Design of Experiments:
Part 2

Dan Frey
Asociate Professor of Mechanical Engineering and Engineering Systems

Plan for Today

- Adaptive experimentation
- Quasi-experimental design
- Philosophy of Science and Epistemology

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.

My Observations of Industry

- Farming equipent 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
  - Unforseen 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…”

Minority Views on “One at a Time”

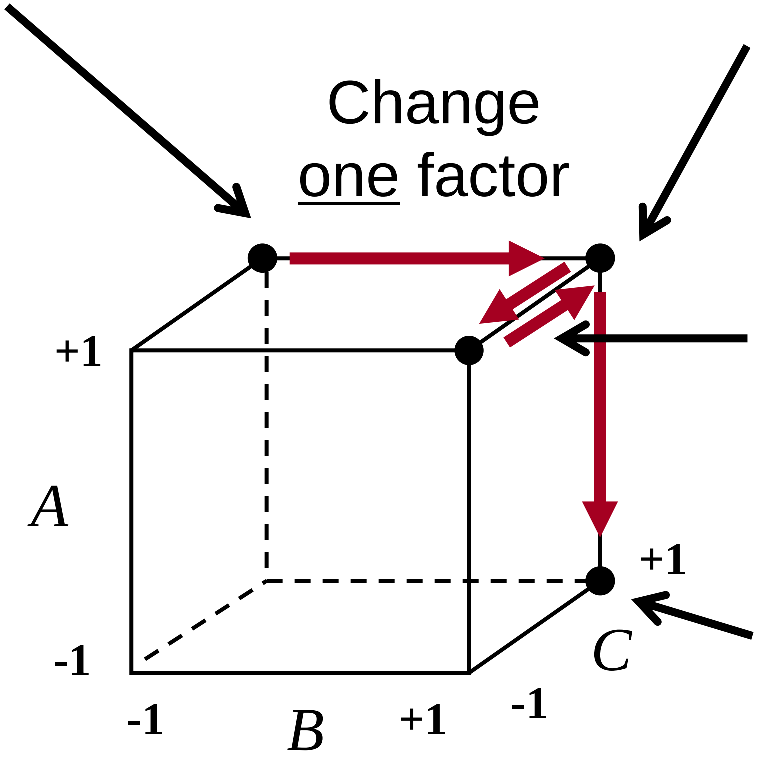
“…the factorial design has certain deficiencies … It devotes observations to exploring regions that may be of no interest…These deficiencies … suggest that an efficient design for the present purpose ought to be sequential; that is, ought to adjust the experimental program at each stage in light of the 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 One Factor at a Time Experiments



Do an experiment

Change
one factor

If there is an *apparent* improvement, retain the change

If the response gets worse, go back to the previous state

Stop after every factor has been changed exactly once

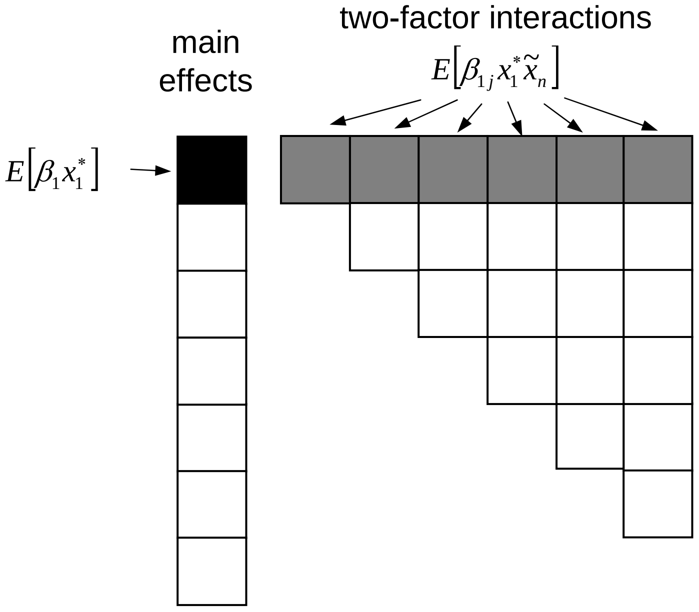
The First Step in *a*OFAT

$$E(y(x_1^*,\tilde{x}_2,\ldots,\tilde{x}_n)) = E\left[\beta_1 x_1^*\right] + (n-1)E\left[\beta_{1j} x_1^* \tilde{x}_j\right]$$

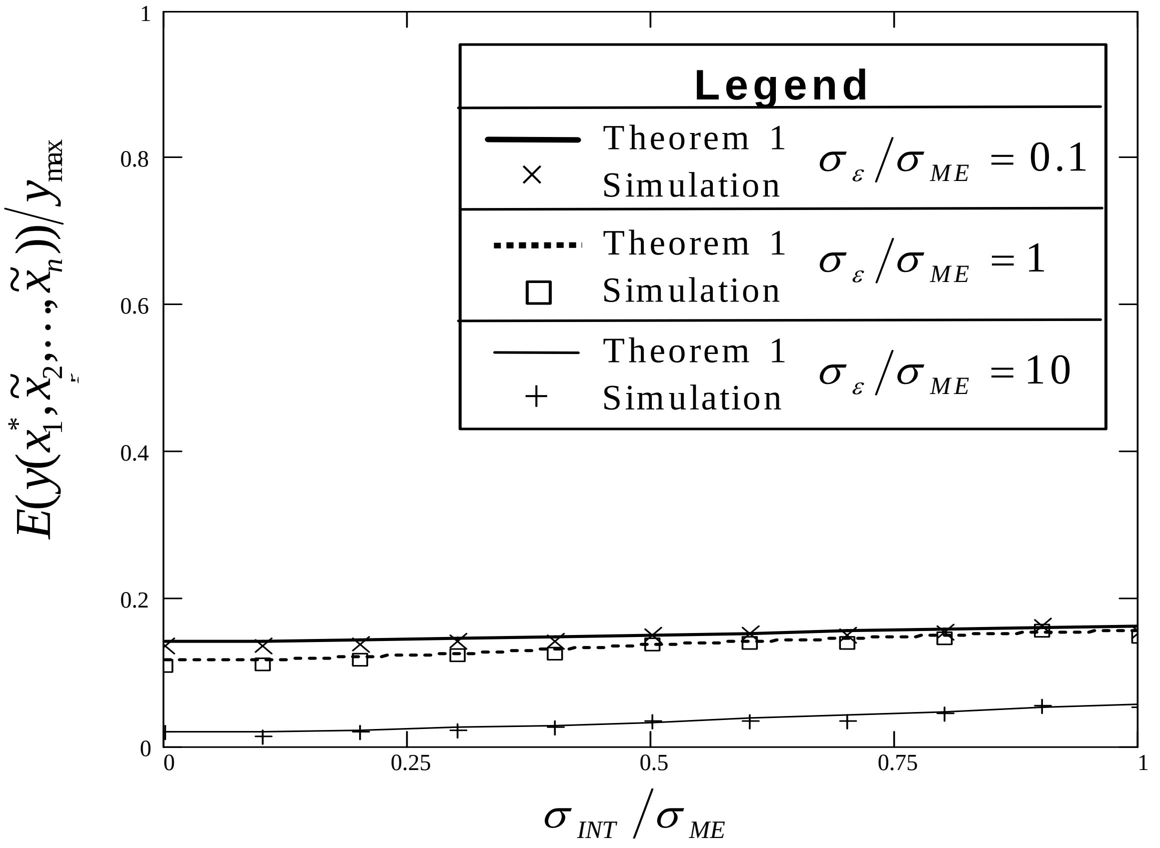
$$E\left[\beta_1 x_1^*\right] = \sqrt{\frac{2}{\pi}}\,\frac{\sigma_{ME}^2}{\sqrt{\sigma_{ME}^2 + (n-1)\sigma_{INT}^2 + \frac{1}{2}\sigma_{\varepsilon}^2}}$$

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

$$\Pr\left(\beta_1 x_1^* > 0\right) = \frac{1}{2} + \frac{1}{\pi}\sin^{-1}\frac{\sigma_{ME}}{\sqrt{\sigma_{ME}{}^2 + (n-1)\sigma_{INT}{}^2 + \frac{1}{2}\sigma_{\varepsilon}{}^2}}$$



Performance after the First Step (*n*=7)

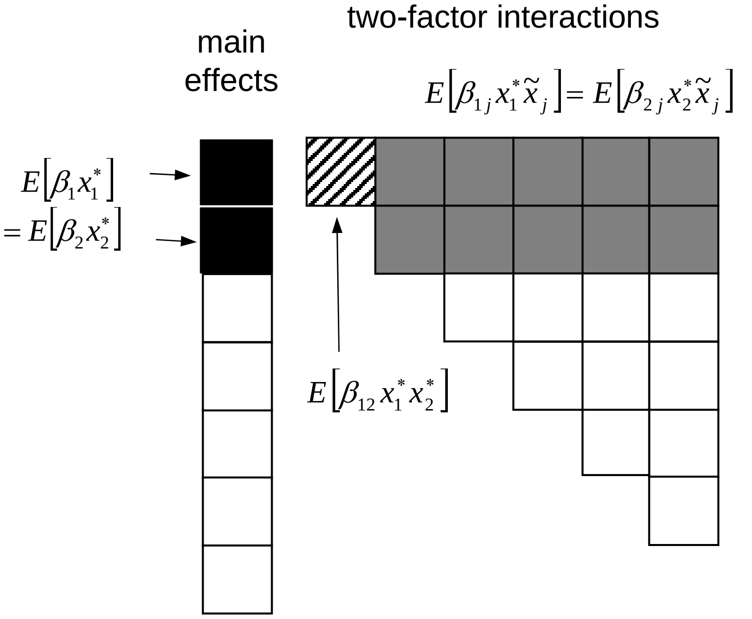


The Second Step in *a*OFAT

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

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

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



main effects

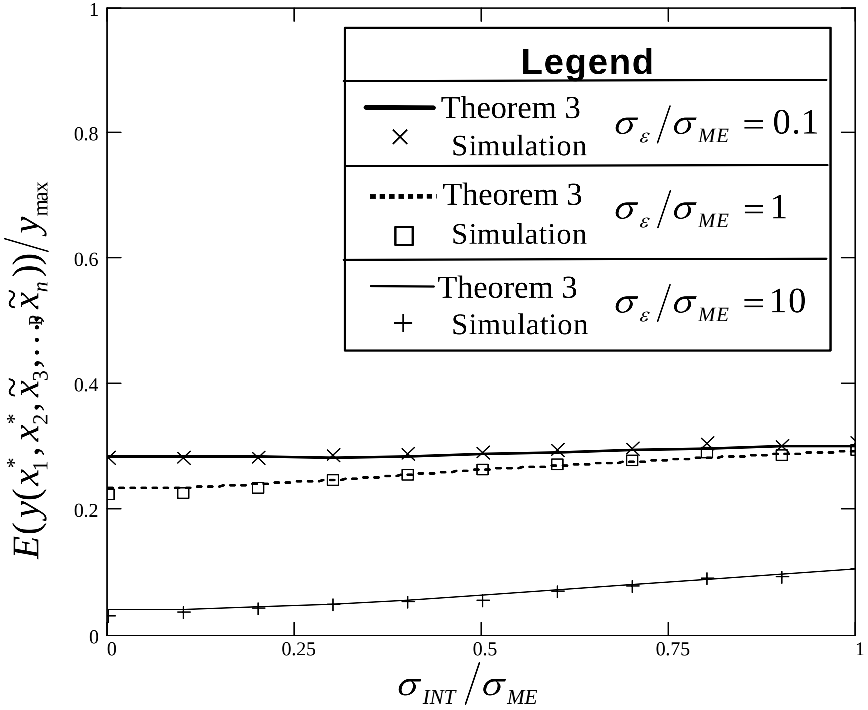
two-factor interactions

$E\left[\beta_{1j} x_1^* \tilde{x}_j\right] = E\left[\beta_{2j} x_2^* \tilde{x}_j\right]$

$E\left[\beta_1 x_1^*\right]$
$= E\left[\beta_2 x_2^*\right]$

$E\left[\beta_{12} x_1^* x_2^*\right]$

Performance after the Second Step (*n*=7)



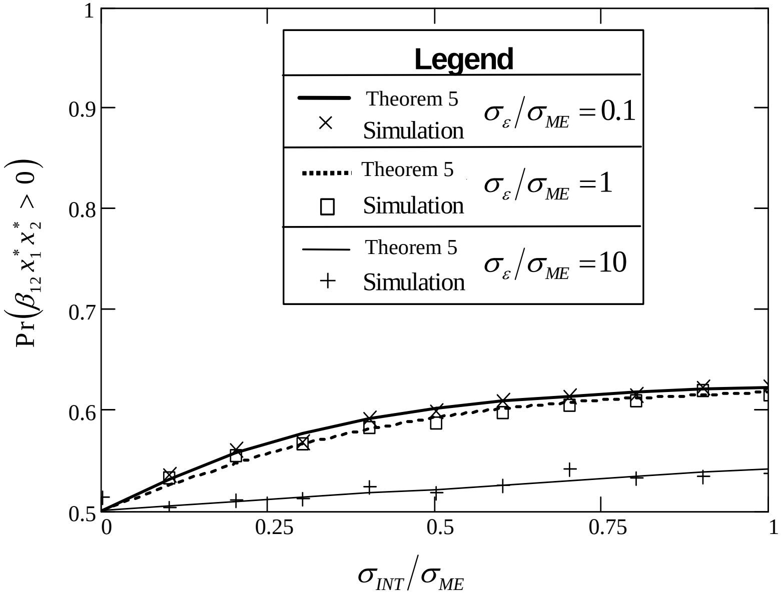
Expected improvement after the second variable is set in adaptive OFAT given a system with seven factors.

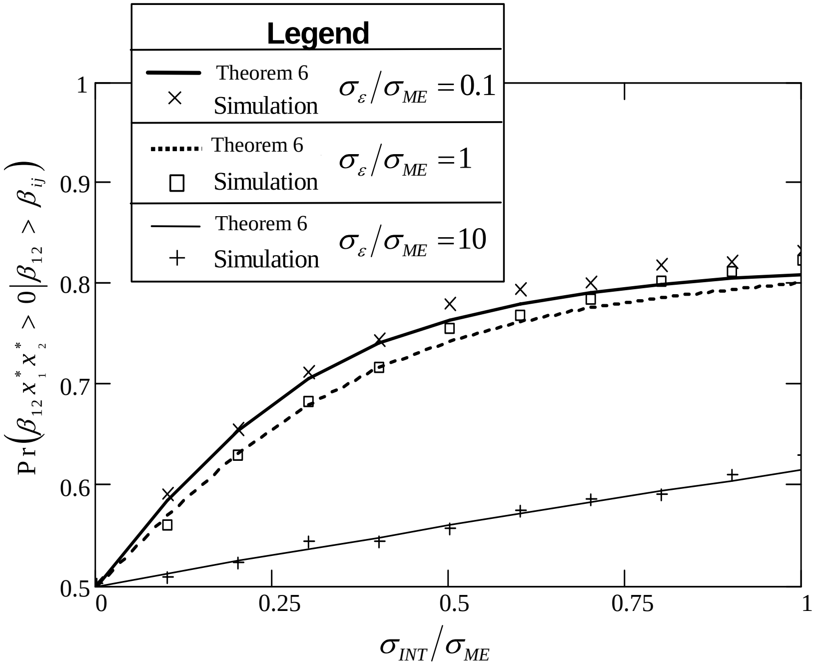
Probability of Exploiting the First Two-Factor Interaction (n=7)

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

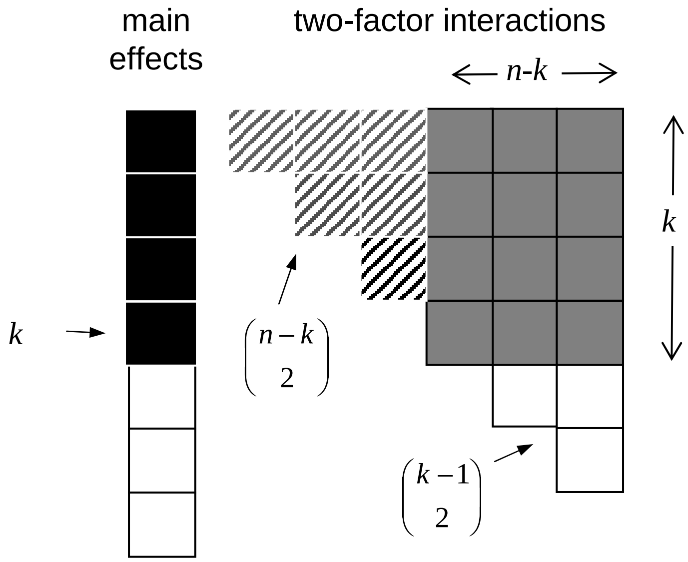
$$\Pr\left(\beta_{12}x_{_1}^* x_{_2}^* > 0 \middle| \beta_{12} > \beta_{ij}\right) >$$

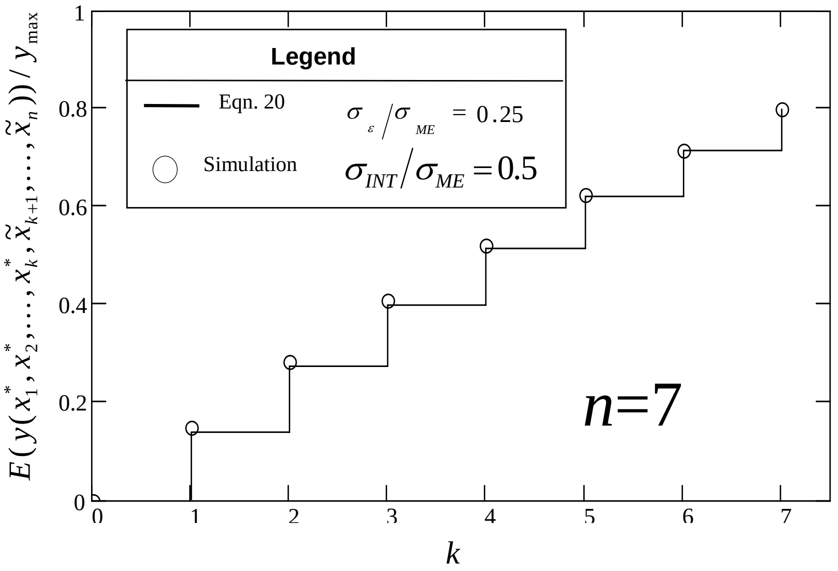
$$\frac{1}{\pi}\binom{n}{2}\int_0^\infty \int_{-x_2}^\infty \frac{\left[erf\left(\frac{1}{\sqrt{2}}\frac{x_1}{\sigma_{INT}}\right)\right]^{\binom{n}{2}-1} e^{\frac{-x_1{}^2}{2\sigma_{INT}{}^2}+\frac{-x_2{}^2}{2\left(\sigma_{ME}{}^2+(n-2)\sigma_{INT}{}^2+\frac{1}{2}\sigma_\varepsilon{}^2\right)}}}{\sigma_{INT}\sqrt{\sigma_{ME}{}^2 + (n-2)\sigma_{INT}{}^2 + \frac{1}{2}\sigma_\varepsilon{}^2}} dx_2 dx_1$$





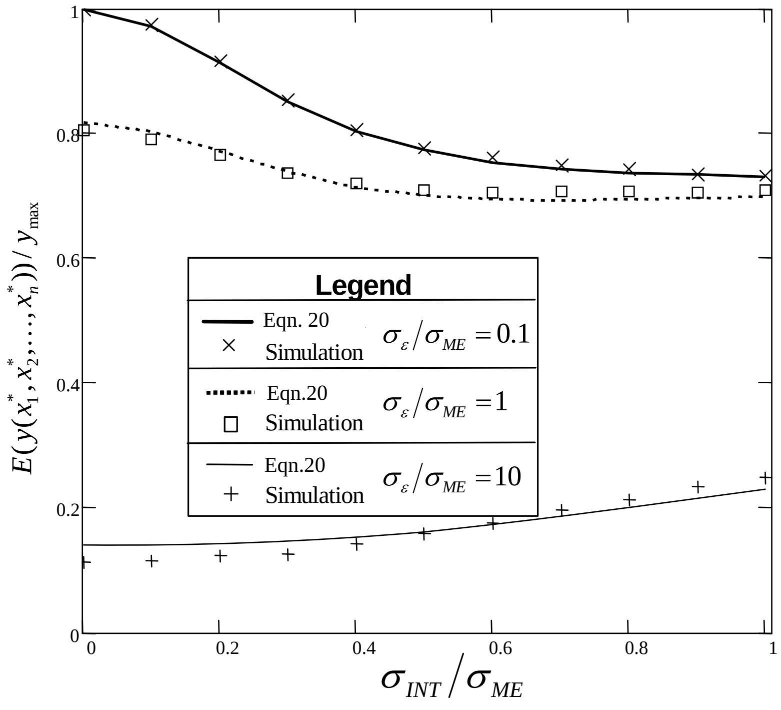
Performance after Multiple Steps



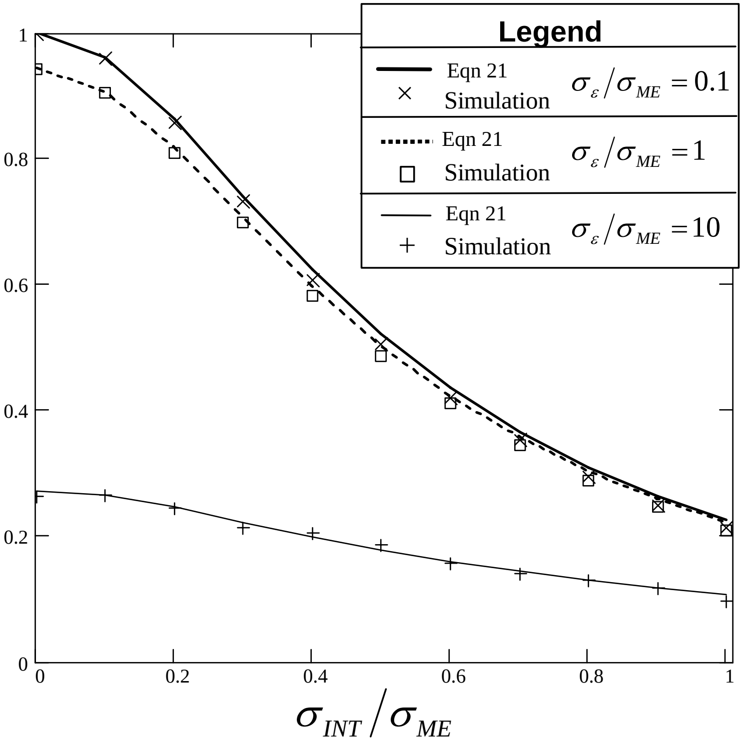


$$\Pr\left(\beta_{ij}x_i^*x_j^*>0\right)\geq\Pr\left(\beta_{12}x_1^*x_2^*>0\right)$$

Final Outcome (*n*=7)

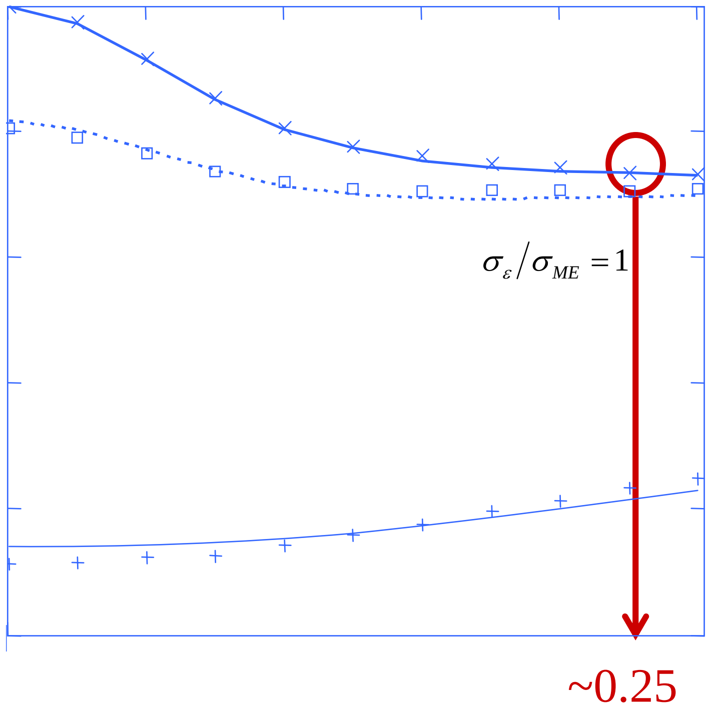


Adaptive OFAT

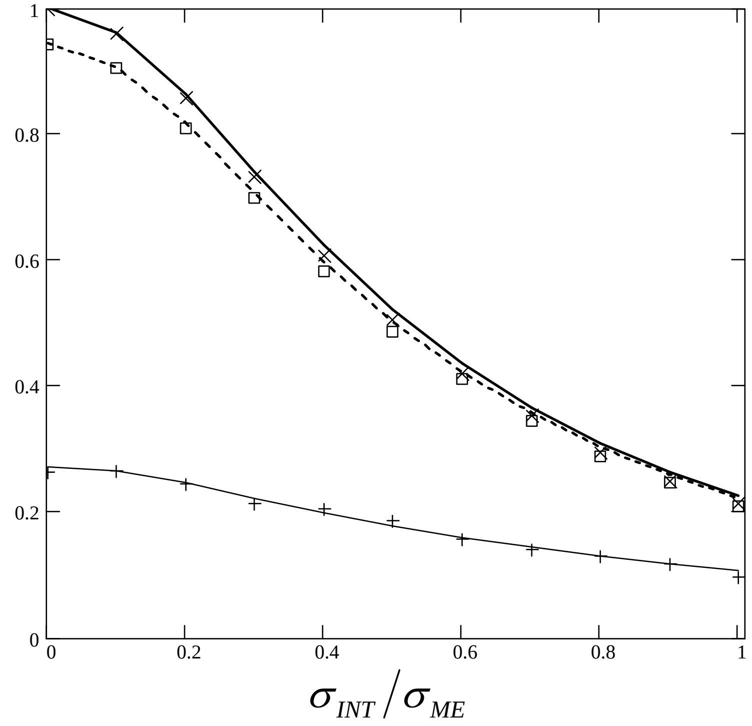


Resolution *III* Design

Final Outcome (n=7)



Adaptive OFAT



Resolution *III* Design

Electric Airplane Experiment

From Unified Engineering 16.01-16.04

		Level	
Factor	Description	-	+
A	Propeller diameter	7 in.	8 in.
B	Propeller pitch	4 in.	5 in.
C	Gear ratio	1:1	1:1.85
D	Wing Area	450 in^2	600 in^2
E	Cells in battery	7	8
F	Motor Type	SP400 7.2V	SP480 7.2V
G	Number of motors	1	2

Electric Airplane – Active Effects

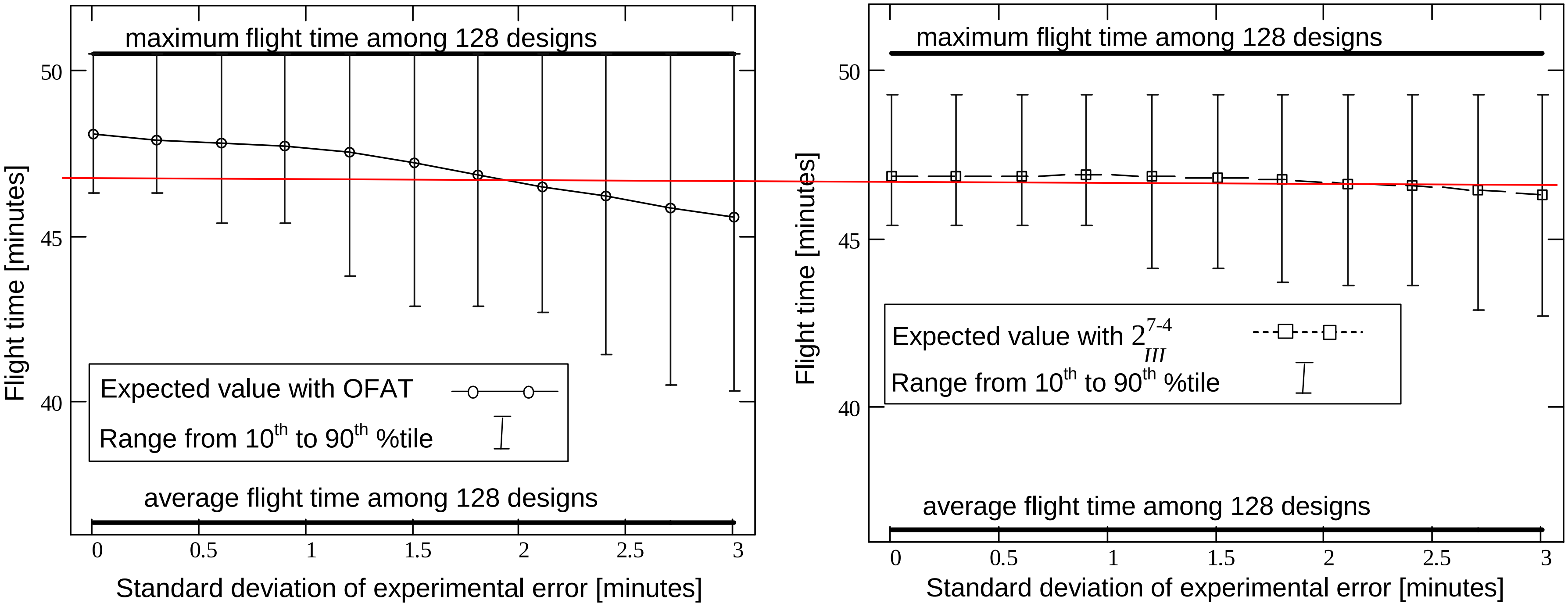
Term	Coefficient
C	9.71
G	5.10
E	3.58
F	-3.24
D*G	1.91
A*C	1.43
C*F*G	-1.13
E*G	0.90
B*C	0.83
D*E*G	0.83
C*D*E*F	0.79
B	-0.79
B*G	0.38
A*F	-0.35

If there is are two motors, the increase in wing area is advantageous

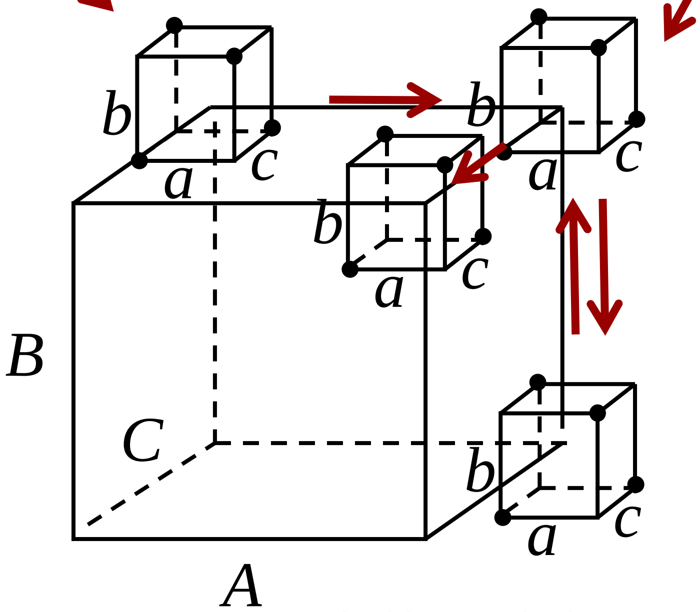
If the motor is geared down, the increase propeller diameter is advantageous

Response = Maximum flight time

Electric Airplane Results



Adaptive “One Factor at a Time” for Robust Design



Run a resolution III on noise factors

Change one factor

Again, run a resolution III on noise factors. If there is an improvement, in transmitted variance, retain the change

If the response gets worse, go back to the previous state

Stop after you’ve changed every factor once

Courtesy of ASME. Used with permission.

Frey, D. D., and N. Sudarsanam, 2006, “An Adaptive One-factor-at-a-time Method for Robust Parameter Design: Comparison with Crossed Arrays via Case Studies,” accepted to *ASME Journal of Mechanical Design*.

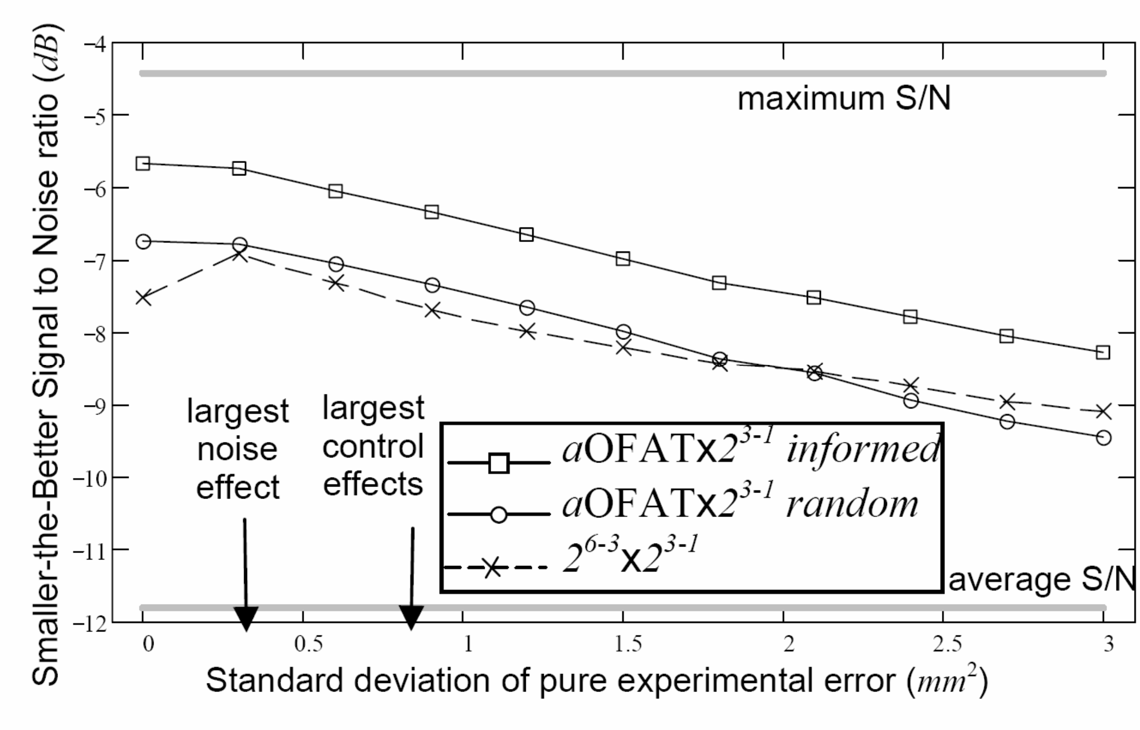
A Manufacturing Case Study

Image removed due to copyright restrictions.

- Sheet metal spinning
- 6 control factors (number of passes of the tool, etc.)
- 3 noise factors (material properties, etc.)
- Goal = more uniform geometry

Kunert, J., et. al., 2004, "An experiment to compare the combined array and the product array for robust parameter design," accepted to *J. of Quality Technology*.

A Manufacturing Case Study

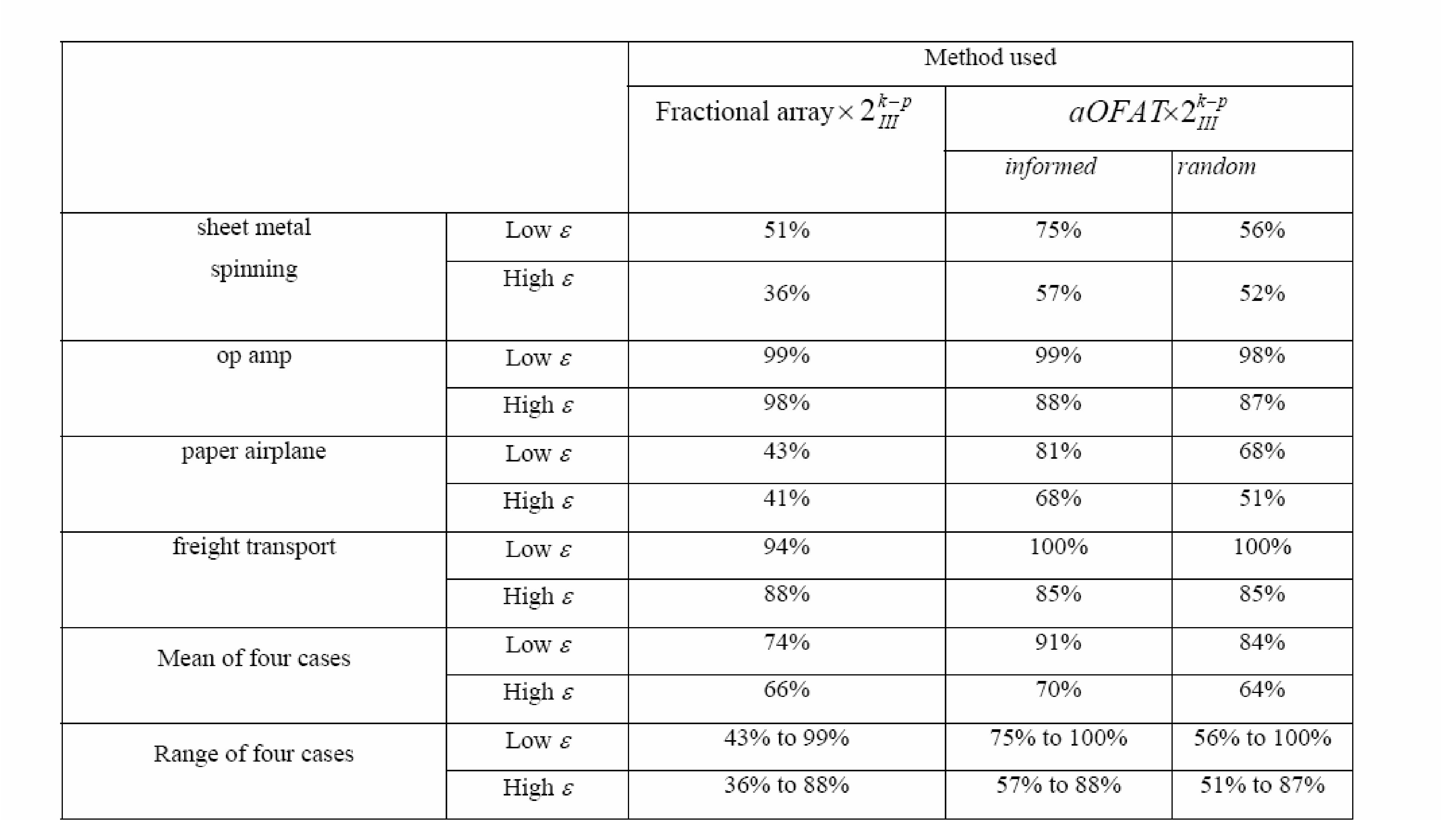


Courtesy of ASME. Used with permission.

- aOFAT worked better if experimental error not too high
- Especially true if an informed starting point was used

Frey, D. D., N. and Sudarsanam, 2006, “An Adaptive One-factor-at-a-time Method for Robust Parameter Design: Comparison with Crossed Arrays via Case Studies,” accepted to *ASME Journal of Mechanical Design*.

Results Across Four Case studies



		Method used		
		Fractional array × 2_{III}^{k-p}	$aOFAT \times 2_{III}^{k-p}$	
			informed	*random*
sheet metal spinning	Low ε	51%	75%	56%
	High ε	36%	57%	52%
op amp	Low ε	99%	99%	98%
	High ε	98%	88%	87%
paper airplane	Low ε	43%	81%	68%
	High ε	41%	68%	51%
freight transport	Low ε	94%	100%	100%
	High ε	88%	85%	85%
Mean of four cases	Low ε	74%	91%	84%
	High ε	66%	70%	64%
Range of four cases	Low ε	43% to 99%	75% to 100%	56% to 100%
	High ε	36% to 88%	57% to 88%	51% to 87%

Courtesy of ASME. Used with permission.

Frey, D. D., N. and Sudarsanam, 2006, “An Adaptive One-factor-at-a-time Method for Robust Parameter Design: Comparison with Crossed Arrays via Case Studies,” accepted to *ASME Journal of Mechanical Design*.

Conclusions : Adaptive Experimentation

- If the goal is maximum improvement rather than maximum precision in estimation
- And experimental error is not too large
- And sequential experiments are possible
- Then adaptive experimentation provides significant advantages over factorial plans
- Mostly because it exploits two-factor interactions, especially the largest ones
- Proven to be effective for robust design

Plan for Today

- Adaptive experimentation
- ➡ Quasi-experimental design
- Philosophy of Science and Epistemology

Quasi-Experimental Design

- Treatments do not meet fully the criteria of an experiment
  - Not actually applied by the experimenter, but occurred "naturally"
  - OR not randomized
  - OR no control group
- Primary techniques
  - Comparison group design
  - Interrupted time series
- Key issue – entertain seriously the alternative hypotheses

A Comparison Group Study

- One area of the country chosen (Boston)
- Three groups of homes known to have been weatherized in a certain year
- Control = randomly slected homes not among those known to be weatherized

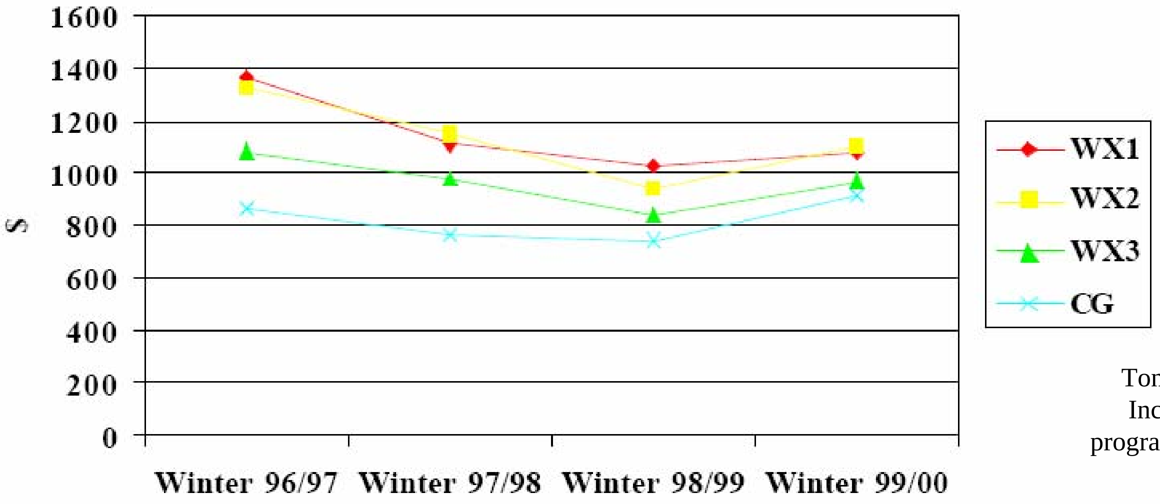


Figure 1. Average Home Primary Heating Fuel Bills per Household

Tonn, B. et al, 2002, Weatherizing the Homes of Low-Income Hoem Energy Assistance Program Clients: A programmatic Assessment, Oak Ridge National Labs Report ORNL/CON-486.

An Interupted Time Series Study

- Small town puts a smoking ban in place
- Reduced incidence of admissions for myocardial infarction observed

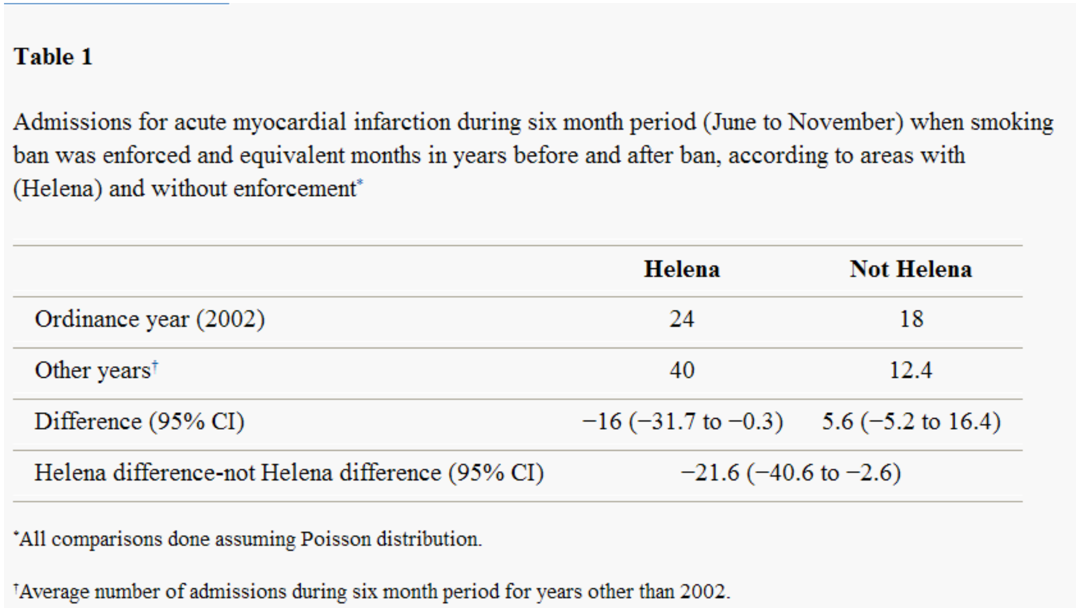


Table 1

Admissions for acute myocardial infarction during six month period (June to November) when smoking ban was enforced and equivalent months in years before and after ban, according to areas with (Helena) and without enforcement[*]

	Helena	**Not Helena**
Ordinance year (2002)	24	18
Other years[†]	40	12.4
Difference (95% CI)	−16 (−31.7 to −0.3)	5.6 (−5.2 to 16.4)
Helena difference-not Helena difference (95% CI)	−21.6 (−40.6 to −2.6)	

[*] All comparisons done assuming Poisson distribution.

[†] Average number of admissions during six month period for years other than 2002.

Sargent, RP, Shepard, RM, Glantzm, SA, 2004,"Reduced incidence of admissions for myocardial infarction associated with public smoking ban: before and after study," *BMJ*. 328(7446): 977–980.

An Educational Study

- Teaching Method A is used by the majority of high school teachers in a district.
- The option of adopting method B is offered and some percentage accept
- Training is provided and some extra pay
- Method B results in X% better pre-test to post-test improvements in raw score

Volunteer Effects

- Ask people to volunteer for a new activity
- Those who volunteer (as compared with those who don't) are on average
  - Higher IQ
  - Younger
  - More approval seeking
  - Different in psychological adjustment
    - If a behavioural study, better adjusted
    - If a medical study, mal-adjusted

Rosnow, R. L. et al., "Volunteer Effects on Experimental Outcomes", *Educational and Psychological Measurement* 1969; 29; 825

Ceiling Effects

Pre-test score of
method B group
40

Pre-test score of
method A group (non-volunteer)
80

Post-test score of
method B group
70

Pre-test score of
method A group (non-volunteer)
95

If the paper reports "the pre-post- test gains for method B were significantly higher" what would you say about that?

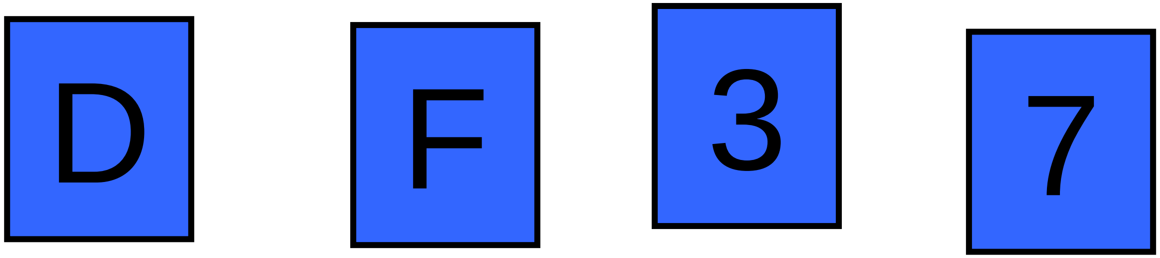
Plan for Today

- Adaptive experimentation
- Quasi-experimental design

➡ Philosophy of Science and Epistemology

Concept Question

- Each card has a letters on one side and a number on the other
- Hypothesis-*if* a card has a D on one side it *must* have a 3 on the other side
- You are a scientist investigating this hypothesis
- You are allowed to turn over any two cards
- Which would you choose to turn over?



1) D&F 2) D&3 3) D&7 4) F&3 5) F&7 6) 3&7

David Hume:
The Problem of Induction

- "We have no other notion of cause and effect, but that of certain objects, which have been always conjoin'd together, and which in all past instances have been found inseparable. We cannot penetrate into the reason of the conjunction. We only observe the thing itself, and always find that from the constant conjunction the objects acquire a union in the imagination."

Hume, David, 1740, *A Treatise of Human Nature: Being an Attempt to introduce the experimental Method of Reasoning into Moral Subjects.*

Nelson Goodman:
The New Riddle of Induction

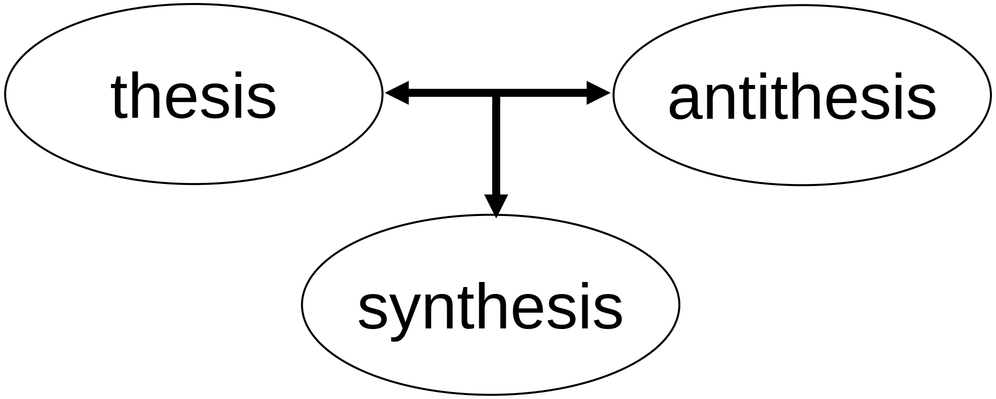
- Some regularities in the world establish habits, some do not
- Proposes the color "grue"
  - Applies to all things examined before a certain time *t* just in case they are green
  - But also to other things just in case they are blue and not examined before time *t*
- But there is a virtuous circle that makes induction work

Goodman, N., 1983, *Fact, Fiction, and Forecast*, Harvard University Press.

Ernst Mach: Phenomenolism

- *Analysis of Sensations* (1885)
- Postulates "elements" such as individual sounds, temperatures, pressures, spaces, times, and colors
- All material things including our own bodies are nothing but complexes of elements that have been constructed by the human mind
- Material bodies do not produce sensations, but rather complexes of sensations are associated together by the human mind to produce material bodies

Hegel's Dialectic



thesis

antithesis

synthesis

- A dialectic of existence
  - – First, existence must be posited as pure Being (Sein)
  - – But pure Being, upon examination, is found to be indistinguishable from Nothing (Nichts)
  - – Being and Nothing are united as Becoming

Hegel, Georg Wilhelm Friedrich. Lectures on the History of Philosophy. London.

Logical Positivism

- “…it began in the 1920’s and flourished for about twenty or thirty years …they were convinced that a genuine contingent assertion about the world must be verifiable through experience and observation.”

  Cambridge Dictionary of Philosophy

- Analytic / synthetic dichotomy. Analytic truths are true (or false) by virtue of some rules of language (including math).

Popper: Falsificationism

- The criterion of demarcation of empirical science from pseudo science and metaphysics is falsifiability.
- The strength of a theory can be measured by the breadth of experimental results that it precludes

Sir Karl Popper, *Logik der Forschung*

Suhs Independence "Axiom"

Maintain independence of the functional requirements

$$\left\{\mathbf{FR}\right\} = \left[\mathbf{A}\right]\left\{\mathbf{DP}\right\} \qquad \text{where} \qquad \mathbf{A}_{i,j} = \frac{\partial \mathbf{FR}_i}{\partial \mathbf{DP}_j}$$

$$\begin{bmatrix} X & 0 & 0 \\ 0 & X & 0 \\ 0 & 0 & X \end{bmatrix} \qquad \begin{bmatrix} X & 0 & 0 \\ X & X & 0 \\ X & X & X \end{bmatrix} \qquad \begin{bmatrix} X & X & X \\ X & X & X \\ X & X & X \end{bmatrix}$$

uncoupled

decoupled

coupled

Acceptable

Avoid

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

Kuhn: Scientific Revolutions

- Paradigm = a set of scientific and metaphysical beliefs that make up a theoretical framework within which scientific theories can be tested, evaluated, and revised.
- “Normal science” = refinement within a paradigm
- “Revolution” = older paradigm overthrown

Kuhn, Thomas S. *The Structure of Scientific Revolutions*.
University of Chicago Press, 1996 (3rd edition).

Lakatos

- Resolves Popper and Kuhn
- Research programs
  - Progressive programs generate bold predictions and useful new work although they may have some counterevidence against them
  - Degenerate programs seek to defend their theory against all evidence and may even do so successfully, but make no useful predictions

Concept Question

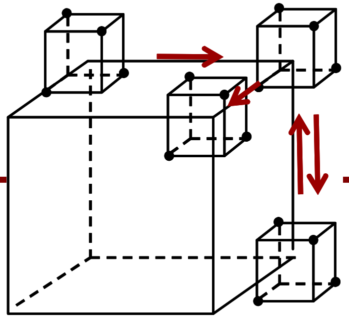
- Each card has a letters on one side and a number on the other
- Hypothesis-*if* a card has a D on one side it *must* have a 3 on the other side
- You are a scientist investigating this hypothesis
- You are allowed to turn over any two cards
- Which would you choose to turn over?

D F 3 7

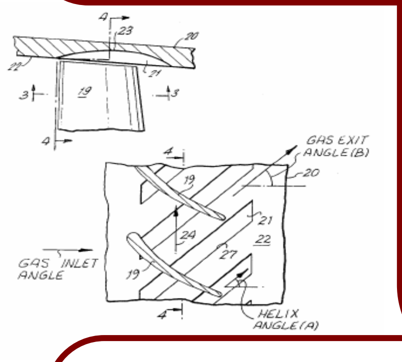
1) D&F 2) D&3 3) D&7 4) F&3 5) F&7 6) 3&7

Overview Research

Concept Design



Outreach to K-12

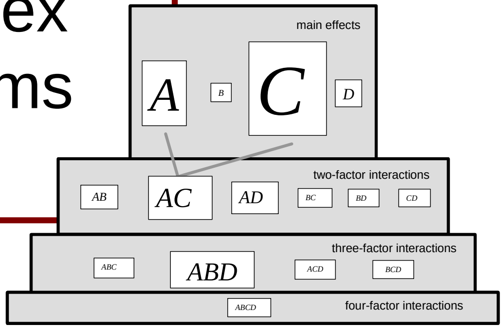


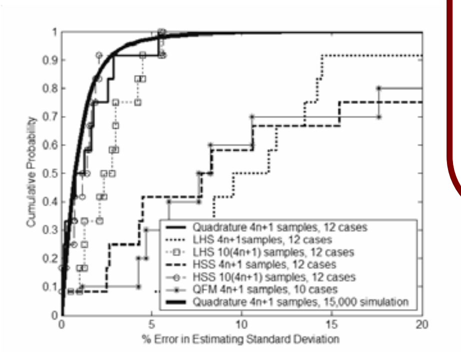
Adaptive Experimentation and Robust Design

$$\Pr\left(\beta_{12}x_1^*x_2^* > 0 \middle| \beta_{12} > \beta_{ij}\right) > \frac{1}{\pi}\binom{n}{2}\int_0^{\infty}\int_{-x_2}^{\infty}\frac{\left[erf\left(\frac{1}{\sqrt{2}}\frac{x_1}{\sigma_{INT}}\right)\right]^{\binom{n}{2}-1}e^{\frac{-x_1^2}{2\sigma_{INT}^2}+\frac{-x_2^2}{2\left(\sigma_{ME}^2+(n-2)\sigma_{INT}^2+\frac{1}{2}\sigma_{\varepsilon}^2\right)}}}{\sigma_{INT}\sqrt{\sigma_{ME}^2+(n-2)\sigma_{INT}^2+\frac{1}{2}\sigma_{\varepsilon}^2}}dx_2dx_1$$

PBS show
“Design Squad”

Complex Systems





Methodology Validation

Next Steps

- Wednesday 2 May
  - Design of Computer Experiments
- Friday 4 May
  - Exam review
- Monday 7 May – Frey at NSF
- Wednesday 9 May – Exam #2