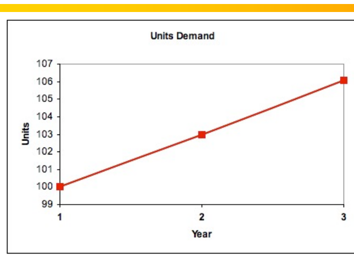
4

TU Delft Delft University of Technology

Data and Assumptions

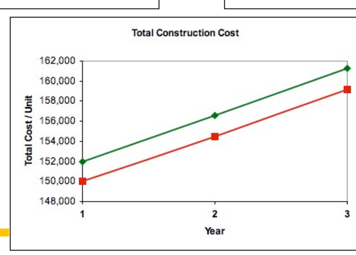
- Current demand expectations:
  - 100 condo units
  - 100 apartment units
- Current selling price expectations:
  - $205,000/unit as condo
  - $200,000/unit as apartment
- Current construction cost estimate per unit (not inclusive of land)
  - $152,000/unit as condo
  - $150,000/unit as apartment
- Demand, price, and construction cost all projected to increase linearly at 3%/year with annual volatility ~20%

Market Projections



LIGHT GREEN: condo

DARK RED: apart.



Analysis and Suggested Design

- Go with condo building design
  - NPV $12.4M > $11.6M (8% discounting)
- All 309 units developed in phase 1 (incl. 3% growth)
- Sold over 3 years (1 year/phase)

NPV as Condo Project

		Phase 1	Phase 2	Phase 3
Year	0	1	2	3
Sales Price/Unit		205,000	211,150	217,485
Units Demand		100	103	106
Constr & Sales/Unit		152,000	156,560	161,257
Planned Capacity Deployment		309	0	0
Units Sold		100	103	106
Sales Revenue		20,500,000	21,748,450	23,053,357
Total Constr & Sales Costs		46,968,000	0	0
Net Cash Flow		-26,468,000	21,748,450	23,053,357
PV of Cash Flow		-24,507,407	18,645,790	18,300,498
NPV (exclu land)	12,438,881			

NPV as Apartment Project

		Phase 1	Phase 2	Phase 3
Year	0	1	2	3
Sales Price/Unit		200,000	206,000	212,180
Units Demand		100	103	106
Constr & Sales/Unit		150,000	154,500	159,135
Planned Capacity Deployment		309	0	0
Units Sold		100	103	106
Sales Revenue		20,000,000	21,218,000	22,491,080
Total Constr & Sales Costs		46,350,000	0	0
Net Cash Flow		-26,350,000	21,218,000	22,491,080
PV of Cash Flow		-24,398,148	18,191,015	17,854,144
NPV (exclu land)	11,647,011			

MIT ESD Massachusetts Institute of Technology Engineering Systems Division — 7 — TU Delft Delft University of Technology — MIT

The Situation

- Management wants design offering best expected (or "average") future performance over range of possible scenarios
- Not convinced suggested design is best
  - Market demand and price may change; differ from projections
  - Costs may increase, construction delayed
- Your team is asked to investigate alternative design(s) that can improve expected performance compared to current design

MIT ESD Massachusetts Institute of Technology Engineering Systems Division — 8 — TU Delft Delft University of Technology — MIT

Appendix D – Assumptions for Discounted Cash Flow Model

Multi-Family Residential Development Project: All input parameters on this page… Assume zero time-to-build

Years per Phase	1.00			
Construction Cost Expected Growth Rate (each phase) (g_{CC})	3%			
Construction Cost Expectations (as of completion, each phase):				
	Phase 1	Phase 2	Phase 3	
Base Cost (BC_t)	130,000	133,900	137,917	*per unit*
Added Cost Expectation to Finish & Sell:				
As Condos (FC_{Ct})	22,000	22,660	23,340	*per unit*
As Apts (FC_{At})	20,000	20,600	21,218	*per unit*
Total Construction and Sales Cost Expectations				
As Condos (CC_{Ct})	152,000	156,560	161,257	*per unit*
As Apts (CC_{At})	150,000	154,500	159,135	*per unit*
Land Cost Expectation				
Up-front Cost Expectation to Enable Switching Flexibility (C_{switch})	0	*per unit*		
Up-front Cost Expectation to Enable Expansion Flexibility (C_{expand})	0	*per unit*		
Selling Price Expected Growth Rate (each phase) (g_P)	3%			
Selling Price Expectations:				
	Phase 1	Phase 2	Phase 3	
As Condos (P_{Ct})	205,000	211,150	217,485	*per unit*
As Apts (P_{At})	200,000	206,000	212,180	*per unit*
Units Demand Expected Growth Rate (each phase) (g_D)	3%			
Units Demand Expectations (each phase)	Phase 1	Phase 2	Phase 3	
As Condos (D_{Ct})	100	103	106	
As Apts (D_{At})	100	103	106	
Planned capacity deployment strategy (each phase)	Phase 1	Phase 2	Phase 3	
As Condos (K_{Ct})	309	0	0	
As Apts (K_{At})	309	0	0	
Maximum capacity allowed (K_{max})	309			
OCC Built Property (r_P)	8.00%	<== Note: Should include spec premium.		
OCC Construction Costs (r_C)	8.00%	<== Note: Should be near riskfree rate.		
*Resulting Canonical OCC (r)**	8.00%	<==Note: Devlpt project contains operatior		
	*Blended rate over all phases and uses (approximation of G			
Overall Volatility Factor (σ)	20%			
Uncertainty Factor in Cost (σ_{CC}):	Volatility factor around expectated growth rate			
As Condos	20%			
As Apts	20%			
	Uncertainty factor around initial cost value			
As Condos	20%			
As Apts	20%			
Uncertainty Factor in Price (σ_P):	Volatility factor around expectated growth rate			
As Condos	20%			
As Apts	20%			
	Uncertainty factor around initial selling price			
As Condos	20%			
As Apts	20%			
Uncertainty Factor in Units Demand (σ_D)	Volatility factor around expectated growth rate			
As Condos	20%			
As Apts	20%			
	Uncertainty factor around initial units demand			
As Condos	20%			
As Apts	20%			

Static Benchmark Model

The table above presents the numerical assumptions used in the DCF model, which were not shown in details to participants. NPV is the objective function used to evaluate the static benchmark model. The condo-only design is used as example to demonstrate equations and modeling assumptions. The same analysis is completed for the apartment-only design, although not presented here. The static NPV is calculated as:

$$NPV_C = \sum_{t=0}^{T} \frac{R_{Ct} - C_{Ct}}{(1+r)^t}$$

Here R_{Ct} represents the revenue at time $t = 0, 1, \ldots, T$, with $T = 3$ years, while C_{Ct} represents the total construction and sales cost at time t. Parameter r represents the canonical Opportunity Cost of Capital (OCC), or the discount rate used to discount cash flows in both cases (r = 8%). The spreadsheet below supports presentation of the NPV calculation using the formula above.

NPV as Condo Project				
		Phase 1	Phase 2	Phase 3
Year	0	1	2	3
Sales Price/Unit		205,000	211,150	217,485
Units Demand		100	103	106
Constr & Sales/Unit		152,000	156,560	161,257
Planned Capacity Deployment		309	0	0
Units Sold		100	103	106
Sales Revenue		20,500,000	21,748,450	23,053,357
Total Constr & Sales Costs		46,968,000	0	0
Net Cash Flow		-26,468,000	21,748,450	23,053,357
PV of Cash Flow		-24,507,407	18,645,790	18,300,498
NPV (exclu land)	12,438,881			

The rows "*Sales Price/Unit*", "*Units Demand*", "*Constr & Sales/Unit*", and "*Planned Capacity Deployment*" present the vectors for price, demand, construction cost, and planned capacity deployment outcomes respectively for the condo-only design:

$$\boldsymbol{P_C} = [P_{C1}, P_{C2}, P_{C3}],$$

$$\boldsymbol{D_C} = [D_{C1}, D_{C2}, D_{C3}]$$

$$\boldsymbol{CC_C} = [CC_{C1}, CC_{C2}, CC_{C3}]$$

$$\boldsymbol{K_C} = [K_{C1}, K_{C2}, K_{C3}]$$

Here the price in each year/phase is calculated as $P_{Ct} = P_{C1}(1 + g_P)^t$, with $g_P = 3\%$ as the projected annual growth rate for sales price, and $P_{C1} = 205{,}000$. Demand and construction costs are calculated similarly using projected demand and construction cost growth rates g_D and g_{CC}, as well as initial projected demand D_{C1} and cost CC_{C1}. Construction cost CC_{Ct} can be further divided as the base construction cost (BC_t) plus the finishing cost (FC_{Ct}) in each year. Planned capacity deployment K_{Ct} is determined by the analyst. The base case model assumes that $\boldsymbol{K_C} = [309, 0, 0]$, meaning that all condo units are developed in year 1. Other deployment strategies can be explored. The row "*Units Sold*" shows the number of units sold, assumed to follow demand projections.

Subsequent rows calculate the cash flows and present value cash flows for the project. The vectors for "*Sales Revenue*" and "*Total Constr & Sales Costs*" correspond to the revenues and costs outcomes used in the NPV calculation above. In vector form, they are represented as:

$$\boldsymbol{R_C} = [R_{C1}, R_{C2}, R_{C3}],$$
$$\boldsymbol{C_C} = [C_{C1}, C_{C2}, C_{C3}]$$

The sales revenue (R_{Ct}), total construction and sales cost (C_{Ct}), net cash flow (CF_{Ct}), and present value of cash flow (PV_{Ct}) at each year/time t are calculated as:

$$R_{Ct} = MIN(D_{Ct}, K_{Ct})P_{Ct}$$
$$C_{Ct} = K_{Ct}CC_{Ct}$$
$$CF_{Ct} = R_{Ct} - C_{Ct}$$
$$PV_{Ct} = CF_{Ct}/(1 + r)^t$$

The sales revenue in each year is the minimum between the demand and the offered capacity. In the static case, capacity is deployed in year 1 to fit demand projections exactly. In the stochastic case described below however, the *MIN*(*) function ensures that revenues do not go beyond available capacity when demand is higher than offered capacity. The NPV is calculated by summing PV_{Ct} over the T years of the project, as shown in the topmost NPV equation.

Stochastic Model

In the stochastic case, price, demand, and construction cost variables P_{Ct}, D_{Ct}, and CC_{Ct} are modeled as random variables. The stochastic evolution of these random variables over time is modeled using GBM. The equations above remain the same, except that price, demand, and construction cost vectors are replaced by the stochastic vectors below:

$$\boldsymbol{P_C}^S = [P_{C1}{}^S, P_{C2}{}^S, P_{C3}{}^S],$$
$$\boldsymbol{D_C}^S = [D_{C1}{}^S, D_{C2}{}^S, D_{C3}{}^S]$$
$$\boldsymbol{CC_C}^S = [CC_{C1}{}^S, CC_{C2}{}^S, CC_{C3}{}^S]$$

For example, $P_{Ct}{}^S = P_{C1}(1 + g_{Pt}{}^S)^t$ where in this case the growth parameter is modeled according to the standard Itô process:

$$g_{Pt}{}^S = g_P dt + \sigma_P dZ_t$$

The parameter σ_P represents the uncertainty factor around annual price projections, and dt is a small time increment of one period (here g_P = 3%, dt = 1 year, σ_P = 20%). The random variable $dZ_t \sim U(-1, 1)$ is the standard Wiener process modeling the stochastic error around the trend growth rate. For simplification and computational efficiency in Excel®, dZ_t is sampled from a uniform distribution between –1 and 1. This random variable is however typically sampled from a standard normal distribution $\sim N(0, 1)$ to account for the fact that larger deviations from the trend are typically less likely. The growth parameters for demand and construction costs $g_D{}^S$ and $g_{CC}{}^S$ are modeled in a similar fashion, also assuming GBM.

In the stochastic model, it is assumed that units unsold in the previous phase can be sold in the next phase. Also, the objective function becomes the mean NPV response over M simulated NPV outcomes, which is the typical estimator for sample means:

$$ENPV_C = E\left[NPV_C\right] = \frac{1}{M}\sum_{m=1}^{M} NPV_{Cm}$$

In this thesis, M = 2,000 to balance statistical significance, and computational efficiency. Two thousand scenarios run in about 1-2 seconds on an Apple MacBook running Excel® 2004 under Mac OS X 10.5.8, 2.4 GHz Intel Core 2 Duo processing power, and 4GB Random Access Memory. From these simulations, one can calculate the standard deviation, minimum, maximum, and percentile values (e.g. P5 and P95) of the distribution. Similarly, one can represent the distribution of NPV outcomes as a CDF, as shown in Figure 6.3.

Flexible Model

Similar equations as those used for the stochastic case govern the case with flexibility. The main difference is the introduction of programmatic decision rules, modeled using *IF*(*) statements in Excel®. Each row in the figure below is visited in turn to give an example of such implementation.

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			

The row "*Next Phase Developed As:*" specifies the type of unit developed next phase, either condo or apartment. An example decision rule may be "*if observed cash flows this year are higher for condo than for apartments, then develop next phase as condo, else develop next phase as apartments*". This decision rules compares the observed cash flows if the phase is developed as condos to the case where it is developed as apartments, using the inflexible stochastic model described previously. The rule suggests developing next phase depending on the unit type that generated the highest cash flow in the current phase. This is not the only possible decision rule here. In fact, participants suggested a variety of decision rules, all implemented separately.

The row "*Develop Current Phase?*" determines whether a phase is developed or not. A possible decision rule here can be "*if construction cost per unit is lower than sales price in the current phase, develop, else do not develop*". The row "*Additional Capacity*" determines how many extra units are built within each phase to suit a capacity expansion flexible strategy. "*Total Capacity Added*" accounts for the planned capacity deployment, plus any additional unit added within a phase. The remaining rows of the DCF model are calculated as described in the static and stochastic models above.

Appendix E – Debrief Material and Example Survey Questions

LimeSurvey - Post-experimental Debrief and Survey, MIT http://sk-1.tbm.tudelft.nl/limesurvey/admin/admin.php?action=...

Post-experimental Debrief and Survey, MIT

"Empirical Testing of Conceptual Design Procedures to Improve Expected Future Performance of Infrastructure Systems"

Instructions: please read the following debrief section, and fill in the appropriate survey information.

Debrief

Thank you very much for your participation! You were involved in an experiment testing out different conceptual design procedures to determine what factors contribute best in helping a team of engineers and designers identify opportunities for flexibility in design and management. The goal is ultimately to improve the anticipated future performance of an infrastructure system compared to an initial benchmark design. Therefore, this study investigates whether 1) you were able (or not) to identify major opportunities for flexibility in a simple infrastructure design problem that 2) can or are known to improve the anticipated performance of the system. Your ideas will be implemented using a computer-based economic model to determine which ones are most valuable.

There are many sources of flexibility that can be implemented in this design problem to improve future expected performance compared to the benchmark design provided. The best flexible design alternatives in general are those that 1) reduce exposure to risks if/when downside events occur, and 2) provide opportunities to capture upside benefits – if needed.

Your team performed an experiment involving two sessions. Session 1 controls for any prior knowledge, experience, or ideas you may have about flexibility in infrastructure design when no particular guidance is offered. It simply asks you to improve the anticipated performance of the suggested design by thinking about alternative design configurations for this particular problem. You may or may not offer ideas involving flexibility at this point; this session is entirely based on your prior training. Session 2 measures any difference if one or a combination of the following tools is introduced to help you think about the uncertainties affecting this system, and the kinds of flexibility that could help you deal with them to improve performance.

Session 2 combines one of two possible variants (called "levels") for each of two factors of interest. The first factor focuses on the **type of education received on flexibility in infrastructure design**. One level provides no particular education on this topic, and relies on your prior training. The other level provides a brief lecture on flexibility in infrastructure design giving intuitive stories about various implementations of flexibility. It discusses why flexibility can improve the anticipated future performance of infrastructure systems. It also introduces important concepts to identify and implement flexibility in new infrastructure design projects.

The second factor focuses on a brainstorming mechanism that sustains creativity by **changing your perspective** on the design task by means of provocation. One level essentially turns "off" this feature, so that no change of perspective mechanism is provided. The other level turns "on" this feature by forcing you through different prompts to see different angles of the problem to identify major sources of uncertainty, and valuable sources of flexibility. It guides you through a thought process suggested to analyze flexibility in new infrastructure design problems.

There are 23 questions in this survey

Process Satisfaction - Session 1

The following questions relate to the FIRST SESSION ONLY.

2 [F]Please select ONE ANSWER for all questions below *

Please choose the appropriate response for each item:

	1 - Strongly disagree	2	3	4 - Neutral	5	6	7 - Strongly agree
I feel satisfied with the way the first session was conducted.	○	○	○	○	○	○	○
I appreciated the techniques we used in the first session.	○	○	○	○	○	○	○
I liked the way the first session progressed today.	○	○	○	○	○	○	○
I feel satisfied with the methods we used in the first session.	○	○	○	○	○	○	○
I feel satisfied about the way we carried out activities in the first session.	○	○	○	○	○	○	○

Results Satisfaction - Session 1

The following questions relate to the FIRST SESSION ONLY

3 [F]Please select ONE ANSWER for all questions below *

Please choose the appropriate response for each item:

	1 - Strongly disagree	2	3	4 - Neutral	5	6	7 - Strongly agree
I feel happy with what we achieved in the first session.	○	○	○	○	○	○	○
I feel satisfied with the things we achieved in the first session	○	○	○	○	○	○	○
I am happy with the results of the first session.	○	○	○	○	○	○	○
Our accomplishments in the first session give me a feeling of satisfaction.	○	○	○	○	○	○	○
When the first session was over, I felt satisfied with the results.	○	○	○	○	○	○	○

Quality Assessment - Session 1

The following questions relate to the FIRST SESSION ONLY

4 [F]Please select ONE ANSWER for all questions below *

Please choose the appropriate response for each item:

	1 - Strongly disagree	2	3	4 - Neutral	5	6	7 - Strongly agree
The result of the first session had the required quality.	○	○	○	○	○	○	○
What we achieved in the first session met the goal.	○	○	○	○	○	○	○
We achieved what we intended to achieve in the first session.	○	○	○	○	○	○	○
The result in the first session has the quality we intended to achieve.	○	○	○	○	○	○	○
The result in the first session was in line with the goal set for this workshop.	○	○	○	○	○	○	○
I feel confident the top five (5) design alternatives recommended in the first session provide better anticipated performance than the current benchmark design	○	○	○	○	○	○	○

Appendix F – Example Original Transcript Coding Analysis

4.2.1. Negative or bad scenarios (e.g. start with a smaller initial design, and reduce risk of over-capacity and losses)
4.2.2. Positive or good scenarios (e.g. engineer ability to switch product output easily, write legal contract to enable physical expansion later on if needed)
4.2.3. Completely unexpected scenario (e.g. plan ahead for emergency procedure in case of hurricane)
4.3. Design a show house first to see the market reaction
4.4. Build the condo or apartment over time without put your money in once.
4.5. Try to find reliable customers and build relation with them in order to prevent future downside.
4.6. Build both type of units at the first phase and wait for market response
4.7. Consider design to be robust for 1 week of raw material shortage
4.8. In the case of earthquake or hurrican, one of the solution is to improve building technology to reduce the possible negative impact.
4.9. Engineer the units to be built first on the common elements within both type of units to then to built the differentiators within them
4.10. Enough capital for further construction in a short time.
4.11. Consider flexibilty to incorporate other type of units (i.e. a two floor apartment)
4.12. Try to identify the structure of family or their income to predict possible change in the future.

5. Management

5.1. How should you manage and decide when it is appropriate to use, or exercise, the flexibilities in this system?
5.2. Examples:
5.2.1. If demand is lower than capacity for two years, I will shutdown operations for 6 months
5.2.2. If market price gets above a certain threshold, I will expand production capacity
5.3. Effective leader with more centralized power in decision-making process.
5.4. If raw materials go below desin assumptions, build more units on the earlier phases
5.5. If the cost of input goes down or the revenue goes up to some extent, I will expand production.
5.6. Start with a 50-50 assumption and adapt to latest sales history when deciding on future units
5.7. Consider a risk averse strategy and build more units if demand improves only

6. Recommendations

6.1. Summarizing all ideas above, what design alternative(s) would you recommend to improve future expected performance of this infrastructure system compared to the currently suggested design?
6.2. I'd consider the alternative of delaying the type of units as much as possible, building a flexible design capable of delaying the decision as much as possible
6.3. Predication of future trend in family such as size and income to figure their capacity and accomodate to future needs.
6.4. I'd consider several sources (contractors, raw materials) and delay the decision as much as possible
6.5. Consider pre-sales as a source of minimizing demand uncertainty
6.6. Try to figure out the market demand. For example, survey; build show condo and apartment to see consumers' preference; pre-sell to see the demand
6.7. Engineer to accomodate other facilities (pool, gym, playground) in the complex if the demands improves with these extra features

new 8/1/10 4:45
Comment:
UncSource.-
Strategy.phasing (rigid)
Enabler.management
DecisionRule.-

new 8/1/10 4:46
Comment:
UncSource.market response
Strategy.switching
Enabler.management
DecisionRule.-

new 8/1/10 4:48
Comment:
Strategy.switching
Enabler.design
DecisionRule.-

new 8/1/10 4:52
Comment:
Strategy.phasing
Enabler.management
DecisionRule.if raw materials go below design assumption, build more units on the earlier phases

new 8/1/10 4:52
Comment:
Strategy.capacity expansion
Enabler.management
DecisionRule.if cost of input goes down or the revenue goes up, expand production

new 8/1/10 4:54
Comment:
UncSource.demand
Strategy.phasing
Enabler.management
DecisionRule.use latest sales history for the next phase

new 8/1/10 4:54
Comment:
UncSource.demand
Strategy.phasing
Enabler.management
DecisionRule.if demand improves build more units

Appendix G – Summary of Complete Ideas from Transcript Analysis

Switch between condos and apartments:

- Develop each phase independently and decide in light of how market uncertainty (e.g. construction cost, demand, or price) manifests in a previous phase whether to develop next phase as condo or apartment;
- Decide to have an initial mixture of condos and apartments, and develop remaining units within the same phase as condo or apartment, or;
- Develop units as condo, but sell as apartment if demand or price is too low to sell as condo.

Phase the development planning

- Deploy capacity over time rather than all at once. The number of units in each phase is decided upon based on observations of one of the market uncertainties in the current and/or previous phase.

Expand unit capacity

- Build more units if needed within the same phase in light of realized market uncertainty (e.g. if demand is higher than expected, build and sell more units).

Reduce unit capacity

- Build fewer units than originally planned within the same phase.

Adjust unit capacity "just in time"

- Develop exactly however many units of a given type is demanded, in any given phase.

Temporarily abandon the project

- Not develop a particular phase based on prior observation of some market uncertainty (e.g. if demand is lower than expected in the previous phase, do not develop the subsequent phase).
- This flexibility must clearly consider the possibility to resume development.

Completely abandon the project

- Completely abandon the project based on some manifestation of market uncertainty. This can happen in any phase, including the first and last phases.

Appendix H – Transcript Analysis and Response Measurements Dataset

Expt	Rep	Complete flex. ideas S1	Complete ideas? S1	ENPV attained S1	Good idea? S1	Total ENPV flex.(millions) S1	Total DENPV (millions) S1	Complete NEW flex. Ideas S2	Complete NEW idea? S2	Percent. total complete ideas from S2	ENPV attained S2	NEW good idea? S2	Percent. total good ideas from S2	Total ENPV flex.(millions) S2	Total DENPV (millions) S2	Change total DENPV (millions) S1 to S2	Percentage total DENPV from S2
1	1		NO	9.3	NO	9.3	0.0	switching	YES		10.5	YES		10.5	1.2	1.2	100%
TOTAL			0		0				1	100%		1	100%				
1	2	switching	YES	12.4	YES	12.4	3.1					NO		13.0	3.7	0.6	54%
TOTAL			1		1				0	0%		0	0%				
1	3		NO	9.3	NO	9.3	0.0		NO		9.3	NO		9.3	0.0	0.0	0%
TOTAL			0		0				0	0%		0	0%				
1	4		NO	9.3	NO	9.3	0.0		NO			NO		9.3	0.0	0.0	0%
TOTAL			0		0				0	0%		0	0%				
1	5		NO	9.3	NO	9.3	0.0	switching	YES		10.8	YES		10.8	1.5	1.5	100%
TOTAL			0		0				1	100%		1	100%				
1	6		NO	9.3	NO	9.3	0.0		NO			NO		9.3	0.0	0.0	0%
TOTAL			0		0				0	0%		0	0%				
1	7	capacity "just in time"	YES	12.7	YES	12.7	3.4		NO			NO		12.7	3.4	0.0	0%
TOTAL			1		1				0	0%		0	0%				
1	8		NO	9.3	NO	9.3	0.0		NO			NO		9.3	0.0	0.0	0%
TOTAL			0		0				0	0%		0	0%				
2	1		NO	9.3	NO	9.3	0.0	temp. abandonment	YES		9.7	YES		11.3	2.0	2.0	100%
								capacity expansion	YES		11.3	YES					
TOTAL			0		0				2	100%		2	100%				
2	2	switching	YES	12.3	YES	12.3	3.0	temp. abandonment	YES		11.7	YES		13.2	3.9	0.9	57%
								capacity expansion	YES		11.6	YES					
TOTAL			1		1				2	67%		2	67%				
2	3	switching	YES		NO	14.6	5.3	capacity "just in time"	YES		11.4	YES		14.6	5.3	0.0	50%
TOTAL			1		0				1	50%		1	100%				
2	4		NO	0.0	NO	9.3	0.0	temp. abandonment	YES		10.5	YES		12.8	3.5	3.5	100%
								switching	YES		12.8	YES					
								capacity reduction	YES		0.0	NO					
TOTAL			0		0				3	100%		2	100%				
2	5		NO		NO	9.3	0.0	switching	YES		14.4	YES		14.4	5.1	5.1	100%
TOTAL			0		0				1	100%		1	100%				
3	1		NO		NO	9.3	0.0	phasing	YES		12.3	YES		12.3	3.0	3.0	100%
TOTAL			0		0				1	100%		1	100%				
3	2		NO		NO	9.3	0.0	abandonment	YES		10.2	YES		10.2	0.9	0.9	100%
TOTAL			0		0				1	100%		1	100%				
3	3	capacity reduction	YES	11.3	YES	9.3	0.0	abandonment	YES		10.9	YES		11.3	2.0	2.0	100%
TOTAL			1		1				1	50%		1	50%				
3	4		NO	9.3	NO	9.3	0.0		NO		9.3	NO		9.3	0.0	0.0	0%
TOTAL			0		0				0	0%		0	0%				
3	5		NO	9.3	NO	9.3	0.0	switching	YES		14.7	YES		14.7	5.4	5.4	100%
TOTAL			0		0				1	100%		1	100%				
4	1		NO	9.3	NO	9.3	0.0	capacity expansion	YES		11.0	YES		13.0	3.7	3.7	100%
								phasing/switching	YES		12.6	YES					
TOTAL			0		0				2	100%		2	100%				
4	2		NO	9.3	NO	9.3	0.0	phasing	YES		9.3	NO		11.4	2.1	2.1	100%
								switching	YES		11.4	YES					
								abandonment	YES		9.8	YES					
TOTAL			0		0				3	100%		2	100%				
4	3		NO	9.3	NO	9.3	0.0	abandonment	YES		12.2	YES		12.2	2.9	2.9	100%
TOTAL			0		0				1	0%		1	0%				
4	4		NO	9.3	NO	9.3	0.0	temp. abandonment	YES		11.2	YES		11.4	2.1	2.1	100%
								capacity expansion	YES		11.4	YES					
								switching	YES		9.3	NO					
								phasing	YES		11.1	YES					
TOTAL			0		0				4	100%		3	100%				
4	5	perm. abandonment	YES	10.2	YES	10.2	0.9	capacity "just in time"	YES		11.0	YES		12.2	2.9	2.0	76%
TOTAL			1		1				1	50%		1	50%				
4	6		NO	9.3	NO	9.3	0.0	capacity expansion	YES		11.2	YES		11.2	1.9	1.9	100%
								phasing	YES		11.1	YES					
TOTAL			0		0				2	100%		2	100%				
4	7		NO	9.3	NO	9.3	0.0	switching	YES		15.2	YES		16.0	6.7	6.7	100%
								phasing	YES		12.2	YES					
TOTAL			0		0				2	100%		2	100%				
4	8		NO	9.3	NO	9.3	0.0	capacity expansion	YES		11.0	YES		14.6	5.3	5.3	100%
								switching	YES		14.6	YES					
								temp. abandonment	YES		9.3	NO					
TOTAL			0		0				3	100%		2	100%				

Appendix I – Complementary Response Measurements Datasets

Complete dataset of ΔG measurements for all treatment groups with $x_1 = E$ and $x_2 = I \in \{-1, +1\}$. R_k is the k^{th} replicate of a treatment. Expressed in units of good ideas.

Treatment	E	I	R_1	R_2	R_3	R_4	R_5	R_6	R_7	R_8
1	–1	–1	1	0	0	0	1	0	0	0
2	–1	+1	2	2	1	2	1			
3	+1	–1	1	1	1	0	1			
4	+1	+1	2	2	1	3	1	2	2	2

Complete dataset of $\Delta ENPV$ measurements for all treatment groups with $x_1 = E$ and $x_2 = I \in \{-1, +1\}$. R_k is the k^{th} replicate of a treatment. Expressed in millions.

Treatment	E	I	R_1	R_2	R_3	R_4	R_5	R_6	R_7	R_8
1	–1	–1	1.2	0.6	0.0	0.0	1.5	0.0	0.0	0.0
2	–1	+1	2.0	0.9	0.0	3.5	5.1			
3	+1	–1	3.0	0.9	2.0	0.0	5.4			
4	+1	+1	3.7	2.1	2.9	2.1	2.0	1.9	6.7	5.3

Sample dataset of eight ΔRS measurements for each treatment group with $x_1 = E$ and $x_2 = I \in \{-1, +1\}$. R_k is the k^{th} participant for a treatment. Expressed in points of satisfaction.

Treatment	E	I	R_1	R_2	R_3	R_4	R_5	R_6	R_7	R_8
1	–1	–1	9	25	-3	-5	15	-6	3	0
2	–1	+1	-6	5	3	-9	4	1	5	0
3	+1	–1	-1	9	10	15	9	0	15	5
4	+1	+1	16	18	5	11	15	-5	6	15

Sample dataset of eight Δ*QA* measurements for each treatment group with $x_1 = E$ and $x_2 = I \in \{-1, +1\}$. R_k is the k^{th} participant for a treatment. Expressed in quality points.

Treatment	*E*	*I*	R_1	R_2	R_3	R_4	R_5	R_6	R_7	R_8
1	–1	–1	11	23	-6	-5	8	-4	4	4
2	–1	+1	-3	4	5	-10	3	4	6	0
3	+1	–1	-5	6	12	16	3	0	15	12
4	+1	+1	14	17	10	14	8	-4	8	14

Complete dataset of Δ*FI* measurements for all treatment groups with $x_1 = E$ and $x_2 = I \in \{-1, +1\}$. R_k is the k^{th} replicate of a treatment. Expressed in influence points.

Treatment	*E*	*I*	R_1	R_2	R_3	R_4	R_5	R_6	R_7	R_8
1	–1	–1	0.1	0.1	-0.1	-0.1	0.0	0.0	0.1	-0.2
2	–1	+1	0.2	0.0	-0.1	0.0	-0.1			
3	+1	–1	0.2	0.4	0.3	0.5	0.2			
4	+1	+1	0.0	0.3	0.2	-0.2	0.1	0.0	0.0	0.2

Appendix J – List of Flexibility and Uncertainty Related Words

Word	Flexibility	Uncertainty
accuracy/volatility		✓
alternative	✓	
assumed		✓
assumption		✓
capacity	✓	
capture	✓	
change	✓	
competition		✓
conversion	✓	
convertibility	✓	
cost		✓
decision	✓	
delay	✓	
demand		✓
differentiation	✓	
distribution		✓
downside		✓
drop		✓
efficiency	✓	
environment		✓
estimate		✓
expansion	✓	
exposure		✓
fire		✓
flexibility	✓	
flexible	✓	
fluctuating		✓
fluctuation		✓

future		✓
interchangeable	✓	
management	✓	
market		✓
multi-purpose	✓	
option	✓	
overoptimistic		✓
performance		✓
phase	✓	
phased	✓	
possible		✓
price		✓
probability		✓
profit		✓
projections/marketing		✓
range		✓
return		✓
revenue		✓
scalability	✓	
shell	✓	
statistics		✓
uncertainty		✓
unfinished	✓	
upgrade	✓	
value		✓
volatility		✓
TOTAL	**31**	**23**

Appendix K – Code for Permutation-Based Statistical Analysis

```
% This script calculates main and interaction effects using the regress() function under 5,000 simulated
% randomization/permutation of the dataset to generate a distribution of main and interaction effects.
% This distribution is used to measure the p-value of the observed
% experimental main and interaction effects

% NOTE: using the regress() function produces different B0 coefficient each
% time, which are taken as the grand mean in the model. This is essentially because I do not have
% have the same number of replicates in each group, which offsets the mean estimate. This does not
% affect results however, but is important to note.

clear all; close all;

% Number of trials for permutations
trials=5000; % this number should depend on the total number of permutations that can be made, although evaluating all may require too much co

% Data complete idea delta
complete1 = [1  0       0       0       1       0       0       0]; % E-I-
complete2 = [2  2       1       3       1];             % E-I+
complete3 = [1  1       1       0       1];             % E+I-
complete4 = [2  3       1       4       1       2       2       3]; % E+I+

% Data good idea delta
good1 = [1      0       0       0       1       0       0       0];
good2 = [2      2       1       2       1];
good3 = [1      1       1       0       1];
good4 = [2      2       1       3       1       2       2       2];

% Data absolute ENPV delta
ENPV1 = [1.2 0.6 0.0 0.0 1.5 0.0 0.0 0.0];
ENPV2 = [2.0 0.9 0.0 3.5 5.1];
ENPV3 = [3.0 0.9 2.0 0.0 5.4];
ENPV4 = [3.7 2.1 2.9 2.1 2.0 1.9 6.7 5.3];

% Choose what data to compare.
% INPUT NEEDED HERE
data1 = ENPV1;
data2 = ENPV2;
data3 = ENPV3;
data4 = ENPV4;

group1 = ones(1,length(data1));
group2 = 2*ones(1,length(data2));
group3 = 3*ones(1,length(data3));
group4 = 4*ones(1,length(data4));
data =  cat(2, data1, data2, data3, data4); % contains *all* experimental data
group = cat(2,group1, group2, group3, group4);

% Compute relevant terms to calculate main and interaction effects,
% assuming the linear regression model
% Y = B0 + B1X1 + B2X2 + B12X1X2
% Y is the response of interest (either complete ideas, good ideas, or ENPV)
% Output values from the regression:
% B0 is interpreted as the grand mean of the model
% B1 is the main effect of factor 1, here the education factor (E)
% B2 is the main effect of factor 2, here the ideation factor (I)
% B12 is the interaction effect of factors 1 and 2, the education and ideation factors
% Input values to the regression:
% X1 can take on values -1 for factor level E-, and +1 for factor level E+
% X2 can take on values -1 for factor level I-, and +1 for factor level I+
% X1X2 simply multiplies the values of X1 and X2 to model the independent variable of the interaction effect

% Matrix X encodes the values of the independent variables, or factor
% levels, associated with each response in the vector "data"
% Note the last column is made of constant "1" terms to stabilize the
% regression model in Matlab
% The first column is factor E, and responses in vector "data" are
% associated to -1 if produced under factor level E-, or +1 if produced
% under factor level E+
% See spreadsheet "summaryData.xls" under tab 'regression'
X1 =    [-1; -1; -1; -1; -1; -1; -1; -1; -1; -1; -1; -1; -1; +1; +1; +1; +1; +1; +1; +1; +1; +1; +1; +1; +1; +1];
X2 =    [-1; -1; -1; -1; -1; -1; -1; -1; +1; +1; +1; +1; +1; -1; -1; -1; -1; -1; +1; +1; +1; +1; +1; +1; +1; +1];
X1X2 =  X1.*X2;
constant = ones(26,1);
X = [X1 X2 X1X2 constant];

% Compute the "observed" main and interaction effects for the existing dataset
b = regress(data',X);

obs_E = b(1); % main effect for E
obs_I = b(2); % main effect for I
obs_EI = b(3); % interaction effect for E and I
obs_M = b(4); % grand total mean

% The idea below is to shuffle the whole data set and calculate the main and interaction effects every
% time it is reshuffled using the regression model. This then creates a distribution
% of main and interaction effects, and therefore we can evaluate the
% likelihood of the observed main and interaction effects with the existing
% dataset. This gives us the significance level of these main and
% interaction effects. The underlying assumptions are that if sampling was truly random, and treatments have no effect, than
% randomly permuting the data at any position in the dataset would yield
% treatment results that show no difference. Hence, the main and interaction effects should be zero
% under the null hypothesis, and the observed main and interaction effects
% would be the same under any re-arrangement of the dataset.

% Simulate the distribution of statistics of interest.
% For each trial, associate any of the data from the entire dataset to any of
% the four groups by permuting randomly the order sequence. Calculate the
% statistics of interest (e.g. main and interaction effects). Record the statistic. This produces a
```

```
% distribution of the statistic of interest arising from all possible permutations
% evaluated, with which we compare the likelihood of the observed mean
% difference to produce the one tail and two-tail p-values below
for i=1:trials
    % Shuffle randomly the position of each data, and re-arrange the
    % dataset
    r=randperm(length(group));
    rdata = data(r);

    % Calculate and record the statistics of interest with the newly shuffled dataset
    % (prefix "r" for randomized, reshuffled)
    r_b = regress(rdata',X);

    r_E(i)  = r_b(1); % main effect for E
    r_I(i)  = r_b(2); % main effect for I
    r_EI(i) = r_b(3); % interaction effect for E and I
    r_M(i)  = r_b(4); % grand total mean
end

% Calculate standard deviation of each distribution, and see if standard error approximation above is close
std_r_M = std(r_M);

std_r_E = std(r_E);
std_r_I = std(r_I);
std_r_EI = std(r_EI);
std_r_M = std(r_M);

% Choose effect to show on histogram.
% Grand total mean is obs_M
% Main effects are obs_E, obs_I
% Interaction effect is obs_EI

obs_eff = obs_E; % INPUT NEEDED HERE!

switch obs_eff
    % Main effects
    case obs_E
        r_eff = r_E;
    case obs_I
        r_eff = r_I;

    % Interaction effect
    case obs_EI
        r_eff = r_EI;

    % Grand total mean
    case obs_M
        r_eff = r_M;
end

% Show distribution of main and interaction effects on histogram
n_bins = 100; % number of bins. The more bins, the more resolution
w_bins = (max(r_eff)-min(r_eff))/n_bins;
bins = [min(r_eff):w_bins:max(r_eff)];
[r_eff_hist, r_eff_out] = hist(r_eff, bins);
hist(r_eff, bins); % show the actual plot in figure
hold on
plot([obs_eff obs_eff], [0 max(r_eff_hist)],'--+r'); % plot location of main effect statistic as dashed line

switch obs_eff
    % Main effects labels
    case obs_E
        xlabel('Simulated main effect coefficients B1 for factor x1');
    case obs_I
        xlabel('Simulated main effect coefficients B2 for factor x2');

    % Interation effect label
    case obs_EI
        xlabel('Simulated interaction effect coefficients B12 between factors x1 and x2');
    % Main effect label
    case obs_M
        xlabel('Simulated grand total mean');
end
ylabel('Count');

% Calculate one-tail and two-tail p-values. If test statistic is negative,
% look to the left of distribution for one-tail p-value. If positive, look
% to the right. Two-tail p-value is sum of areas to the left and right of
% distribution using +/- test statistic value
if obs_eff >= 0
        obs_eff % show statistic of interest
        p_one = mean(r_eff>obs_eff) % calculate and show the one-tail p-value using the right of the distribution
        p_two = mean(r_eff>obs_eff) + mean(r_eff<-obs_eff) % calculate and show the two-tail p-value by considering the area to the left of th
end
if obs_eff < 0
        obs_eff % show statistic of interest
        p_one = mean(r_eff<obs_eff) % calculate and show the one-tail p-value using the left of the distribution
        p_two = mean(r_eff<obs_eff) + mean(r_eff>-obs_eff) % calculate and show the two-tail p-value by considering the area to the right of t
end
```