INTERNATIONAL CONFERENCE ON ENGINEERING DESIGN, ICED'07
28 - 31 AUGUST 2007, CITE DES SCIENCES ET DE L'INDUSTRIE, PARIS, FRANCE

USING DESIGN OF EXPERIMENTS (DOE) FOR DECISION ANALYSIS

Victor Tang[1], Kevin N. Otto[2], and Warren P. Seering[1]
1 Massachusetts Institute of Technology, USA
2 Robust Systems and Strategy USA, and Massachusetts Institute of Technology USA

ABSTRACT
We take an engineering design approach to a problem of the artificial - corporate decision-analysis under uncertainty. We use Design of Experiments (DOE) to understand the behaviour of systems within which decisions are made and to estimate the consequences of alternative decisions. The experiments are a systematically constructed class of *gedanken* (thought) experiments comparable to "what if" studies, but organized to span the entire space of controllable and uncontrollable options. We therefore develop a debiasing protocol to forecast and elicit data. We consider the composite organization, their knowledge, data bases, formal and informal procedures as a measurement system. We use Gage theory from Measurement System Analysis (MSA) to analyze the quality of the data, the measurement system, and its results. We report on an *in situ* company experiment. Results support the statistical validity and managerial efficacy of our method. Method-evaluation criteria also indicate the validity of our method. Surprisingly, the experiments result in representations of near-decomposable systems. This suggests that executives scale corporate problems for analyses and decision-making. This work introduces DOE and MSA to the management sciences and shows how it can be effective to executive decision making.

Keywords: Decision Analysis, Design of Experiments, Gage R&R, Complex Systems, Business Process

1 INTRODUCTION
This article is about a new idea: we can study corporate problems and their potential solutions under uncertainty using engineering methods; specifically, DOE [1] and MSA [2]. These are proven methods in engineering and manufacturing, but are absent in management decisions. Our hypothesis is that as in an engineering system, corporate problems and their potential outcomes depend on the behaviour of organizational systems under uncertainty, and these systems can be studied with experiments (real or simulated). We also consider decisions as intellectual artefacts than can be designed, evaluated, and their outcomes predicted using engineering methods.

DOE presents us with a method to understand the behaviour of the corporate systems within which decisions are made and to estimate the consequences of alternative choices as scenarios. The experiments are a set of systematically designed *gedanken* experiments structured to span the entire space of controllable and uncontrollable options.

In any experiment, data quality depends on instruments and how they are used. We consider the composite of the organization, their knowledge, data bases, formal and informal procedures as a measurement system. Gage theory from MSA presents us with an engineering method to uncover weaknesses that contribute to low-quality data, and take corrective action.

Executives jealously guard decision-making as a power-reserved. Our objective is not to make decisions. Rather, it is to provide a more complete and systematic analysis than is currently practical and provide the results of this analysis to corporate leaders in a form that is particularly useful to them.

2 LITERATURE SURVEY
Decision theory is an interdisciplinary field of study to understand, improve, and predict the outcomes of decisions under uncertainty. It draws from systems analysis, mathematics, economics, psychology, and management. Scholars identify three research streams: the normative, descriptive, and prescriptive streams. We follow Keeney [3] and summarize the three strands in Table 1.

Table 1. Summary of normative, descriptive, and prescriptive theories

	normative	descriptive	prescriptive
focus	how people should decide with logical consistency	how and why people decide the way they do	prepare people to decide and help them make better decisions
criterion	theoretical adequacy	empirical validity	efficacy and usefulness
scope	all decisions	classes of decisions tested	specific decision problems
theoretical foundations	axioms of utility theory	psychology of beliefs and preferences	logic, normative and descriptive theories
operational focus	analysis of alternatives. determine preferences.	prevent systematic errors in decision-making	processes, procedures. end-end decision life-cycle.
judges	"theoretical sages"	experimental researchers	applied analysts

Normative theory is deals with the logical consistency of decision-making. A person's choices are *rational,* when their behaviour is consistent with the normative axioms of expected utility theory (EU) [4]. These axioms establish ideal standards for rationality. Though elegant, normative theory is not without paradoxes or inconsistencies [5], [6]. Moreover, perfect rationality far exceeds people's cognitive capabilities; therefore, they *satisfice* and do not maximize [7]. Simon [7] in his seminal work on organizational behaviour argues that bounded rationality is the fundamental operating mechanism in decision-making. We adopt this perspective of bounded rationality for our work.

Descriptive theory concentrates on representations of how and why people make the decisions they do. For example, Prospect Theory posits that we think of value as changes in gains or losses [8]. In Social Judgment Theory, the decision maker aggregates cues and correlates them against the environment. Naturalistic Decision Making opts for descriptive realism. For example, Klein [9] studies contextually-complex decisions characterized by urgency, volatile and risky conditions such as combat. His work reveals that these decision-makers rely on a few factors and mental simulations that can be completed in a limited number of steps. We note that corporate decisions are also characterized by urgency, volatile and risky conditions, but with insufficient simulations.

Prescriptive Decision Theory is about the practical application of normative and descriptive theories. Decision analysis is the discipline that seeks to help people and organizations make more insightful decisions and act more intelligently under uncertainties. Decision analysis includes the design of alternative choices, i.e. the task of "…logical balancing of the factors that influence a decision … these factors might be technical, economic, environmental, or competitive…" [10]. Decision analysis is boundedly rational. "There is no such thing as a final or complete analysis; there is only an economic analysis given the resources available [11]." We illustrate the diversity of decision analysis with a sample of four prescriptive theories (Table 2).

Table 2. Summary of four prescriptive theories

	Utility Theory	Imprecision	Real Options	AHP
preference basis	utility	preference	monetary value	importance
units	utils	preference	monetary units	unitless
foundations	Subjective expected utility (SEU).	Fuzzy sets and trade-off functions.	Temporal resolution of uncertainty.	Scales for pairwise comparisons.
principles	Normative axioms.	Trade-offs are not additive.	Sequential temporal flexibility.	Linear ordering by importance.
distinctive processes /analyses	Decision representation. Utility function.	Preference mapping for improved insight.	Options: abandon, stage, defer, grow, scale, switch.	Factors hierarchy. Analysis of pairwise comparisons.

Five examples of prescriptive methods are: Ron Howard's method of decision analysis [10] and Keeney's Value Focused Thinking (VFT) [12] both of which use utility theory; Otto and Antonsson's method of imprecision [13], real options [14], and Analytic Hierarchy Process (AHP) [15]. Keefer, Kirkwood, and Corner [16] present a survey of decision analysis. We position our work in this paper as a prescriptive method (Table 3, next page). We present the highlights of our method in juxtaposition with Table 2. The remainder of this article is devoted to the explanation of Table 3.

Table 3. Summary of our DOE-based method

preference basis	▪ more of output is better, or less is better, or require exact specified output
units	▪ natural units specific to the decision situation
foundations	▪ bounded rationality ▪ design of experiments (DOE), Gage R&R ▪ research on bias from descriptive decision theory
principles	▪ unconstrained exploration of entire solution space ▪ unconstrained exploration of entire space of uncertainty
distinctive processes	▪ debiasing of elicited data ▪ determining the quality of the input data ▪ construction of decision alternatives

3 HYPOTHESIS AND RESEARCH METHODS

3.1 Hypothesis and research question

We take an engineering design approach to investigate corporate decisions and their outcomes under uncertainty. Our hypothesis is that as in an engineering system, corporate decisions and their potential consequences depend on the behaviour of business systems, which can be studied with experiments (real or simulated). We also consider decisions as intellectual artefacts that can be designed and tested using engineering methods. The research questions are: Is there support to indicate the efficacy and validity of such an approach? What can we learn about the systems that underpin decisions?

3.2 Protocol for experiments

The canonical model for decision making

The "canonical paradigm" [18] for decision making posits seven steps: (1) recognize a decision is needed, (2) define the problem or opportunity, (3) specify goals and objectives, (4) generate alternatives, (5) analyze alternatives, (6) select an alternative, and (7) learn about the decision. Simon notes: "The classical view of rationality provides no explanation of where alternate courses of action originate; it simply presents them as a free gift to the decision makers" [19]. Analysis has crowded out synthesis. We concentrate on step (4) in order to fill this gap.

Data-collection and forecasting protocol

All decisions are based on forecasts about outcomes and preferences for those outcomes. Forecasts are subject to bias [6]. Overconfidence is one of the most pernicious biases in decision making [20]. It is therefore surprising that there is "... little evidence that debiasing techniques are frequently employed in actual practice [21]". But how do you debias? Scholars suggest:

- ***Counter-argumentation.*** This process requires the explicit articulation of the reasons why a forecast might be correct and also incorrect [22], [23]. Disconfirmatory information has a debiasing effect that enriches people's mental models about the decision that improves their ability to conceptualize alternatives [24]. Therefore, our forecasting protocol that includes counter-argumentation.
- ***Anti-herding.*** Herding refers to people's tendency to succumb to social pressures and produce forecasts that cluster together [25]. To avoid herding, our protocol forbids the disclosure of individual forecasts, but it encourages counter-argumentation. This is our non disclosure rule.
- ***Accountability.*** Accountability is an important factor in reducing bias [26], particularly when it is known before judgements are reached. Accountability can attenuate the bias of overconfidence and improve the accuracy of forecasting.

We embody these principles and our non-disclosure rule into our protocol. This process promotes more critical systems thinking that also diminishes information asymmetry among the team members engaged in the process. They are the hallmarks of our protocol.

Generation and analysis of decision alternatives

We use DOE procedures to construct decision alternatives and to predict outcomes under uncertainty. To construct a decision alternative, one simply specifies an appropriate treatment. This way, we can explore the entire solution space and answer any "what if" questions decision makers wish to pose.

Experiment Validation
We follow Yin [27] and Hoyle, Harris and Judd [28] and subject our experiments to tests of construct, internal, and external validity.

Method validation
Carroll and Johnson [29] specify six criteria for the evaluation of a method. We use these six criteria to evaluate our DOE-based decision analysis method.

4 INDUSTRIAL ILLUSTRATION: *IN SITU* EXPERIMENT

Prior to these experiments, we performed extensive simulations of our method using a comprehensive (>600 equations) system dynamics model of a real company. Then we performed two *in situ* experiments, one in the US and the other Japan. Due to space limitations, we present the American experiment in this article. All experiments and simulations can be found in Tang [30].

4.1 Experiment with a US manufacturing company

We will call the company High-Tech Electronics Manufacturing. HiTEM is a contract manufacturer. It has plants in the US, Asia, and Europe. Adapting the canonical paradigm, we specified and used the experimental protocol below (Table 4) for this company experiment.

Table 4. Experimental protocol

1.	Framing the problem	▪ Understand the decision situation, goals and objectives. ▪ Specify the problem in DOE normal form
2.	Establish the base line	▪ Forecast the business-as-usual (BAU) case
3.	Forecast the sample space	▪ Forecast the sample space in three uncontrollable environments
4.	Analyze the data	▪ Analyze summary statistics and test treatments ▪ Analyze gage R&R statistics
5.	Analyze alternatives	▪ Construct and analyze alternatives
6.	Learning from the decision	▪ Summarize findings and lessons from the experiment ▪ Analyze validity of the experiment and the decision's quality.

4.2 Framing the Problem

The Decision Situation
HiTEM has not made a profit in three years. The newly appointed president must turn a profit in six months. He wanted to know what strategic alternatives, in addition to his own, were possible. He appointed a five-person task force to work with us. Task force members were from manufacturing, marketing, finance, and operations.

Framing decision in our DOE normal form
The "problem" and "outcomes" (Table 5) have already been addressed. The "controllable variables", SG&A, COGS, and sales are the usual expense, cost, and revenue items. The plan was to alter the customer-portfolio mix by shedding customers that do not contribute a designated level of profit. "Financing" meant selling unprofitable plants in Mexico or China for a one-time cash flow.

Table 5. Framing of HiTEM's decision situation in DOE normal form

problem	survival		
outcomes	profitability in 6 months		
controllable variables	. 1. SG&A . 2. COGS	. 3. Capacity Utilization . 4. customer portfolio mix	.5. Sales .6. Financing
uncontrollable variables	. 1. change in demand . 2. senior executive interactions	3. banker actions 4. loss of critical skills	

The next step was to bracket the limits of the controllable variables (Table 6). "Level 3" was specified as doable, but only with a very strong effort. "Level 1" was the lowest acceptable-level of managerial performance. The * entries represent current level of operations, i.e. "business-as-usual" (BAU).

Table 6. Controllable variables and levels

controllable	level 1	level 2	level 3
SG&A	\$54 M + 10%	\$54 M *	\$54 M - 10%
COGS	\$651 M + 2%	\$651 M *	\$651 M – 2%
plant capacity	40% utilization	60% utilization *	80% utilization
customer portfolio mix	No change. Retain current customer mix *	dev. < 10%, A&T < 6%, manufacturing < 4%	dev. < 20%, A&T < 12%, manufacturing < 8%
sales	\$690 M – 5%	\$690 M *	\$690 M + 5%
financing	cash shortfall * of \$10 M annualized	Divest Mexico plant. yields \$12M annualized	Divest China plant. yields \$25M annualized

* BAU

Uncontrollable variables are those management cannot, or are very costly to control, but have a direct impact on the desired outcome (Table 7). "Level 3" is the best, but realistic, condition. "Level 2" is the current condition, denoted by *. "Level 1" is the worst, but possible, uncontrollable condition. To determine the limits in Table 6 and 7, team members were free to consult with their staffs.

Table 7. Uncontrollable variables and levels

uncontrollable	level 1	level 2	level 3
change in demand	change causes > 5% loss of gross profit	no change *	change causes > 5% gain of gross profit
senior executive interactions	no change * same as level 2.	Senior executives rarely deal openly with differences. End-runs are routine and disruptive. *	Senior executives are open and discuss differences. There is strong management unity.
banker actions	banks end business with HiTEM	no change *	banks cooperate with HiTEM and relax terms
loss of critical-skills personnel	lose ≥ 3 from critical skills list	no change *	gain 1 or 2 highly qualified skills

* current environmental conditions

For a specific configuration of the controllable variables, we use an ordered 6-tuple; e.g. (2,1,2,2,3,2) that means variable 1 at level 2, variable 2 at level 1, and so on. We use a 4-tuple for a configuration of uncontrollable variables. So, [(2,2,2,1,2,1);(2,2,2,2)] is BAU in the current decision situation.

Experimental data-set structure

We use an L_{18} array for our core data set (Table 8). We augment our L_{18} with the BAU treatment and the high-leverage "test treatments" 19, 20, 21, and 22, which are obtained using the Hat matrix. We compound the uncontrollable variables into the current (2,2,2,2), worst (1,2,1,1), and best (3,3,3,3) uncontrollable environments as specified by the team.

Establishing the base line: Forecasting BAU and Counter-argumentation

Each team member individually forecasts profit for BAU six months out for the three uncontrollable environments (cells *a, b, c* in Table 8). Disclosing each other's forecast was prohibited. Next, we direct each team member to write three reasons why their forecast is accurate and three reasons why not. We get 15 reasons why the forecasts are accurate and 15 opposing reasons. The team is then directed to read and debate all 30 reasons. After this discussion, they forecast the BAU treatments a second time. The non-disclosure rule still applies. Table 9 is a summary of the data from the above procedure. Note that the dispersion of the data from round 1 to round 2 declines. The group has learned from each other through the information transfer generated from our counter-argumentation process.

Forecasting the sample space

With this learning, we ask each team member to populate the entire data set similar to Table 8, where the rows were randomized differently. Each team member made 23*3= 69 forecasts, 23 treatments in three environments. We had a total of 69*5= 345 forecasts. The non-disclosure rules applied as before

Table 8. Data set structure for the HiTEM experiment

treatment	SG&A	COGS	capacity	portfolio	sales	financing	current	worst	best	uncontrollable factors
	controllable factors						uncontrollable factors' levels			
							level 2	level 1	level 3	← *cust./demand change*
							level 2	level 2	level 3	← *senior exec. interactions*
							level 2	level 1	level 3	← *banker actions*
							level 2	level 1	level 3	← *critical skills*
BAU	2	2	2	1	2	1	*a*	*b*	*c*	*BAU treatment*
1	1	1	1	1	1	1				*L_{18} treatment 1*
2	1	2	2	2	2	2				*L_{18} treatment 2*
3	3	3	3	3	3	3				*L_{18} treatment 3*
4	2	1	1	2	2	3				*L_{18} treatment 4*
5	2	2	2	3	3	1				*L_{18} treatment 5*
6	2	3	3	1	1	2				*L_{18} treatment 6*
7	3	1	2	1	3	2				*L_{18} treatment 7*
8	3	2	3	2	1	3				*L_{18} treatment 8*
9	3	3	1	3	2	1				*L_{18} treatment 9*
10	1	1	3	3	2	2				*L_{18} treatment 10*
11	1	2	1	1	3	3				*L_{18} treatment 11*
12	1	3	2	2	1	1				*L_{18} treatment 12*
13	2	1	2	3	1	3				*L_{18} treatment 13*
14	2	2	3	1	2	1				*L_{18} treatment 14*
15	2	3	1	2	3	2				*L_{18} treatment 15*
16	3	1	3	2	3	1				*L_{18} treatment 16*
17	3	2	1	3	1	2				*L_{18} treatment 17*
18	3	3	2	1	2	3				*L_{18} treatment 18*
19	3	1	3	1	1	3				*test treatment # 1*
20	1	3	1	3	3	3				*test treatment # 2*
21	1	3	3	1	1	3				*test treatment # 3*
22	3	2	3	3	1	1				*test treatment # 4*

Table 9. BAU forecasts dispersions decline between round 1 and round 2

BAU forecasts	average profit $M		standard deviation		
	round 1	round 2	round 1	round 2	change
current environment	- 5.5	-5.5	1.3	1.2	declined
worst environment	-10.9	-9.75	2.7	0.5	declined
best environment	- 4.28	-5.13	2.5	1.0	declined

Forecasting the sample space

With this learning, we ask each team member to populate the entire data set similar to Table 8, where the rows were randomized differently. Each team member made 23*3= 69 forecasts, 23 treatments in three environments. We had a total of 69*5= 345 forecasts. The non-disclosure rules applied as before.

4.5 Analyzing the data

ANOVA summary statistics

Table 10 shows the ANOVA and residual statistics for all three environments. For HiTEM, a contract manufacturer, COGS is the dominant controllable variable. The controllable variables are strong predictors ($p<<0.05$) of the profit outcome (except for *capacity* and *financing* in the best environment). The p and R^2 values suggest that the appropriate controllable variables were selected. Residual statistics ($p>>0.05$) show they are not carriers of significant information. In the worst environment, we removed outliers (difficult to forecast treatments). We use Table 8 in its entirety to analyze interactions because it gives us more dof's. There are statistically significant interactions, but their contribution to the outcome is small. This suggests that HiTEM's system behaviour is near decomposable

Table 10. ANOVA for team forecasts for current, worst, and best environments (N=72)

	current environment			worst environment			best environment		
	adj MS	MS_{adj} %	p	adj MS	MS_{adj} %	p	adj MS	MS_{adj} %	p
SG&A	56.82	7.6	0.000	73.8	9.1	0.000	56.6	8.3	0.001
COGS	569.3	76.2	0.000	622.8	76.6	0.000	532.	78.	0.000
capacity	14.6	2.	0.017	36.9	4.5	0.001	8.33	1.2	0.204
portfolio	37.1	5.	0.000	26.6	3.3	0.000	36.4	5.3	0.002
sales	51.5	6.9	0.000	28.2	3.5	0.003	37.3	5.5	0.009
financing	13.5	2.1	0.006	21.7	2.7	0.001	6.5	1.0	0.283
error	2.4	0.3	-	3.0	0.4	-	5.1	0.7	-
total	747.1	100%	-	813.1	100%	-	682.2	100%	-
	R^2 83.8%	R^2_{adj} 81.7%		R^2 81.9 %	R^2_{adj} 79.6%		R^2 69.3 %	R^2_{adj} 65.4%	
residuals	AD 0.310	p 0.548		AD 0.409	p 0.338		AD 0.468	p 0.243	

Table 11. Interactions of controllable variables

2 factor interactions	current environment		worst environment		best environment	
	adj MS %	p	adj MS %	p	adj MS %	p
COGS*sales	1.97%	0.079	-	-	-	-
COGS*capacity utilization	-	-	1.16%	0.08	-	-
customer portfolio*sales	-	-	0.9%	0.05	-	-
customer portfolio*capacity	-	-	-	-	1.31%	0.008
	R^2 90.2 %	R^2_{adj} 88.9 %	R^2 97.6 %	R^2_{adj} 97.2 %	R^2 89.2 %	R^2_{adj} 87.6 %

Gage R&R summary statistics

How "good" are the forecasts and the data produced? We apply the Gage R&R to explore this question. Gage R&R is used to analyze the sources of variation in a measurement system. We consider the team members who are forecasting the outcomes of experiments, their knowledge, data bases, formal and informal procedures, and their network of contacts as a *measurement system*. We adopt the MSA term, *"operator"*, to designate each team member who, instead of measuring a manufactured part, is making a forecast.

To obtain reproducibility and repeatability statistics, we use the four test treatments and the BAU treatment (Table 8). For these treatments, we use each operator's forecast and compare it against the value we derive using our L_{18} array. This comparison is used to test the quality of the forecast data.

- ***Reproducibility.*** Figure 1 (left panel) shows the forecasts for the four test treatments BAU, in the current environments, from our five operators. Operator 4's forecasts show a positive bias, while the other forecasts exhibit much less variation. They show more *reproducibility.*
- ***Repeatability.*** In a similar manner, we subject our four test treatments and BAU treatments to the tests of repeatability. Figure 1 (right panel) shows typical results of an operator. The graphs are "close" suggesting repeatability and that the operator was not guessing randomly.

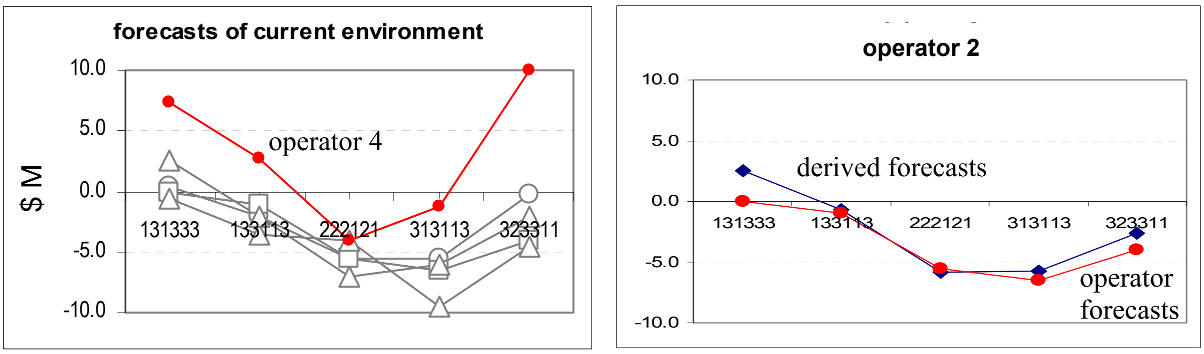

Figure 1. Sources of variability for forecasts

Table 12 shows our Gage R&R statistics. (We removed operator #4's data due to its excessive bias). The p values for treatment, operator, operator*treatment, and repeatability are statistically significant ($p<<0.05$). Of the total variation, 7.07% is from repeatability, 11.00% from reproducibility, and 81.92% from part-part.

Table 12. ANOVA for measurement variances

Two-Way ANOVA Table With Interaction					
Source	DF	SS	MS	F	p
treatment	4	299.10	74.77	38.26	0.000
operator	3	25.01	8.34	4.23	0.029
treatment*operator	12	23.46	1.95	2.49	0.035
repeatability	20	15.72	0.79		
Total	39	363.28			

Gage R&R		
Source	VarComp	% of VarComp
Total Gage R&R	2.01	18.01
Repeatability	0.79	7.07
Reproducibility	1.22	11.00
operator	0.64	5.74
operator*treatment	0.58	5.26
Part-to-Part	9.10	81.92
Total Variation	11.11	100.00

The manufacturing heuristic for a quality measurement system is 90% part-to-part variation, 5% each for repeatability and reproducibility [2]. However, we do not find any literature to inform us whether this heuristic applies in equal measure in our domain. This is an open area for further research.

4.6 Construction and analysis of alternatives

Given the forecasts for the frugal L_{18} set of controllable choices, we can construct forecasts for all possible sets of choices. One simply specifies a treatment that fits the alternative choice. The president was being pressured to improve the BAU (current situation) by changing one single controllable variable. We constructed those alternatives. None meets the profit objective and neither would a two-factor improvement policy (except in the best environment). Table 13 shows a few cases.

Table 13. Derived predictions for strategic alternatives

strategic alternatives vs. BAU		derived profit \$ M		
		environment		
		current	worst	best
2,2,2,1,2,1	BAU	-5.54	-9.40	-2.89
2.3.2.2.2.1	BAU ⊕ [COGS +] ⊕ [portfolio+]	-0.40	-4.24	2.18
3,3,2,1,2,1	BAU ⊕ [COGS +] ⊕ [SG&A+]	-0.86	-4.79	1.65
2,3,3,1,3,1	BAU ⊕ [COGS +] ⊕ [finance+]	-1.40	-4.67	1.82
3,2,2,2,2,1	BAU ⊕ [portfolio+] ⊕ [SGA+]	-2.71	-6.61	0.07
3,2,2,1,2,2	BAU ⊕ [sales+] ⊕ [SGA+]	-3.25	-7.15	-0.68

⊕ means set the variable identified by [variable+] at next higher level in the BAU 6-tuple.

HiTEM's president stated that (3,2.5,2,2,1.5,1.5) was realistically all he could do, i.e. downsize the sales force to reduce SG&A (and expect a decline in sales) and reduce manufacturing labor to reduce COGS. He was less sanguine that he could increase plant capacity sufficiently to influence profitability. He also judged that with a reduced sales force he could not take effective action on the customer-mix issue. Finally to mitigate the anticipated cash shortfall of \$10M, he was prepared to sell unused company real estate if he could not find buyers for the plants Mexico and China. Calculations for these variations yield Table 14.

Table 14. Derived predictions for variations of realistic strategy

variations of Realistic Strategy vs. BAU		derived profit $ M, σ		
		current	worst	best
(2,2,2,1,2,1)	BAU	$ -5.54 M, 1.29	$ -9.40 M, 1.06	$ -2.89 M, 1.59
(3,2.5,2,2,1.5,1.5)	realistic	$ -1.13 M, 1.00	$ -4.46 M, 1.11	$ 1.59 M, 0.44
(3,2.5,2,2,1.5,3)	realistic ⊕ [China divestiture]	$ 0.05 M, 1.24	$ -3.20 M, 0.83	$ 2.38 M, 0.74

The realistic strategy will outperform BAU in every environment. The factors that improve are SG&A, COGS, customer portfolio (e.g. COGS moves from level 2 to level 2.5), and financing - the variables that strongly impact profit, but they cannot turn around the company except in the best environment. Divestiture of the China plant in the realistic strategy can make HiTEM break even. In the current environment, there is less variation in the realistic strategy chosen by the president. It is less risky.

4.7 Findings

The method is useful and the protocol is an effective blueprint for experiments

HiTEM's president and his team were enthusiastic about the method and took immediate action. The following are examples of written feedback from the team.

> "the debate created by having to validate or disprove our actions [was useful]."
>
> "approach will make better decisions."

Table 15 summarizes the findings. It was surprising that forecasting the entire data set and the test treatments took substantially less the time than forecasting the BAU treatment. This suggests that the team can forecast complex scenarios, even under pressure.

Table 15. Summary of findings about our protocol

Framing the problem	The DOE normal form is a useful framework to specify the decision situation, the controllable and the uncertainty space using the uncontrollable variables.
Establish the base line	Counter-argumentation process works well. It promotes individual and team learning.
Forecast treatments	Post counter-argumentation, the team readily forecasts many complex treatments.
Analyze the data	Controllable variables are strong predictors ($p<0.05$ with rare exceptions) of outcome Controllable variables interactions are small. Summary statistics indicate that the team is able to identify the appropriate variables.
Analyze alternatives	Can construct any "what if" alternative and derive the outcome of any alternative
Learning about the decision	President and the team enthusiastic about the method, analysis, and findings. Good external correspondence of derived results

There is support for experimental validity

- *Construct validity.* We have construct validity. We have a conceptual framework for our experiment that is actionable using independent and dependent variables. Our framework is specified by our DOE normal form (Tables 5, 6, 7). The independent variables are the six controllable and the four uncontrollable variables. Profit is the dependent variable. Our protocol and DOE procedures make the framework operational. ANOVA data show that the controllable variables are strong predictors ($p<<0.05$) of the profit outcome except for capacity and financing in the best environment.
- *Internal validity.* Experts must judge the effects of the independent variables on the dependent variables to be consistently credible with their domain knowledge of the phenomenon under study. As domain experts, HiTEM's president and the working team judged that the independent variables exert their influence on the dependent variable (profit) in a valid manner (Figure 2).
- *External validity*. We need to show that this method is generalizable to a larger population. Another field experiment is documented in Tang [28]. Results suggest that the method is sufficiently general for many other corporate decisions. Clearly, however, more experiments are called for.
- *Reliability*. Can we repeat key experimental procedures with consistent results? We used two rounds of forecasts for BAU. Using Gage R&R, we obtained measures for repeatability, reproducibility, and part-part variation. We noted our protocol reduced the dispersion of the forecasts.

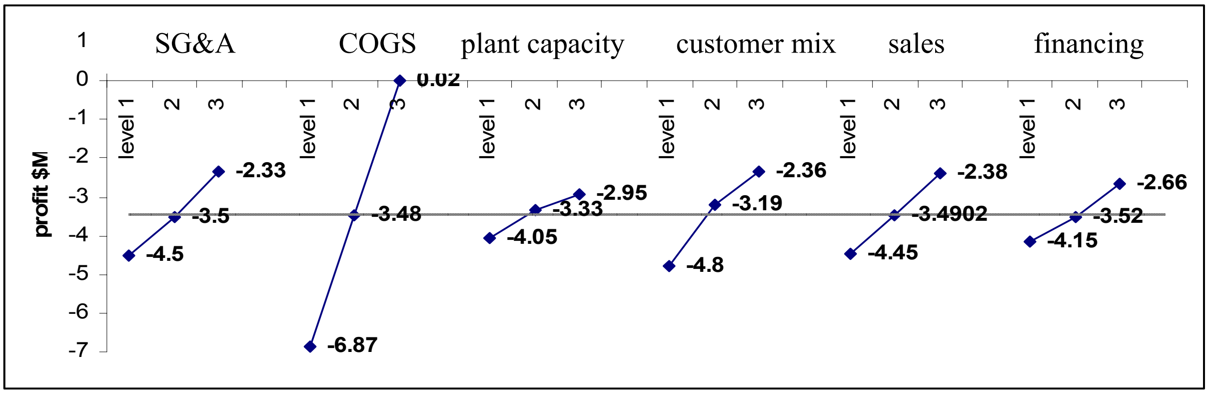

Figure 2. Controllable variables' responses

Our method passes tests of validity. Carroll and Johnson [29] specify six criteria for evaluating methods (Table 16). From this evaluation, we infer the validity of our decision analysis method.

Table 16. Caroll and Johnson's criteria for method evaluation and findings

Discovery. power to uncover new phenomena
- phenomenological behaviour of corporate systems and processes
- system behaviour of the business processes are nearly decomposable
- repeatability and reproducibility of corporate forecasting composite

Understanding . valid constructs that uncover mechanisms
- behaviour of corporate processes determined by controllable and uncontrollable variables.
- the uncertainty space can be characterized with uncontrollable variables

Prediction. ability to make predictions based on rules of logic or mathematics
- derivation of the output of any decision alternative under any specified uncertainty conditions

Prescriptive control. capability to modify the decision process and prescriptions
- construction of alternatives that trades-off performance and risk
- generation of alternative decisions over the entire solution space under any uncertainty condition

Confound control. creating controlled situations to rule out confounding elements
- controllable and uncontrollable variables separate their effects on the outcomes
- high resolution arrays separate the interaction effects from the main effects
- determine (and discriminate) the % contribution of each controllable variable to the outcome

Ease of use. economic and efficient use of time and resources
- written feedback indicates that the method is easy to use and useful to the decision-makers

What actually happened?

During the six months after we were on site, HiTEM's actual performance was (3,2.5,2.5,2.5,1,1) versus the planned "realistic" strategy of (3,2.5,2,2,1.5,1.5). HiTEM reported to the SEC a net income of $1M. In the execution of the "realistic strategy," they were able to improve on two factors but underperformed in two others. Our method predicts $0.41M. This prediction is better than it appears, during the previous two quarters HiTEM's losses exceeded $30 million.

5 DISCUSSION

Corporate business processes that support a decision are complex. The decisions are multi-functional with a variety of stakeholders with diverse and potentially competing interests. Therefore, it was surprising that the experimental data show that the interactions among the controllable variables, although present, were small. The system behaviour was nearly decomposable at our scale of analysis.

This result is consistent with principles of complex systems. Simon [31] noted that: "If we are interested only in certain aggregated aspects of behaviour, it may be that we can predict those aggregates by use of an appropriately aggregated model". And "the dynamic behaviour of a nearly-decomposable system can be analysed without examining simultaneously all the interactions of the elementary parts" [32]. Bar-Yam [33] observes that complex systems at the appropriate scale, i.e. at a level where the descriptions are self-consistent, the detailed behaviour of lower level objects is not relevant at a higher aggregated scale. And "The existence of multiple levels implies that simplicity can also be an emergent property. This means that the collective behaviour of many elementary parts can behave simply on a much larger scale". [33]

Near decomposability of decisions is also consistent with the work of management scholars. Dawes [34] disclosed the "robust beauty" of linear models, i.e., experts are capable of identifying the predictors of an outcome with a linear relationship relative to the outcome. Research studies suggest that verbally reported weights were substantially overstated the importance of minor facts [35]. These research scholars sought experts who use the interactions in their decision-making, "configural" judges. The ANOVA statistics from these configural judges showed significant interaction terms. "Despite their significance, however, these interactions rarely accounted for much judgmental variance ... judges were not necessarily mistaken when they claimed to use information configurally, *but that linear models provided such good approximations to nonlinear processes that the judges' nonlinearity was difficult to detect.*" [36] (Italics are ours) Our experiments also exhibit this phenomenon. Framing a decision problem at the appropriate scale, near decomposability of complex systems, and the near-linear behaviour of complex systems are areas worthy of more study.

6 CONCLUDING REMARKS

This research breaks new ground in corporate decision-analysis using DOE and Gage theory. It expands DOE and MSA research to an entirely new domain: administrative science. We have shown:

- **Engineering design approach to corporate decision analysis using DOE is feasible.**

One, we can explore the entire solution space. Using orthogonal arrays of controllable variables, we can derive outcomes over the entire solution space with the most parsimonious set of experiments. This capability unconstrains the range of "what if" questions decision makers can pose. Two, we can explore outcomes over the entire uncertainty space. Because the uncertainty space is constructed using uncontrollable variables, we can explore any decision alternative and its outcome over the entire space of uncertainty. This unconstrains the range of "what if" questions about uncertainty.

- **Engineering approach to analyze data quality using Gage theory is feasible.**

We can consider the executives who are forecasting, their knowledge, data bases, formal and informal procedures as a *measurement system*. Using Gage theory on this measuring system, we were able to measure repeatability and reproducibility. Regrettably, we can find no body of work to benchmark business-process Gage data. This suggests a new territory for research.

- **Validity tests suggest our *in situ* experiments and our DOE decision-analysis method are valid.**

Validity is inferred from our findings of executive feedback, statistical analyses of company experiments, validation criteria specified by scholars for tests of construct, internal, and external validity, as well as of reliability and Gage R&R analysis.

REFERENCES

[1] Montgomery D.C. *Design and Analysis of Experiments*, 2001. (John Wiley & Sons, New York).

[2] Automobile Industry Action Group (AIAG) and the American Society for Quality Control (ASQC). *Measurement Systems Analysis,* 2002. (AIAG).

[3] Keeney R.L. On the Foundations of Prescriptive Decision Analysis. *Utility Theories: Measurements and Applications,* W. Edwards (editor), 1992. (Springer-Verlag, New York).

[4] Von Neumann J. and O. Morgenstern. *Theory of Games and Economic Behavior*, 1973. (Princeton University Press, Princeton, New Jersey.).

[5] Resnik M.D. *Choices*, 2002. (University of Minnesota Press, Minneapolis).

[6] Kahneman D. and A. Tversky. Choices, Values, and Frames. *Choices, Values, and Frames*, D. Kahneman and A. Tversky (editors), 2003. (Cambridge University Press, Cambridge).

[7] Simon H.A. *Administrative Behavior,* 1997. pp. 118. (The Free Press, New York).

[8] Kahneman D. and A. Tversky. Prospect Theory: An Analysis of Decision under Risk. *Choices, Values, and Frames*, D. Kahneman and A. Tversky (editors), 2003. (Cambridge University Press, Cambridge).

[9] Klein G. *Sources of Power*, 1998. (The MIT Press, Cambridge, Massachusetts).

[10] Strategic Decisions Group. *Decision Analysis Volume 1: General Collection*, R.A. Howard and J.E. Matheson (Editors), 2004. pp. 63. (Strategic Decisions Group).

[11] Ibid. pp. 10.

[12] Keeney R.L. *Value-Focused Thinking*, 1992. (Harvard U. Press, Cambridge, MA).
[13] Otto, K.N. and E.K. Antonsson. Imprecision in Design. *Journal of Mechanical Design*, 1995.117(B) pp. 25-32.
[14] Myers S.C. Determinants of Corporate Borrowing. *Journal of Financial Economics*, 5(2), 1977. pp.147-175.
[15] Saaty T.L. and L. Vargas. *Decision Making with the Analytic Hierarchy Process*, 2006. (Springer-Verlag, New York).
[16] Keefer D.L. C.W. Kirkwood, and J.L. Corner**.** 2004. Perspective on Decision Analysis Applications, 1990-2001, *Decision Analysis* 1(1) 2004, pp. 4-22.
[17] See [7] pp. 126.
[18] Bell D.E. H. Raiffa, and A. Tversky. Descriptive, Normative, and Prescriptive Interactions in Decision Making, *Decision Making,* Bell, D.E. H. Raiffa, A. Tversky (editors), 1995, pp.18. (Cambridge University Press, Cambridge).
[19] Simon H.A. *Administrative Behavior,* 1997. pp. 126. (The Free Press, New York)..
[20] Plous S. Overconfidence. *The Psychology of Judgment and Decision Making*, 1993. (McGraw-Hill Inc. New York).
[21] Yates J.F. E.S. Veinott, and A.L. Patalano, Hard Decisions, Bad Decisions: On Decision Quality and Decision Aiding. *Emerging Perspectives on Judgment and Decision Research* (S.L. Schneider and J. Shanteau (editors), 2003, pp. 34. (Cambridge University Press, Cambridge).
[22] Russo J.E. and P.J. Schoemaker. Managing Overconfidence. *MIT Sloan Management Review*, Winter 1992 33(2), pp. 7-17.
[23] Fishoff B. Debiasing. *Judgment under Uncertainty: Heuristic and Biases,* D. Kahneman, P. Slovic, and A. Tversky (editors), 1999. (Cambridge University Press, Cambridge).
[24] Kray L.J. and A.D. Galinsky. The debiasing effect of counterfactual mind-sets. *Organizational Behavior and Human Decision Processes,* 2003, 91, pp. 69-81.
[25] Sterman J.D. Forecasts and Fudge Factors. *Business Dynamics,* 2000. (McGraw-Hill, Boston).
[26] Lerner J. S. and P. E. Tetlock. Bridging Individual, Interpersonal, and Institutional Approaches to Judgment and Decision Making. *Emerging Perspectives on Judgment and Decision Research* (S.L. Schneider, J. Shanteau (editors), 2003. (Cambridge University Press, Cambridge).
[27] Yin, R.K. *Case Study Research*, 2003. (Sage Publications, Thousand Oaks).
[28] Hoyle R.H. M.J. Harris and C.M. Judd. *Research Methods in Social Relations,* 2002. (Wadsworth, Australia).
[29] Carroll, J.S. and E.J. Johnson. *Decision Research*. 1990. (Sage Publications, Thousand Oaks).
[30] Tang, V. *Corporate Decision Analysis: An Engineering Approach*. Unpublished PhD dissertation, Massachusetts Institute of Technology, 2006. (MIT. Cambridge, Massachusetts).
[31] Simon, H. *Models of Bounded Rationality*, 1997. pp. 106 (MIT Press, Cambridge).
[32] Ibid. pp. 108.
[33] Bar-Yam Y. *Dynamics of Complex Systems*. 1997, pp. 292. (Perseus Books. Reading, MA).
[34] Dawes R.M. The robust beauty of improper linear models in decision making. *American Psychologist*, 34 1979, pp. 571-582.
[35] Slovic P. and S. Lichtenstein. Comparison of Bayesian and regression approaches to the study of information processing in judgement. *Organizational Behavior and Human Performance*, 6 1971, pp. 649-744.
[36] Goldstein W.M. and R.M. Hogarth. Judgement and decision research. *Research on Judgment and Decision Making.* W.M. Goldstein and R.M. Hogarth (editors) 1997, pp. 29. (Cambridge University Press, Cambridge).

Victor Tang
Centre for Innovation in Product Development (CIPD) and Lean Aerospace Initiative (LAI)
Massachusetts Institute of Technology
c/o Building 3 Associates
55 Deerfield Lane South
Pleasantville, New York 10270.U.S.A.
phone: USA 914.769.4049
victang@alum.mit.edu