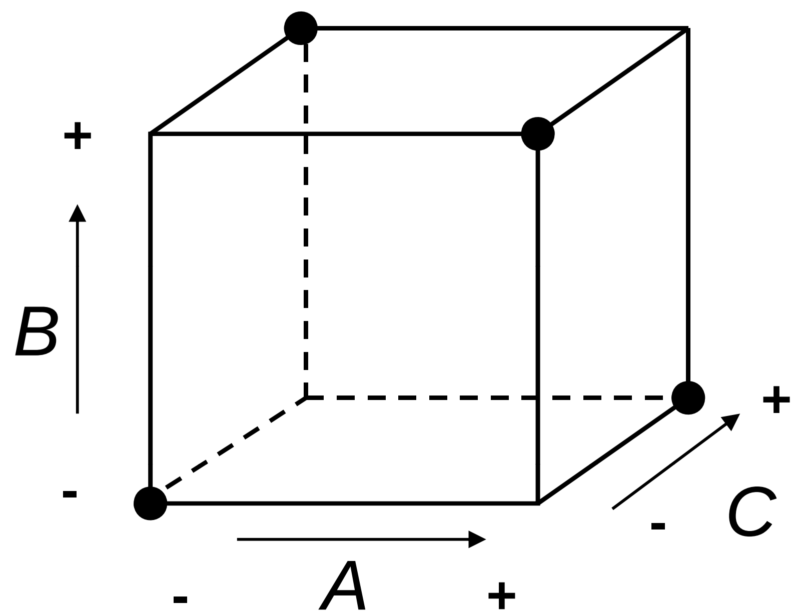
ESD.33 -- Systems Engineering

Experiments:
Strategy, Design, Analysis



Dan Frey

Plan for the Session

- Thomke -- Enlightened Experimentation
- Statistical preliminaries
- Design of experiments
  - History
  - Fundamentals
  - Frey – A role for adaptive one factor at a time

Multiple Roles of Experiments in Systems Engineering

- Promote understanding
- Calibrate our models
- Promote innovation
- Refine the product
- Evaluation and test

3D Printing

1. The Printer spreads a layer of powder from the feed box to cover the surface of the build piston.
2. The Printer then prints binder solution onto the loose powder.
3. When the cross-section is complete, the build piston is lowered slightly, and a new layer of powder is spread over its surface.
4. The process is repeated until the build is complete.
5. The build piston is raised and the loose powder is vacuumed away, revealing the completed part.

3D Computer Modeling

- Easy visualization of 3D form
- Automatically calculate physical properties
- Detect interferences in assy
- Communication!
- Sometimes used in milestones

Thomke’s Advice

- Organize for rapid experimentation
- Fail early and often, but avoid mistakes
- Anticipate and exploit early information
- Combine new and traditional technologies

Organize for Rapid Experimentation

- BMW case study
- What was the enabling technology?
- How did it affect the product?
- What had to change about the process?
- What is the relationship to DOE?

Fail Early and Often

- What are the practices at IDEO?
- What are the practices at 3M?
- What is the difference between a “failure” and a “mistake”?

What is This Prototype For?

Image removed due to copyright restrictions.

From Ulrich and Eppinger, *Product Design and Development*.

What is this Prototype For?

Image removed due to copyright restrictions.

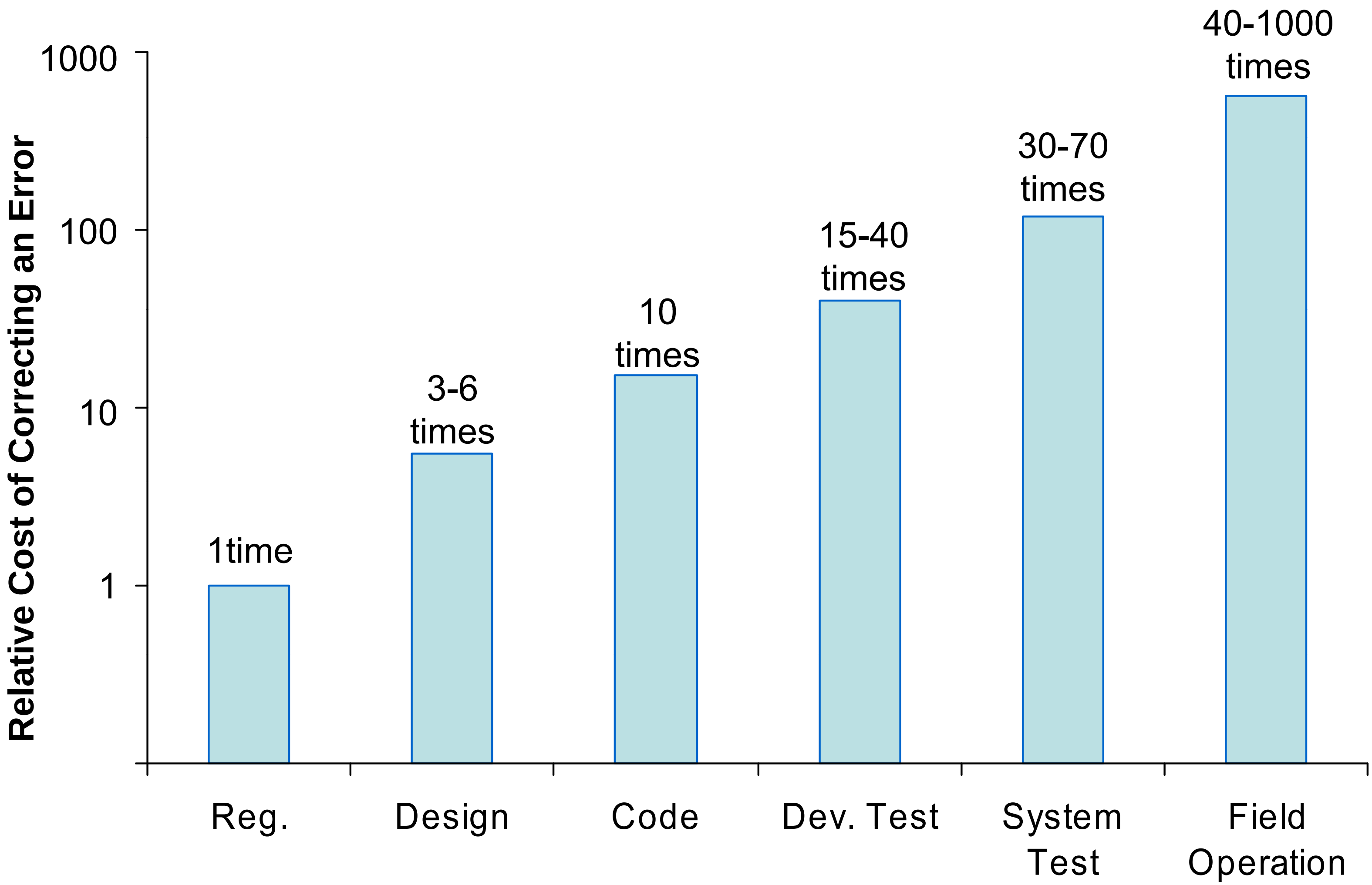
Ball supported at varying locations to determine effect on “feel”.

From Ulrich and Eppinger, *Product Design and Development*.

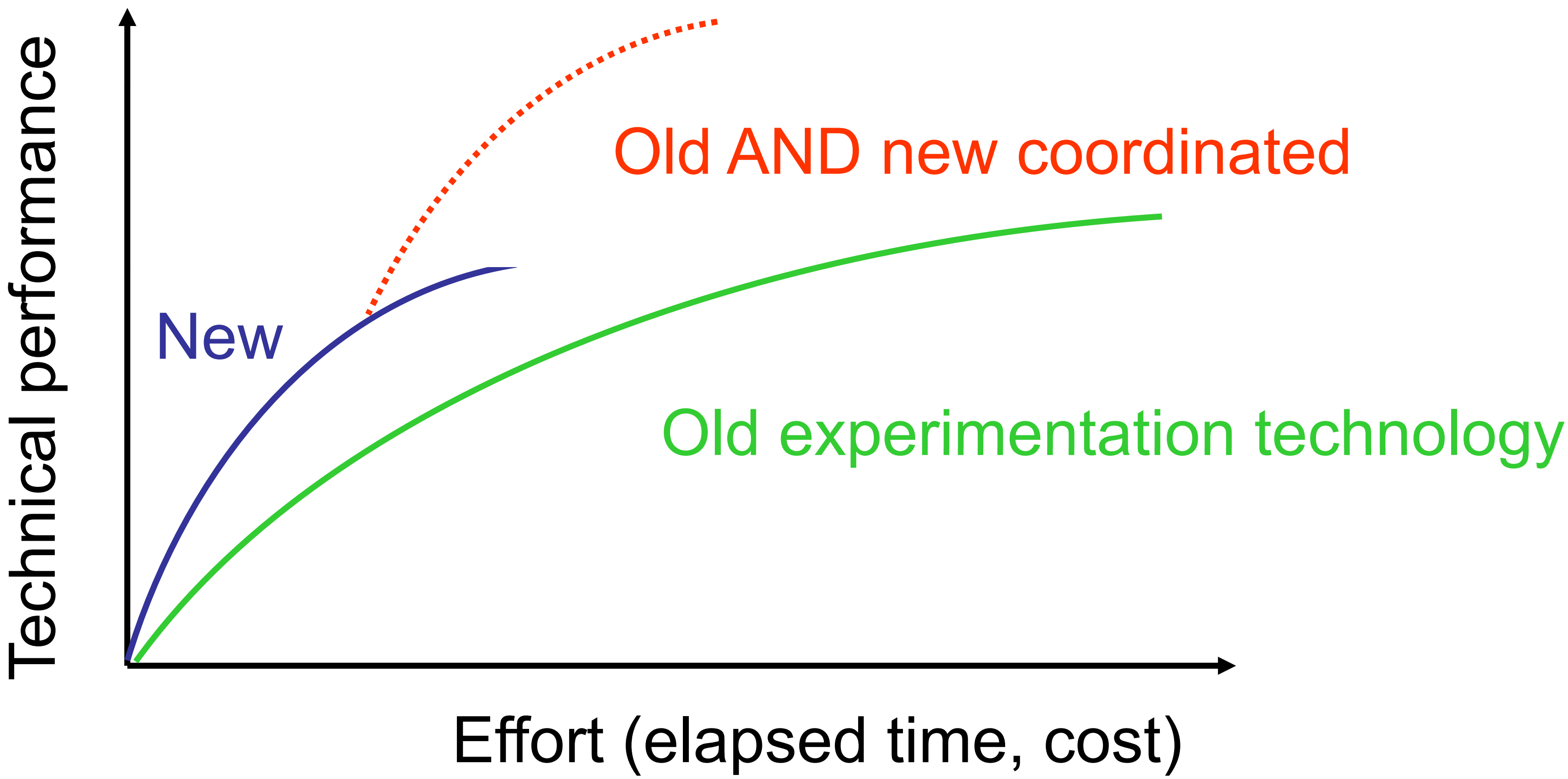
Anticipate and Exploit Early Information

- Chrysler Case study
- What was the enabling technology?
- How did it affect the product or process?
- What is the practice at your companies?

Relative cost of correcting an error



Combine New and Traditional Technologies



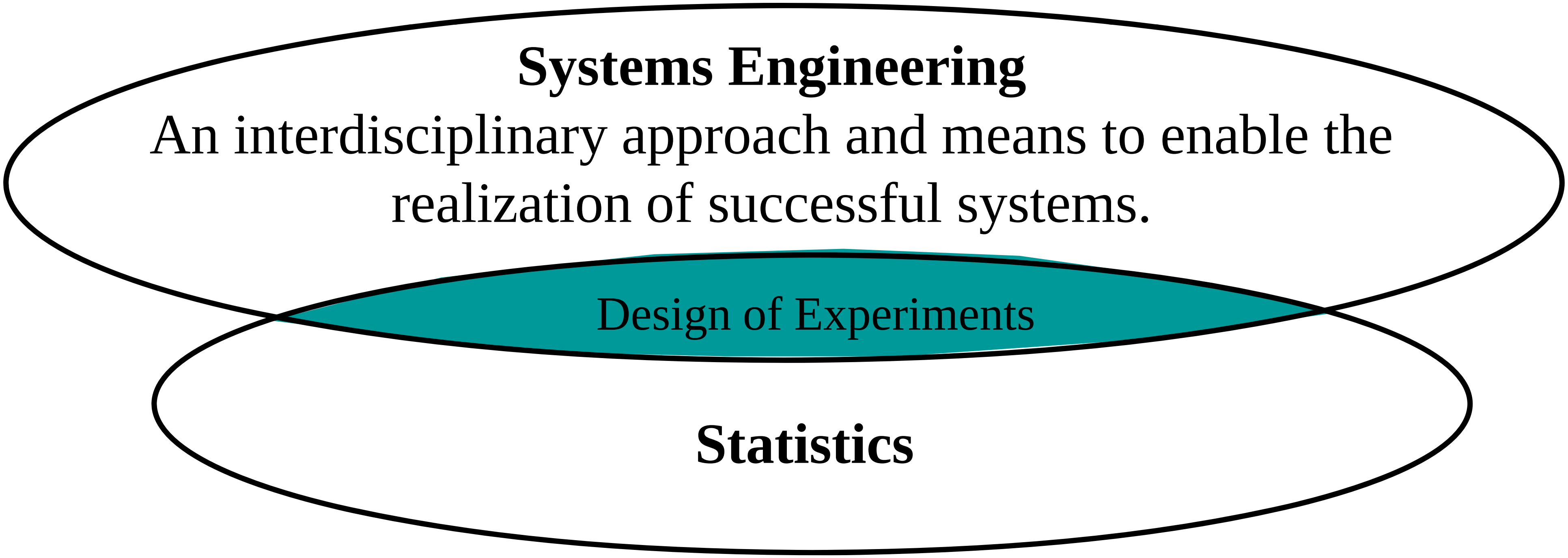
Enlightened Experimentation

- New technologies make experiments faster and cheaper
  - Computer simulations
  - Rapid prototyping
  - Combinatorial chemistry

- Thomke's theses
  - Experimentation accounts for a large portion of development cost and time
  - Experimentation technologies have a strong effect on innovation as well as refinement
  - Enlightened firms think about their system for experimentation
  - Enlightened firms don't forget the human factor

Plan for the Session

- Thomke -- Enlightened Experimentation
- Statistical preliminaries
- Design of experiments
  - History
  - Fundamentals
  - Frey – A role for adaptive one factor at a time



- Statistical THINKING is an important part of SE
- This is especially true regarding experimentation
- However, statistical RITUALS can become counterproductive in SE (and in other pursuits)

Gigernezer's Quiz

Suppose you have a treatment that you suspect may alter performance on a certain task. You compare the means of your control and experimental groups (say 20 subjects in each sample). Further, suppose you use a simple independent means t-test and your result is significant (t = 2.7, $d.f.$ = 18, p = 0.01). Please mark each of the statements below as “true” or “false.” ...

1. You have absolutely disproved the null hypothesis
2. You have found the probability of the null hypothesis being true.
3. You have absolutely proved your experimental hypothesis (that there is a difference between the population means).
4. You can deduce the probability of the experimental hypothesis being true.
5. You know, if you decide to reject the null hypothesis, the probability that you are making the wrong decision.
6. You have a reliable experimental finding in the sense that if, hypothetically, the experiment were repeated a great number of times, you would obtain a significant result on 99% of occasions.

Quiz Results

The percentages of participants in each group who endorsed one or more of the six false statements regarding the meaning of "p = 0.01."

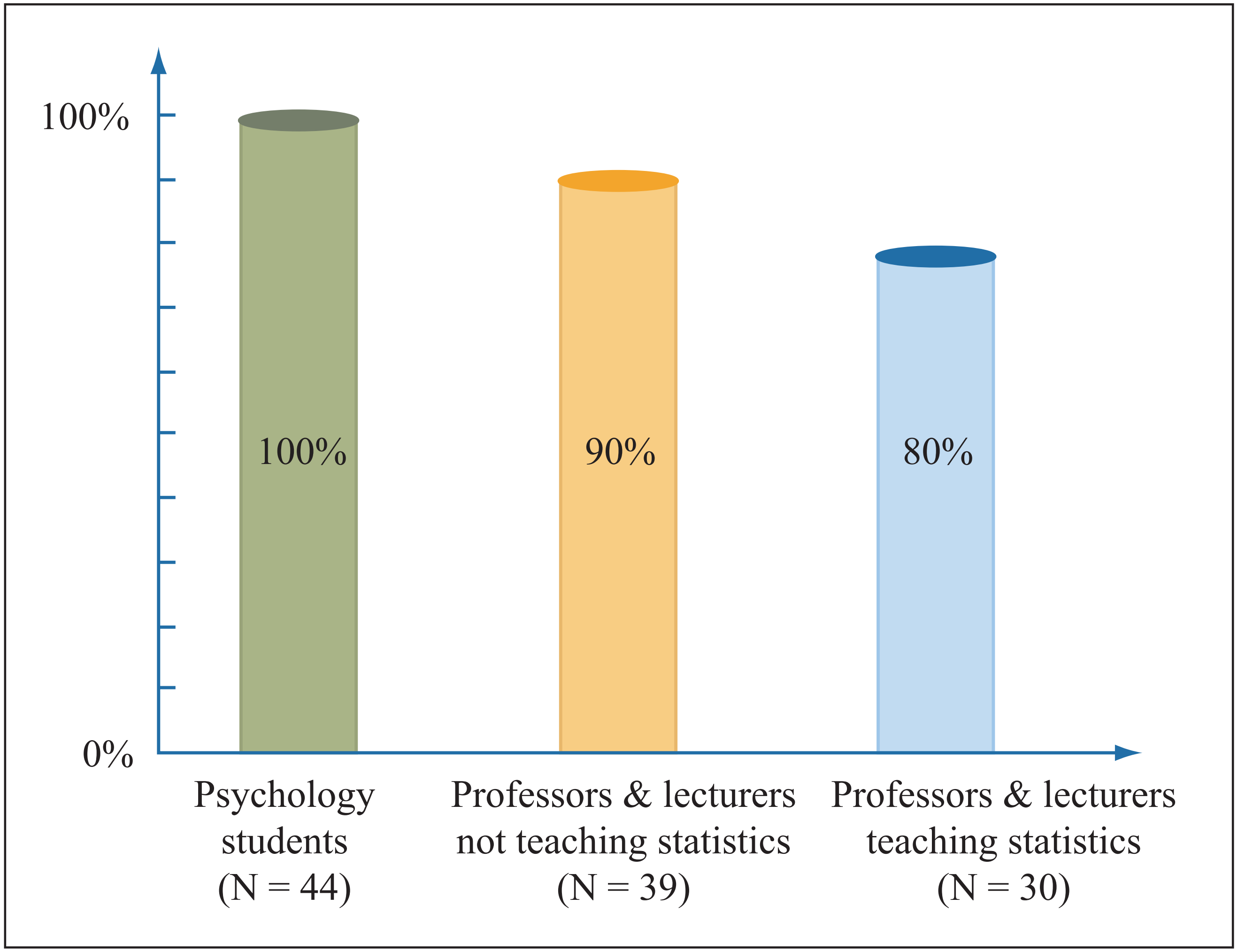


Image by MIT OpenCourseWare.

Gigerenzer, G., 2004, "Mindless Statistics," *J. of Socio-Economics* **33**:587-606.

Type-III error

- "At issue here is the importance of good descriptive and exploratory statistics rather than mechanical hypothesis testing with yes-no answers...The attempt to give an "optimal" answer to the wrong question has been called "Type-III error". The statistician John Tukey (e.g., 1969) argued for a change in perspective..."

Gigerenzer, G., 2004,"Mindless Statistics," *J. of Socio-Economics* **33**:587-606.

Plan for the Session

- Thomke -- Enlightened Experimentation
- Statistical preliminaries
- Design of experiments
  - History
  - Fundamentals
  - Frey – A role for adaptive one factor at a time

Design of Experiments

- Concerned with
  - Planning of experiments
  - Analysis of resulting data
  - Model building
- A highly developed technical subject
- A subset of statistics?
- Or is it a multi-disciplinary topic involving cognitive science and management?

“An experiment is simply a question put to nature … The chief requirement is simplicity: only one question should be asked at a time.”

Russell, E. J., 1926, “Field experiments: How they are made and what they are,” *Journal of the Ministry of Agriculture* **32**:989-1001.

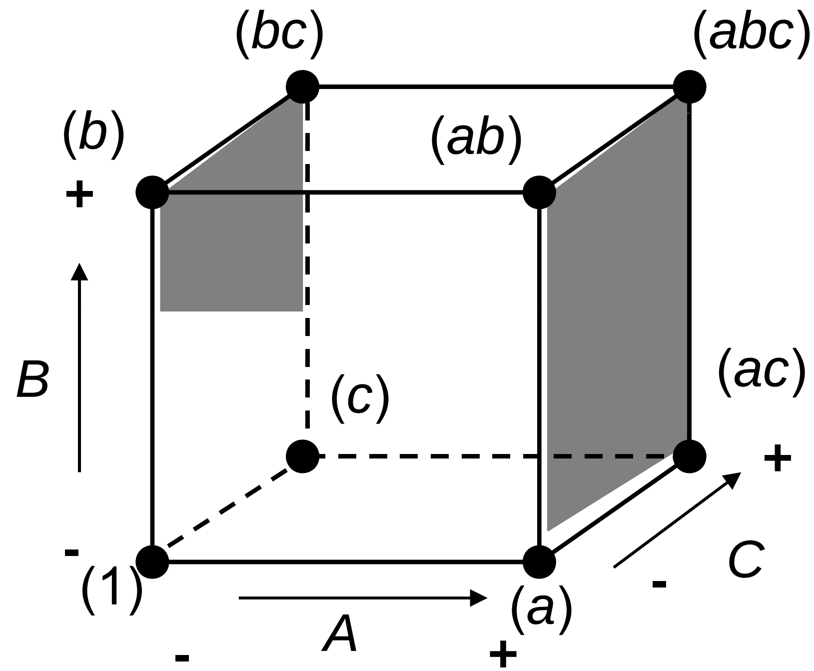
“To call in the statistician after the experiment is done may be no more than asking him to perform a post-mortem examination: he may be able to say what the experiment died of.”

- Fisher, R. A., Indian Statistical Congress, Sankhya, 1938.

Estimation of Factor Effects

Say the independent experimental error of observations (a), (ab), et cetera is σ_ε.

We define the main effect estimate A to be



$$A \equiv \frac{1}{4}\left[(abc)+(ab)+(ac)+(a)-(b)-(c)-(bc)-(1)\right]$$

The standard deviation of the estimate is

$$\sigma_A = \frac{1}{4}\sqrt{8}\sigma_\varepsilon = \frac{1}{2}\sqrt{2}\sigma_\varepsilon$$

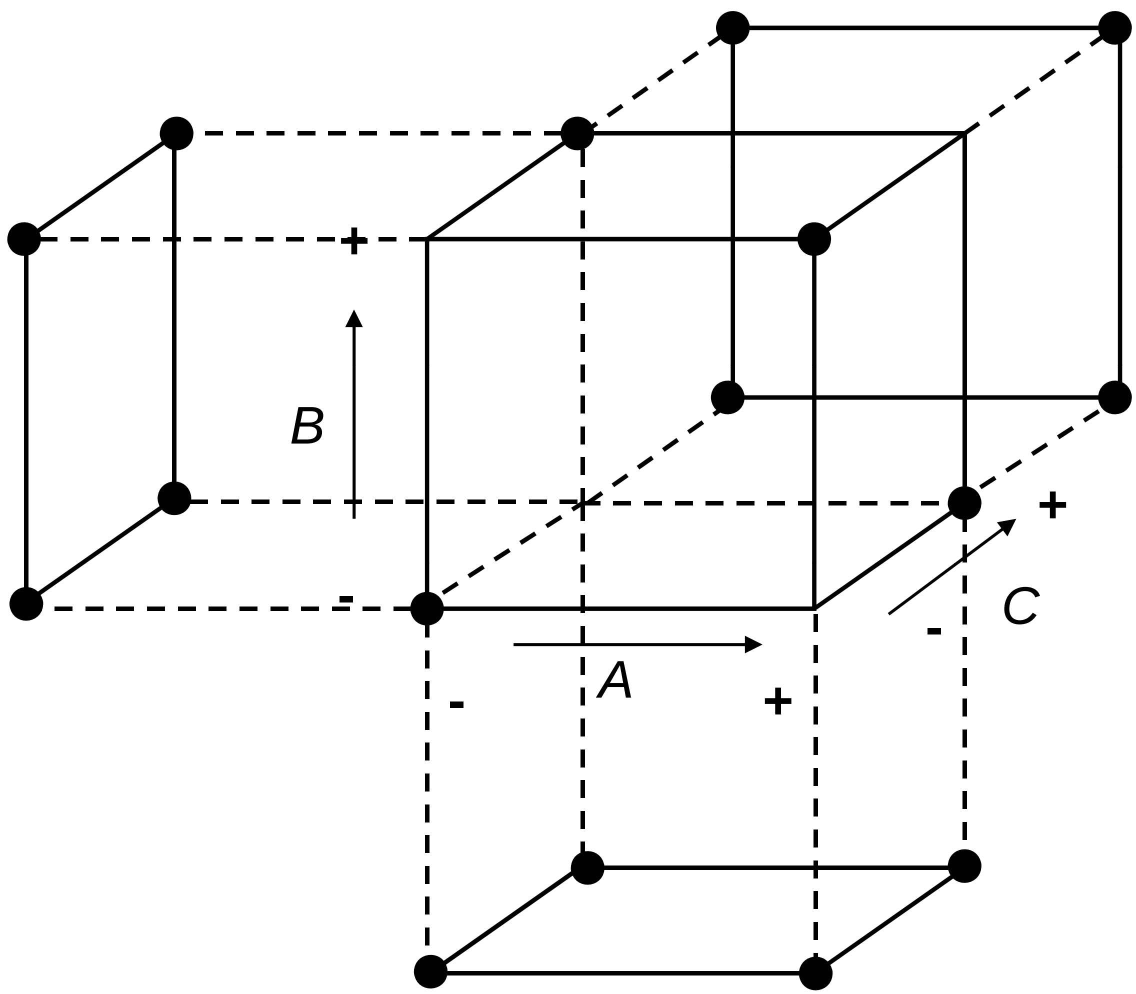
How does this compared to "single question methods"?

Fractional Factorial Experiments

"It will sometimes be advantageous deliberately to sacrifice all possibility of obtaining information on some points, these being confidently believed to be unimportant … These comparisons to be sacrificed will be deliberately confounded with certain elements of the soil heterogeneity… Some additional care should, however, be taken…"

Fisher, R. A., 1926, "The Arrangement of Field Experiments," *Journal of the Ministry of Agriculture of Great Britain*, 33: 503-513.

Fractional Factorial Experiments



$$2_{III}^{3-1}$$

Fractional Factorial Experiments

Trial	A	B	C	D	E	F	G	FG=-A
1	-1	-1	-1	-1	-1	-1	-1	+1
2	-1	-1	-1	+1	+1	+1	+1	+1
3	-1	+1	+1	-1	-1	+1	+1	+1
4	-1	+1	+1	+1	+1	-1	-1	+1
5	+1	-1	+1	-1	+1	-1	+1	-1
6	+1	-1	+1	+1	-1	+1	-1	-1
7	+1	+1	-1	-1	+1	+1	-1	-1
8	+1	+1	-1	+1	-1	-1	+1	-1

2^{7-4} Design (aka “orthogonal array”)
Every factor is at each level an equal number of times (balance).
High replication numbers provide precision in effect estimation.
Resolution III.

Plan for the Session

- Thomke -- Enlightened Experimentation
- Statistical preliminaries
- Design of experiments
  - History
  - Fundamentals
  - Frey – A role for adaptive one factor at a time

My Observations of Industry

- Farming equipment company has reliability problems
- Large blocks of robustness experiments had been planned at outset of the design work
- More than 50% were not finished
- Reasons given
  - Unforeseen changes
  - Resource pressure
  - Satisficing

"Well, in the third experiment, we found a solution that met all our needs, so we cancelled the rest of the experiments and moved on to other tasks…"

Majority View on "One at a Time"

One way of thinking of the great advances of the science of experimentation in this century is as **the final demise of the "one factor at a time" method**, although it should be said that there are still organizations which have never heard of factorial experimentation and use up many man hours wandering a crooked path.

Logothetis, N., and Wynn, H.P., 1994, *Quality Through Design: Experimental Design, Off-line Quality Control and Taguchi's Contributions*, Clarendon Press, Oxford.

Minority Views on "One at a Time"

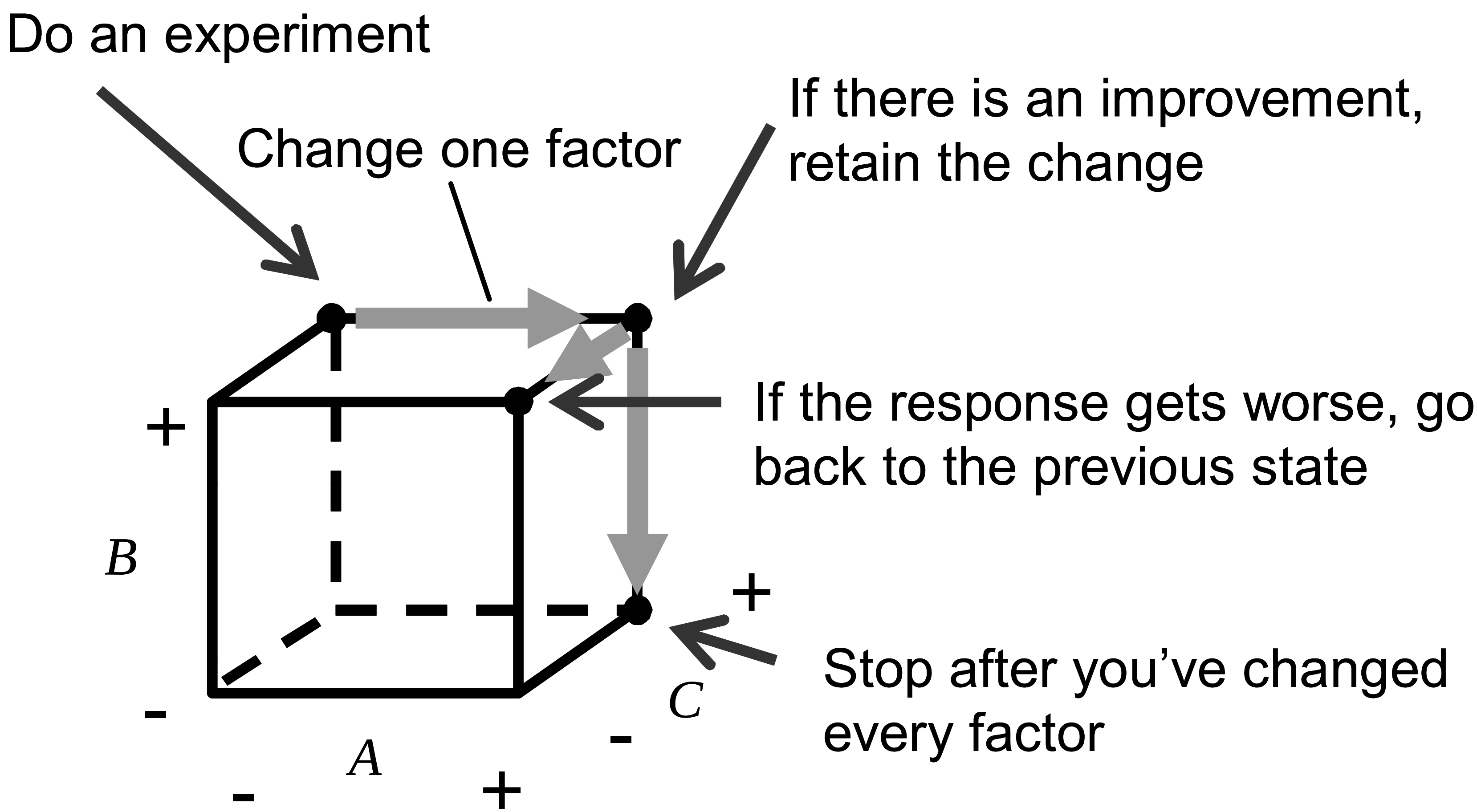
"…the factorial design has certain deficiencies … It devotes observations to exploring regions that may be of no interest…These deficiencies … suggest that an efficient design for the present purpose ought to be sequential; that is, ought to adjust the experimental program at each stage in light of the results of prior stages."

Friedman, Milton, and L. J. Savage, 1947, "Planning Experiments Seeking Maxima", in *Techniques of Statistical Analysis*, pp. 365-372.

"Some scientists do their experimental work in single steps. They hope to learn something from each run … they see and react to data more rapidly …If he has in fact found out a good deal by his methods, it must be true that the effects are at least three or four times his average random error per trial."

Cuthbert Daniel, 1973, "One-at-a-Time Plans", *Journal of the American Statistical Association*, vol. 68, no. 342, pp. 353-360.

Adaptive OFAT Experimentation



Frey, D. D., F. Engelhardt, and E. Greitzer, 2003, "A Role for One Factor at a Time Experimentation in Parameter Design", *Research in Engineering Design* 14(2): 65-74.

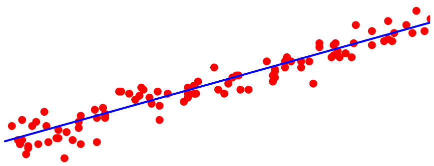
Empirical Evaluation of Adaptive OFAT Experimentation

- Meta-analysis of 66 responses from published, full factorial data sets
- When experimental error is <25% of the combined factor effects OR interactions are >25% of the combined factor effects, adaptive OFAT provides more improvement on average than fractional factorial DOE.

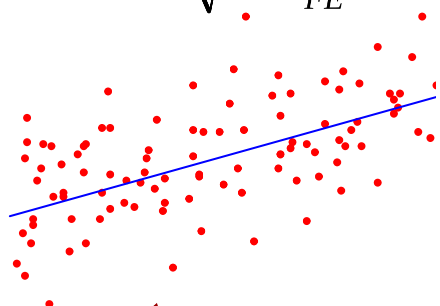
Frey, D. D., F. Engelhardt, and E. Greitzer, 2003, “A Role for One Factor at a Time Experimentation in Parameter Design”, *Research in Engineering Design* 14(2): 65-74.

Detailed Results

$$\sigma = 0.1\sqrt{MS_{FE}}$$



$$\sigma = 0.4\sqrt{MS_{FE}}$$



OFAT/FF

Gray if OFAT>FF

Interaction Strength	Strength of Experimental Error										
	0	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	1
Mild	100/99	99/98	98/98	96/96	94/94	89/92	86/88	81/86	77/82	73/79	69/75
Moderate	96/90	95/90	93/89	90/88	86/86	83/84	80/81	76/81	72/77	69/74	64/70
Strong	86/67	85/64	82/62	79/63	77/63	72/64	71/63	67/61	64/58	62/55	56/50
Dominant	80/39	79/36	77/34	75/37	72/37	70/35	69/35	64/34	63/31	61/35	59/35

A Mathematical Model of Adaptive OFAT

initial observation $\longrightarrow$ $$O_0 = y(\tilde{x}_1, \tilde{x}_2, \ldots \tilde{x}_n)$$

observation with first factor toggled $\longrightarrow$ $$O_1 = y(-\tilde{x}_1, \tilde{x}_2, \ldots \tilde{x}_n)$$

first factor set $\longrightarrow$ $$x_1^* - \tilde{x}_1 sign\{O_0 - O_1\}$$

repeat for all remaining factors

$$\text{for } i = 2 \ldots n$$

$$\left| \begin{array}{l} O_i - y\left(x_1^*, \ldots x_{i-1}^*, -\tilde{x}_i, \tilde{x}_{i+1}, \ldots \tilde{x}_n\right) \\ x_i^* = \tilde{x}_i sign\{\max(O_0, O_1, \ldots O_{i-1}) - O_i\} \end{array} \right.$$

process ends after *n*+1 observations with $$E\left[y\left(x_1^*, x_2^*, \ldots x_n^*\right)\right]$$

Frey, D. D., and H. Wang, 2006, “Adaptive One-Factor-at-a-Time Experimentation and Expected Value of Improvement”, *Technometrics* 48(3):418-31.

A Mathematical Model of a Population of Engineering Systems

$$y(x_1, x_2, \ldots x_n) = \sum_{i=1}^{n} \beta_i x_i + \sum_{i=1}^{n-1} \sum_{j=i+1}^{n} \beta_{ij} x_i x_j + \varepsilon_k$$

system response

$$\varepsilon_k \sim \mathrm{N}\left(0, \sigma_\varepsilon^{\;2}\right)$$

experimental error

$$\beta_i \sim \mathrm{N}\left(0, \sigma_{ME}^{\;2}\right)$$

main effects

$$\beta_{ij} \sim \mathrm{N}\left(0, \sigma_{INT}^{\;2}\right)$$

two-factor interactions

$y_{\max} \equiv$ the largest response within the space of discrete, coded, two-level factors

$$x_i \in \{-1, +1\}$$

Model adapted from Chipman, H., M. Hamada, and C. F. J. Wu, 2001, “A Bayesian Variable Selection Approach for Analyzing Designed Experiments with Complex Aliasing”, *Technometrics* 39(4)372-381.

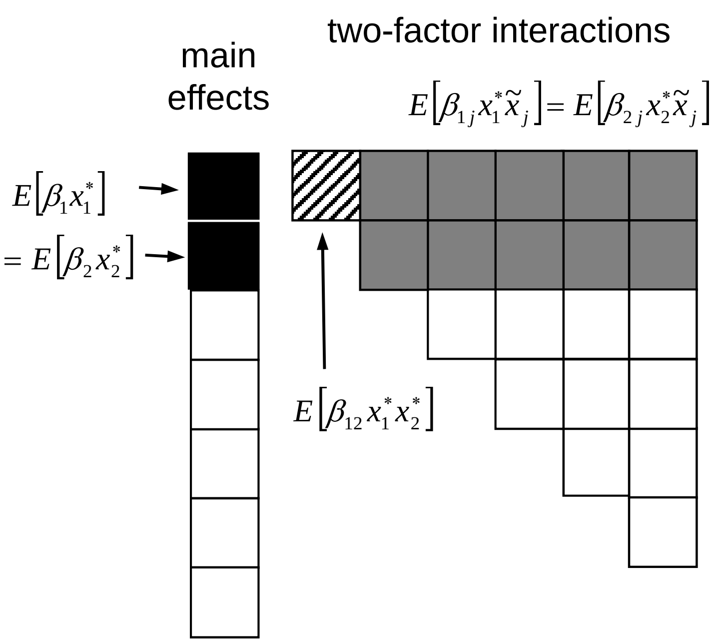
Probability of Exploiting an Effect

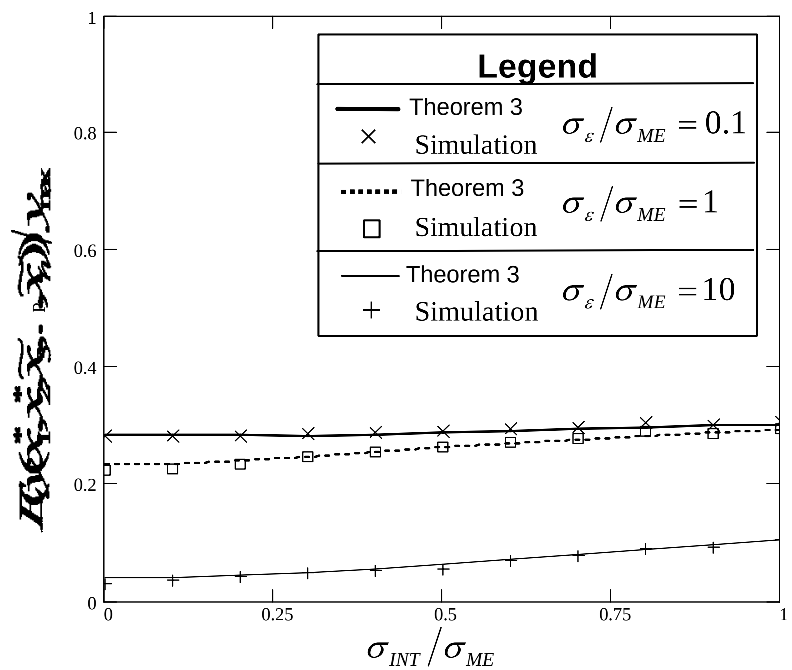
- The i^{th} main effect is said to be "exploited" if
$$\boxed{\beta_i x_i^* > 0}$$
- The two-factor interaction between the i^{th} and j^{th} factors is said to be "exploited" if $\boxed{\beta_{ij} x_i^* x_j^* > 0}$
- The probabilities and conditional probabilities of exploiting effects provide insight into the mechanisms by which a method provides improvements

The Expected Value of the Response After the *Second* Step

$$E(y(x_1^*, x_2^*, \tilde{x}_3, \ldots, \tilde{x}_n)) = 2E\left[\beta_1 x_1^*\right] + 2(n-2)E\left[\beta_{1j} x_1^*\right] + E\left[\beta_{12} x_1^* x_2^*\right]$$

$$E\left[\beta_{12} x_1^* x_2^*\right] = \sqrt{\frac{2}{\pi}} \left[\frac{\sigma_{INT}^2}{\sqrt{\sigma_{ME}^2 + (n-1)\sigma_{INT}^2 + \frac{\sigma_\varepsilon^2}{2}}} \right]$$



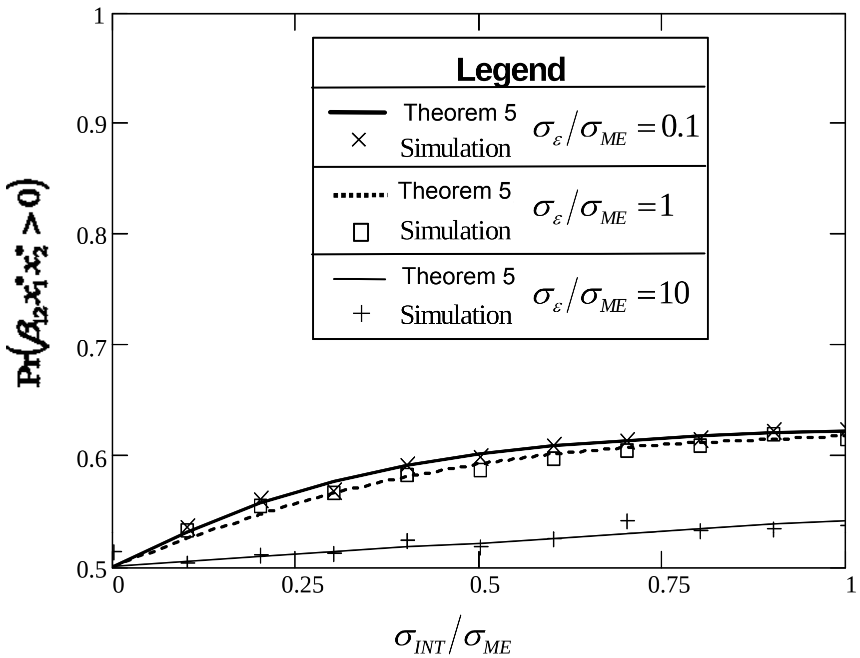


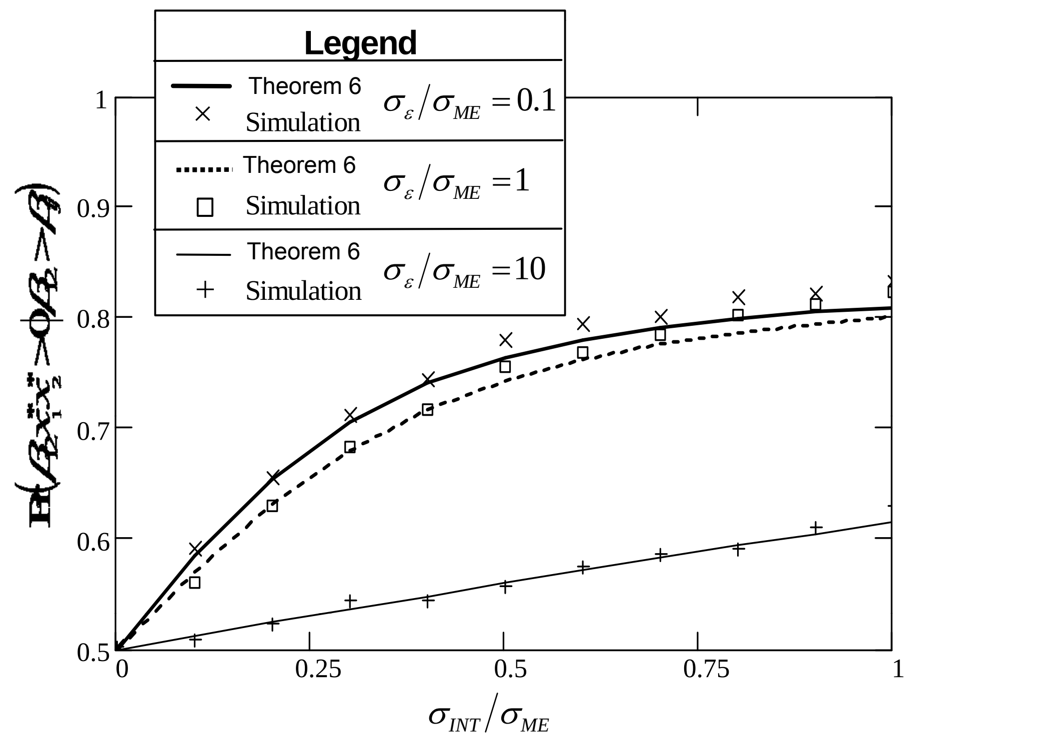
Probability of Exploiting the First Interaction

$$\Pr\left(\beta_{12}x_1^*x_2^*>0\right)=\frac{1}{2}+\frac{1}{\pi}\tan^{-1}\frac{\sigma_{INT}}{\sqrt{\sigma_{ME}^2+(n-2)\sigma_{INT}^2+\frac{1}{2}\sigma_\varepsilon^2}}$$

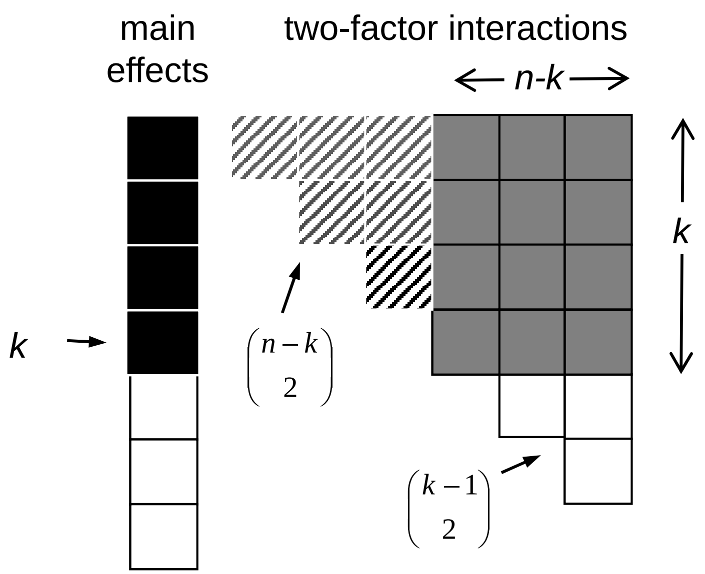
$$\Pr\left(\beta_{12}x_1^*x_2^*>0\,\middle|\,\beta_{12}>\beta_{ij}\right)>$$

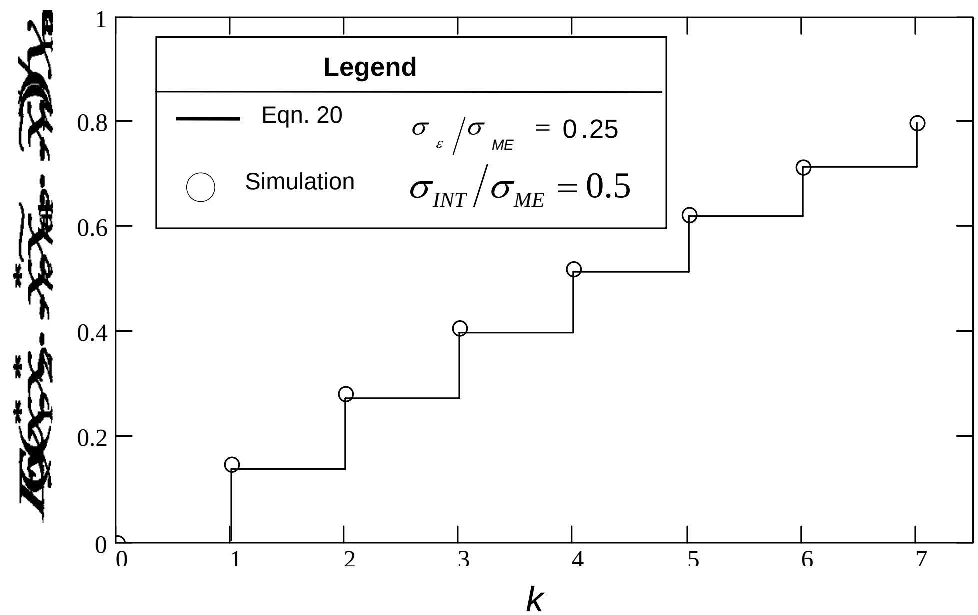
$$\frac{1}{\pi}\binom{n}{2}\int_0^\infty\int_{-x_2}^\infty\frac{\left[erf\left(\frac{1}{\sqrt{2}}\frac{x_1}{\sigma_{INT}}\right)\right]^{\binom{n}{2}-1}e^{\frac{-x_1^2}{2\sigma_{INT}^2}+\frac{-x_2^2}{2\left(\sigma_{ME}^2+(n-2)\sigma_{INT}^2+\frac{1}{2}\sigma_\varepsilon^2\right)}}}{\sigma_{INT}\sqrt{\sigma_{ME}^2+(n-2)\sigma_{INT}^2+\frac{1}{2}\sigma_\varepsilon^2}}dx_2dx_1$$





And it Continues

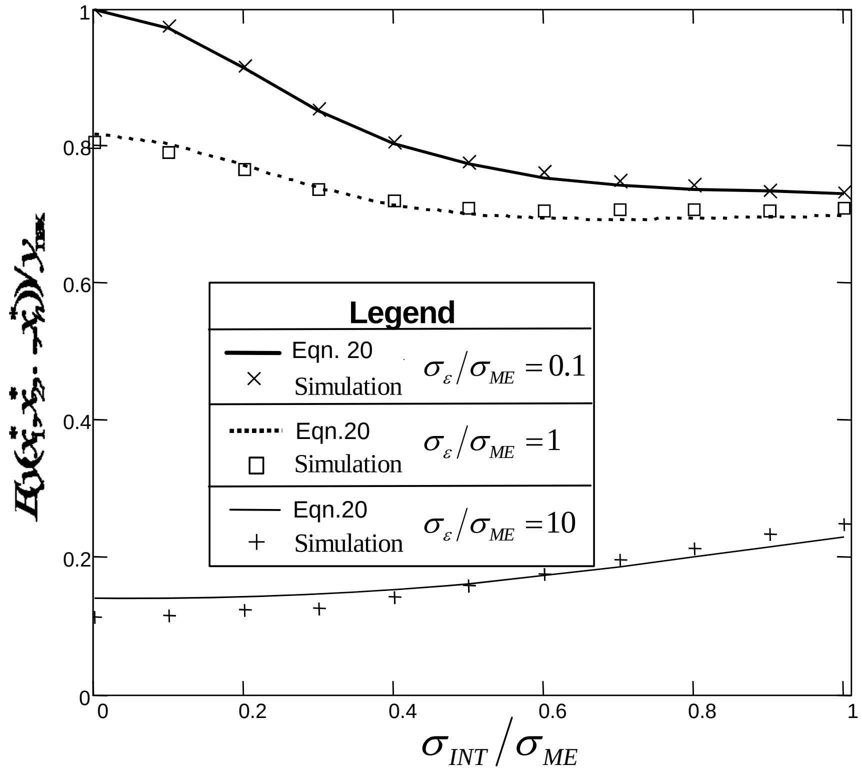




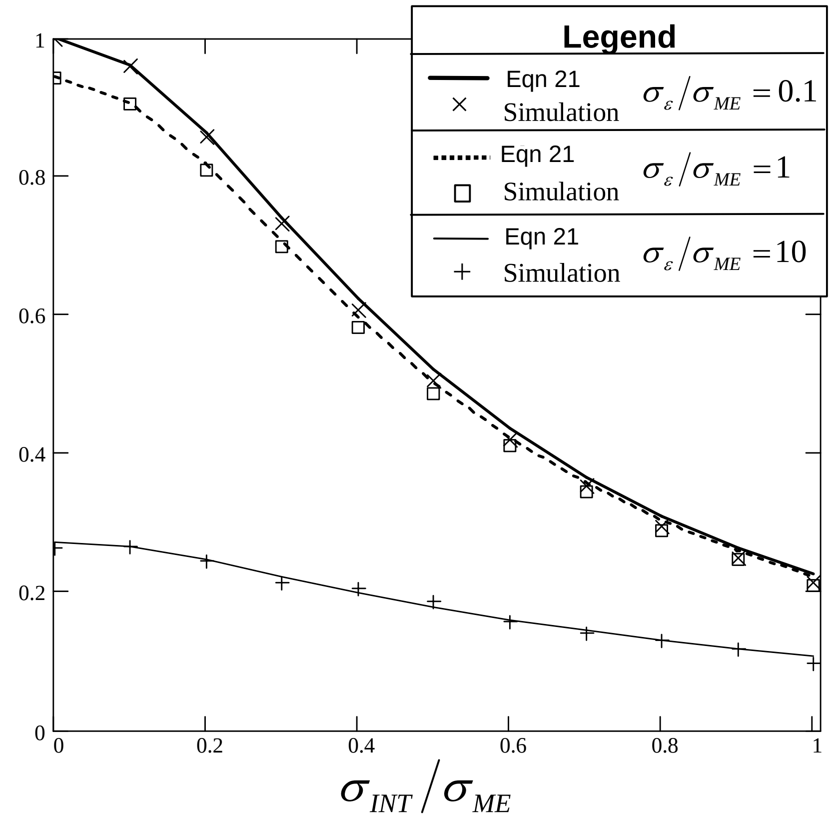
$$\Pr\left(\beta_{ij} x_i^* x_j^* > 0\right) \geq \Pr\left(\beta_{12} x_1^* x_2^* > 0\right)$$

We can prove that the probability of exploiting interactions is sustained. Further we can now prove exploitation probability is a function of *j* only and increases monotonically.

Final Outcome

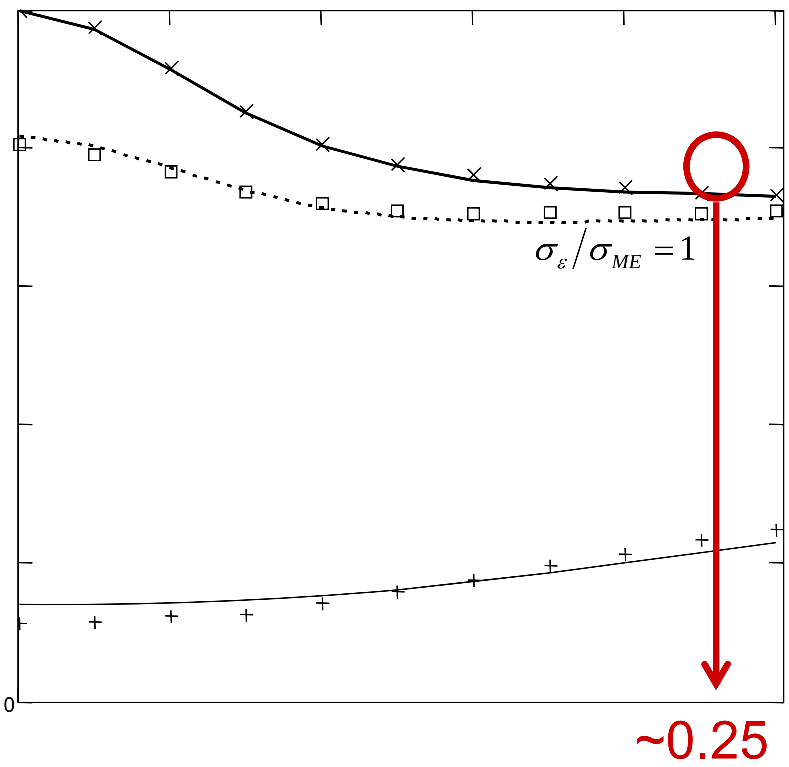


Adaptive OFAT

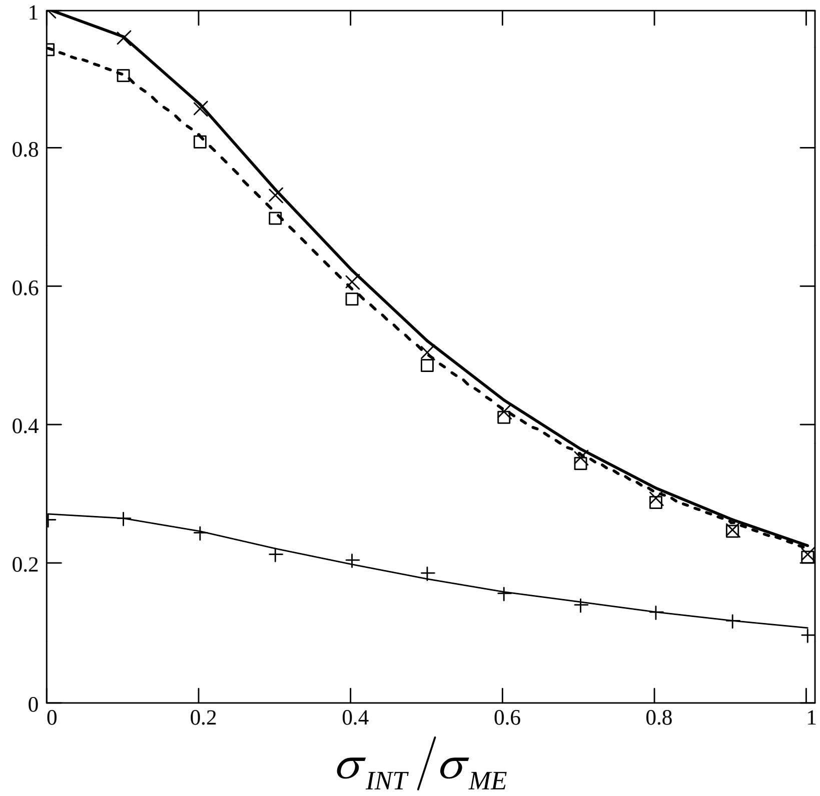


Resolution III Design

Final Outcome



Adaptive OFAT



Resolution III Design

More Observations of Industry

- Time for design (concept to market) is going down
- Fewer physical experiments are being conducted
- Greater reliance on computation / CAE
- Poor answers in computer modeling are common
  - Right model → Inaccurate answer
  - Right model → No answer whatsoever
  - Not-so right model → Inaccurate answer
    - Unmodeled effects
    - Bugs in coding the model

Human Subjects Experiment

- Hypothesis: Engineers using a flawed$_{S}$ imulation are more likely to detect the flaw while using OFAT than while using a more complex design.

- Method: Between-subjects experiment with human subjects (engineers) performing parameter design with OFAT vs. designed experiment.

Results of Human Subjects Experiment

- Pilot with N = 8
- Study with N = 55 (1 withdrawal)
- External validity high
  - 50 full time engineers and 5 engineering students
  - experience ranged from 6 mo. to 40+ yr.
- Outcome measured by subject debriefing at end

Method	Detected	Not detected	Detection Rate (95% CI)
OFAT	14	13	(0.3195,0.7133)
PBL8	1	26	(0.0009,0.1897)

Conclusions

- Experimentation is a critical part of SE
- DOE is a useful set of tools for efficient exploration and model building
- A new model and theorems show that
  - Adaptive OFAT can be more effective if the goal is improvement in system performance rather than model building
  - Adaptive OFAT exploits interactions
  - Adaptive OFAT is more effective in helping human experimenters perceive errors in computer simulations

ESD.33 Systems Engineering
Summer 2010