

Lunches with George

Conrad Fung

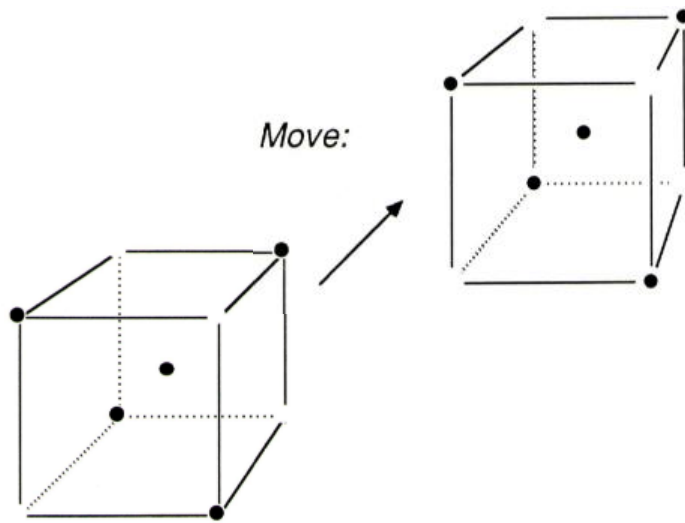
Having the good fortune to be
around at the right moment

To hear whatever he had
on his mind

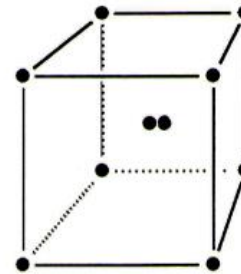
Rapid Fire Phone Calls

- Candidate opening line (rejected):
 - The model, the data, the parsimony
- Ultimate opening line:
 - The model, the data you can't wait to see
 - The theta, beta, sigma, and the rho
 - The Normal, the Poisson, the Cauchy, the t
 - The need to specify what you don't know
 - The likelihood for data you acquire
 - The perspicacious choosing of the prior

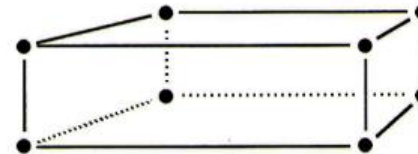
“We’re going to call it Sequential Assembly”



Add Fractions:

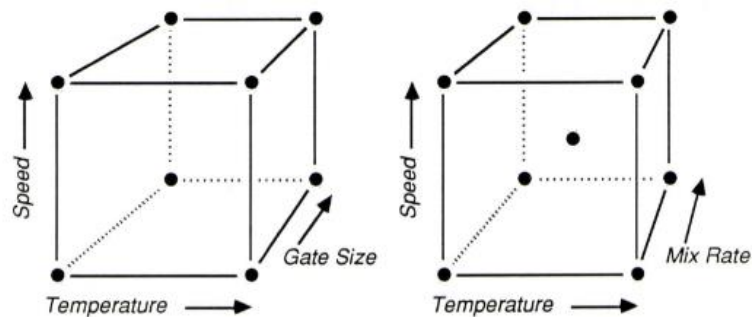


Rescale:

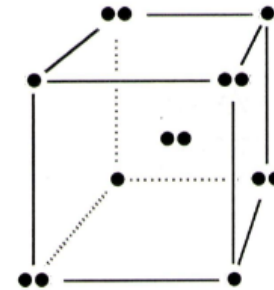


Sequential Assembly continued

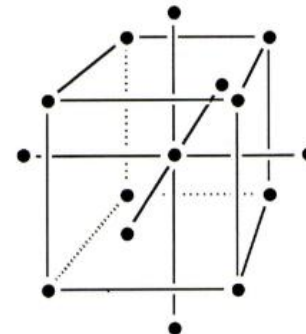
Drop and Add Factors:



Replicate:



Augment:



“Have you been taking out my
commas?”



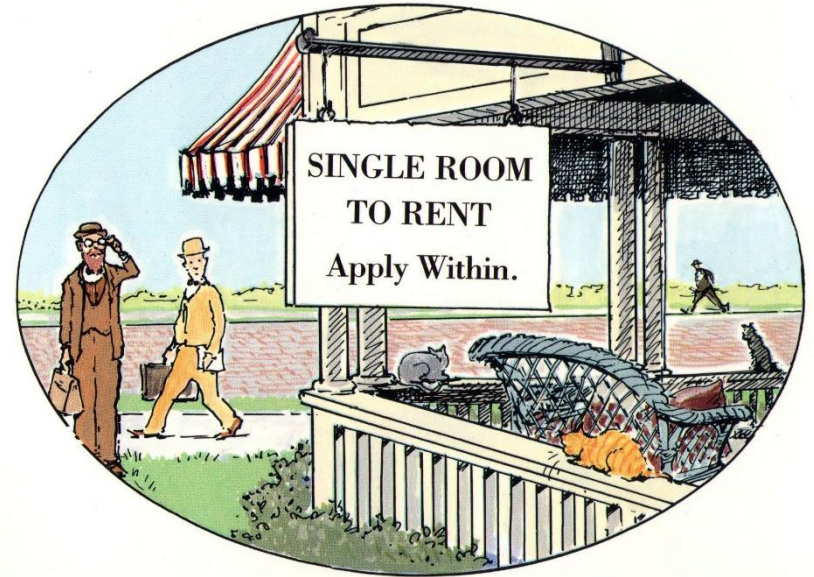
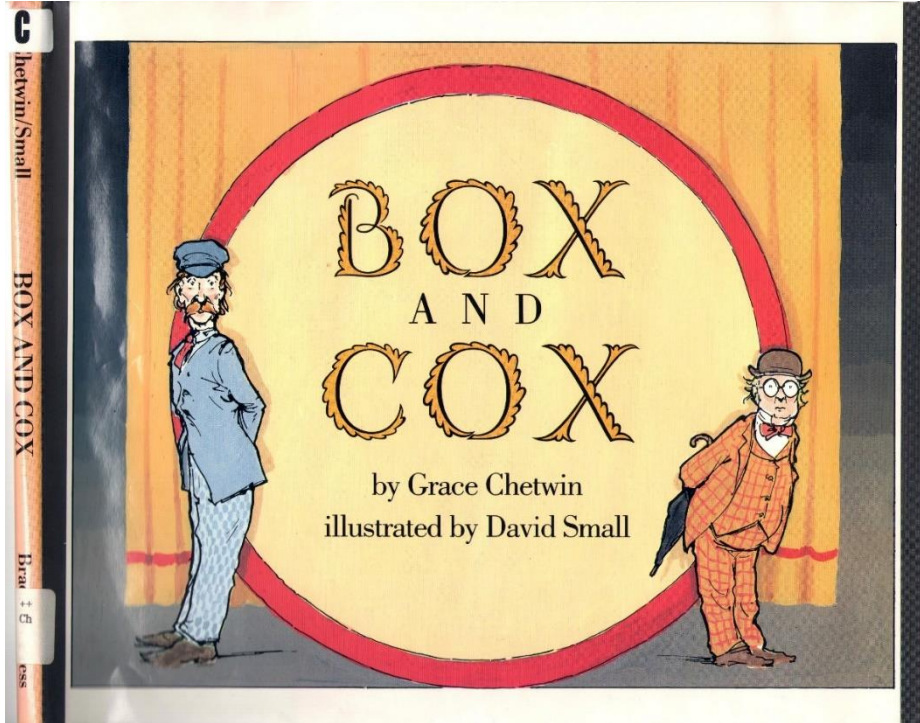
“They have a wonderful author-editor relationship.”

Sampling and Bayes' Inference in Scientific Modelling and Robustness

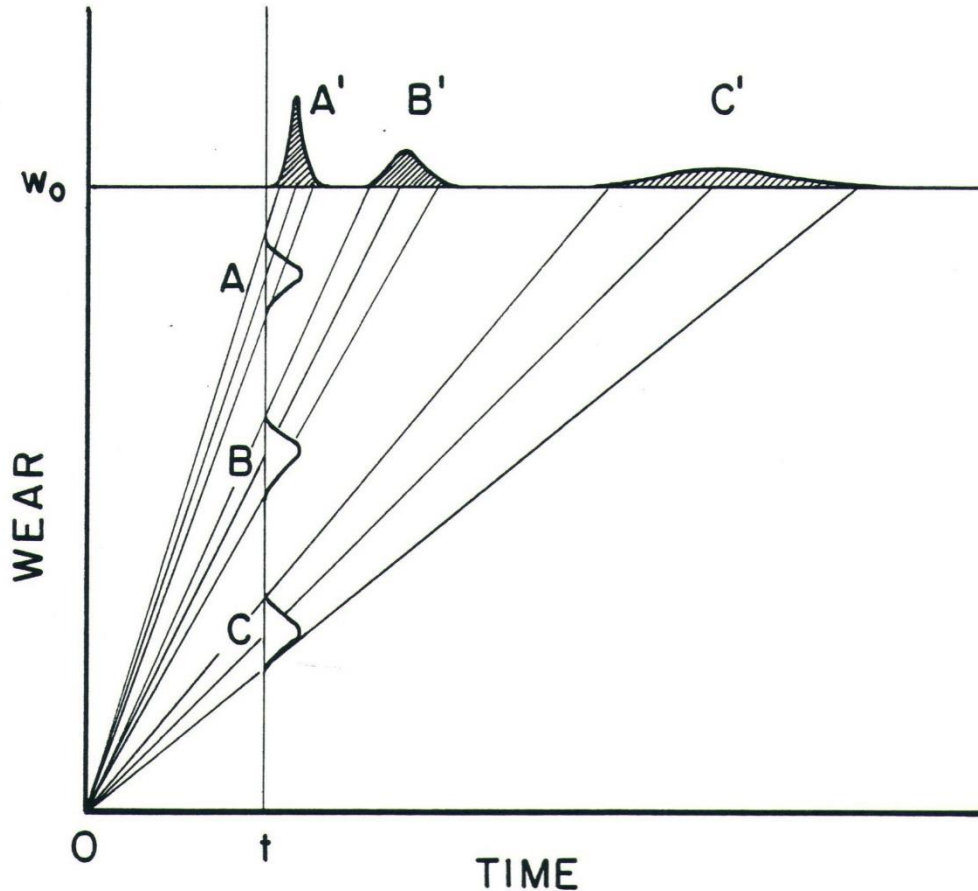
- Induction and deduction:
 - “Although this inductive jump is the only creative part of the cycle and hence is scientifically the most important, the statistician’s role in it may appear inexact and indirect.”
- Statisticians taking refuge in “an imaginary world consisting of only the clean deductive half of the scientific process?”
 - “Deduction alone is sterile – by cutting the iterative process in two you kill it.”
- “The predictive Eve is no longer separated from the Bayesian Adam.”

JRSS A, **143** (4), 1980, pp. 383-430

Insights about Data Transformation



Motivation for a Reciprocal Transformation



The *slopes* represent the *rates* of wear of populations A, B, and C, which should be the underlying quantities of interest.

But conditioning on Time (which is how data is often obtained) shows *amounts* of wear experienced by each population up to that time;

and conditioning on w_0 shows the time to reach this critical level experienced by these populations.

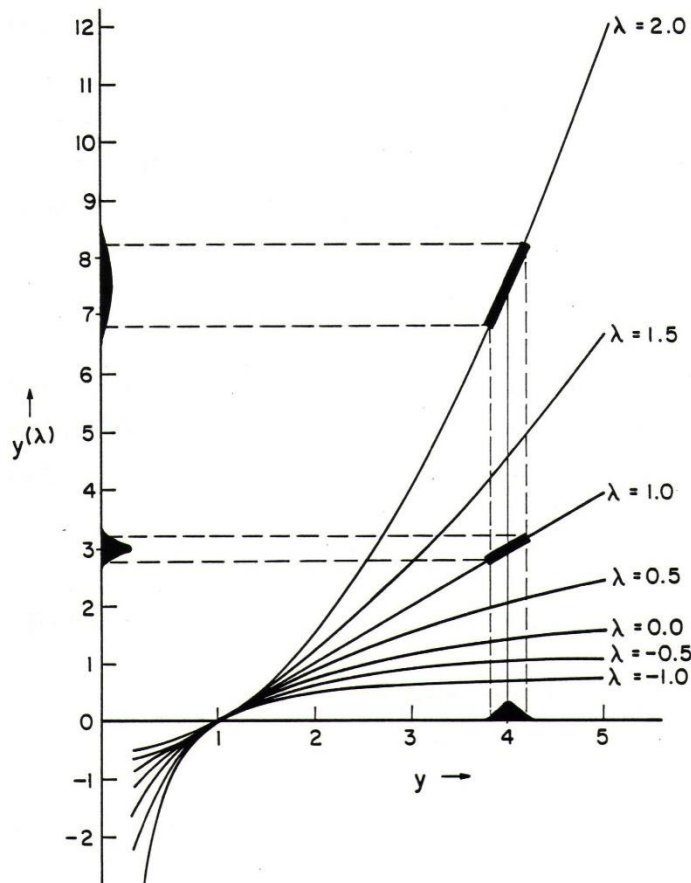
Unfortunate Misunderstanding

- An Analysis of Transformations (1964), Box and Cox
 - They had it covered

$$y^{(\lambda)} = \begin{cases} \frac{y^\lambda - 1}{\lambda} & \lambda \neq 0 \\ \ln y & \lambda = 0 \end{cases} \quad z^{(\lambda)} = \begin{cases} \frac{y^\lambda - 1}{\lambda y^{(\lambda-1)}} & \lambda \neq 0 \\ y \ln y & \lambda = 0 \end{cases}$$

- An Analysis of Transformations Revisited (1981), Bickel and Doksum
 - They claimed Box-Cox type procedures are **unstable** ... the estimates $\hat{\lambda}$ and $\hat{\theta}$ can be highly correlated when estimated jointly, so that the marginal variances of the $\hat{\theta}$ s can be inflated by large factors ...

An Analysis of Transformations Revisted, Rebutted (1982) by Box and Cox



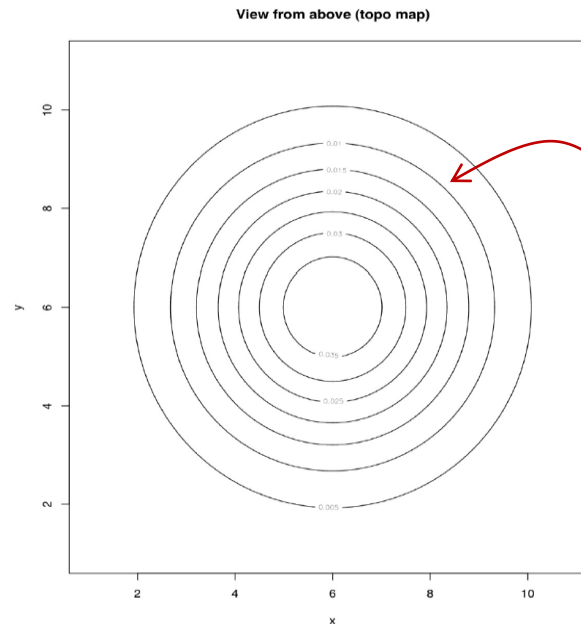
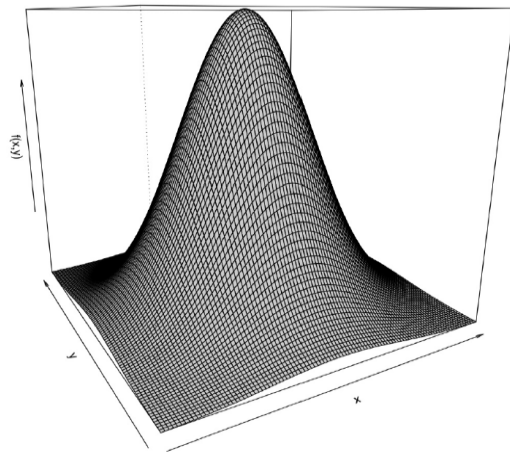
Scale changes are induced by different λ s (in the “y” version of) transformed data, hence the need to take proper account of the Jacobian in transformation analyses.

The “z” version of the transformation takes care of this.

$$z(\lambda) = \begin{cases} \frac{y^\lambda - 1}{\lambda y^{(\lambda-1)}} & \lambda \neq 0 \\ y \ln y & \lambda = 0 \end{cases}$$

Generating Normal Deviates Without Powerful Computing

- Piecewise approximation of a single normal curve is complicated.
- George's insight while still a student: consider two independent normal deviates:



Box-Muller Transformation

- If U_1 and U_2 are independent Uniform (0,1), then X_1 and X_2 are independent Normal (0,1).

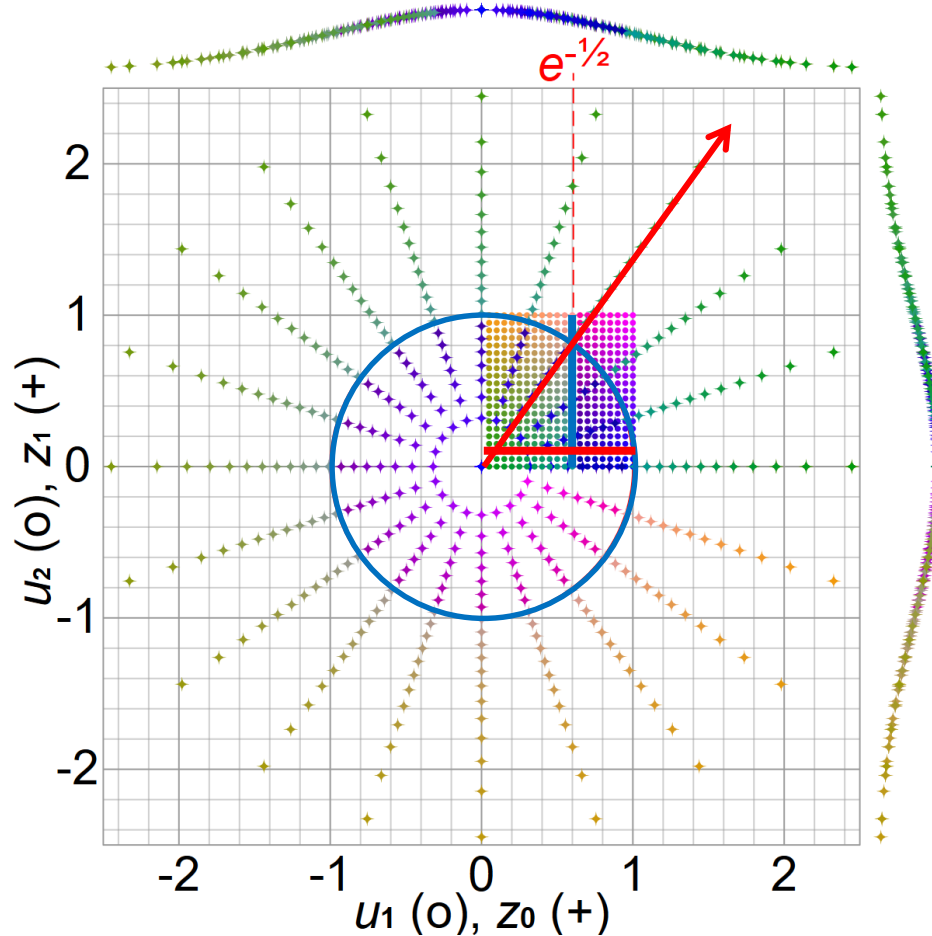
$$\begin{aligned} X_1 &= (-2 \log_e U_1)^{1/2} \cos 2\pi U_2 \\ X_2 &= (-2 \log_e U_1)^{1/2} \sin 2\pi U_2 \end{aligned}$$

These generate coordinates of a circle with radius 1 as U_2 goes from 0 to 1.

↑
If U_1 is uniform on (0,1), then $-2\text{LN}(U_1)$ is Chi-square with 2 d.f. But so is $X_1^2 + X_2^2$.

Ann. Math. Stat., **29** (2), 1958, pp. 610-613

Visualization of Box-Muller Transformation (Wikipedia)



Vertical lines in the $(0,1)$ square map to circles in the z -space as U_2 varies from 0 to 1.

Points on horizontal lines in the $(0,1)$ square map to points on rays extending outward from the origin of the z -space as U_1 varies from 1 to 0.

I was away at DuPont for a few years, wondering if I would ever make it back to Madison. George kept in touch via cassette tapes.

A particularly kind one brought me back.

In the meantime, much had changed in the world of statistics.

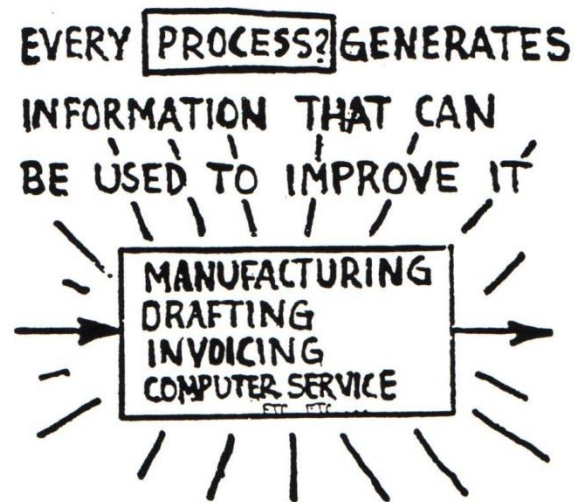
Quality

The “quality movement” was as if custom-made for George

- George was *ready*
- A confluence of his favorite topics
 - Philosophy of learning
 - Design of Experiments
 - Time series, forecasting, and control
 - Robustness
 - Democracy

You can improve almost anything

- Dinner
- Statistical at its heart
 - Embrace variability



“Information Field”

The Era

- 1983: AT&T Bell Labs
 - Blau Godfrey, Raghu Kacker, Anne Shoemaker, Ramón León, Vijay Nair, Madhav Phadke
 - First Quality and Productivity Research Conference
- Taguchi methods
 - “... promises to be a long debate ...”
 - *UPI reporter at the conference*
 - “The difference is they’re doing it and we’re not”
 - *GEPC*

A Polite Acknowledgment

- “We owe to Taguchi our present awareness of the importance of statistics in achieving robust processes and products in industry.”

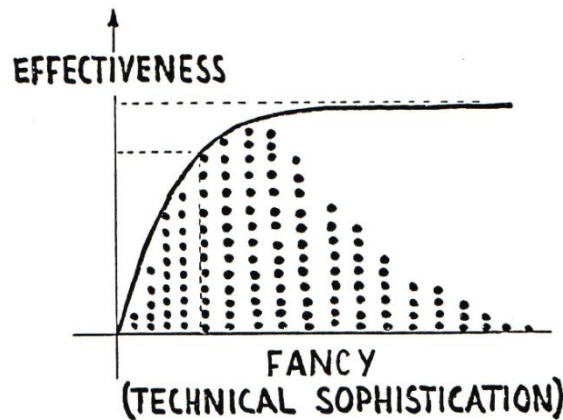
– *GEPB*

Controversy

- Threats from loud, partially informed voices
 - “We don’t need statisticians running around the country debating ideas and theory”
- George’s reply:
 - If we only follow, we will always be behind.
 - If we are to do better, we ought not tie ourselves to any particular subset of procedures.
 - *Investigations should be problem-driven, not method-driven.*

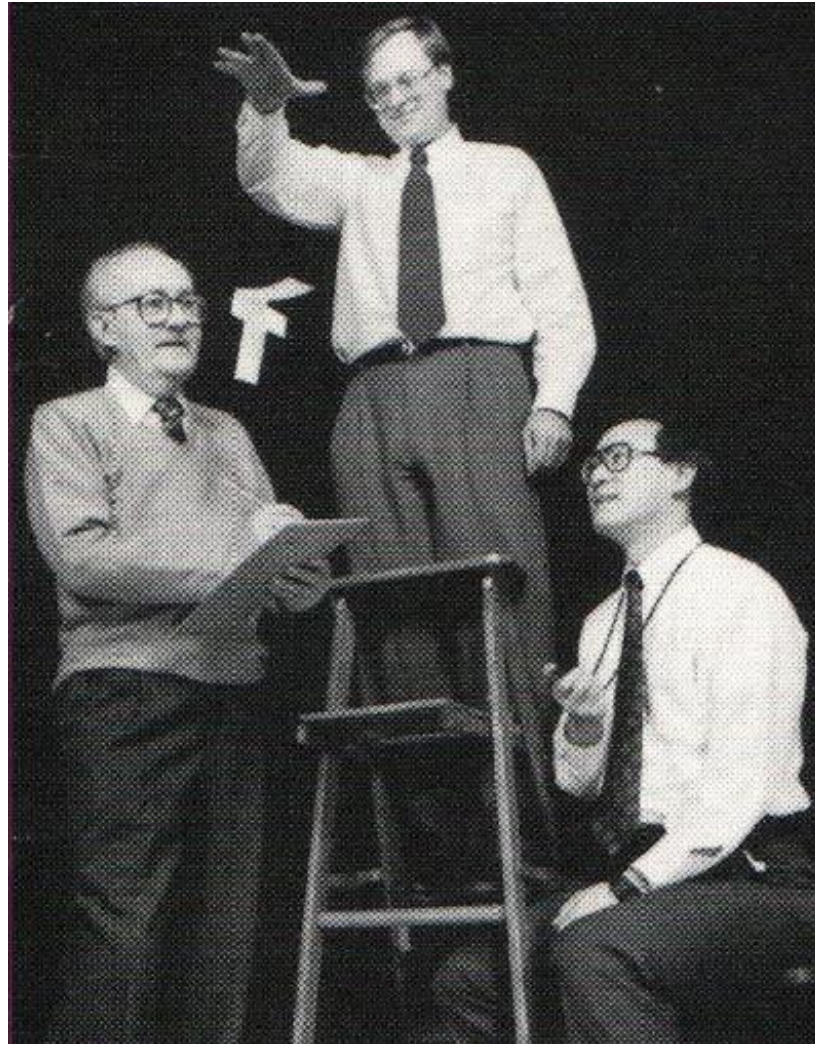
The Era

- Democratization of science via the simplest tools:



STATISTICS IS
TOO IMPORTANT
TO BE LEFT TO THE
STATISTICIANS
(ONLY)

“Engineer’s Key to Quality”



... in the days of overhead
transparencies

How do we find things out?

CRITICAL
EVENT



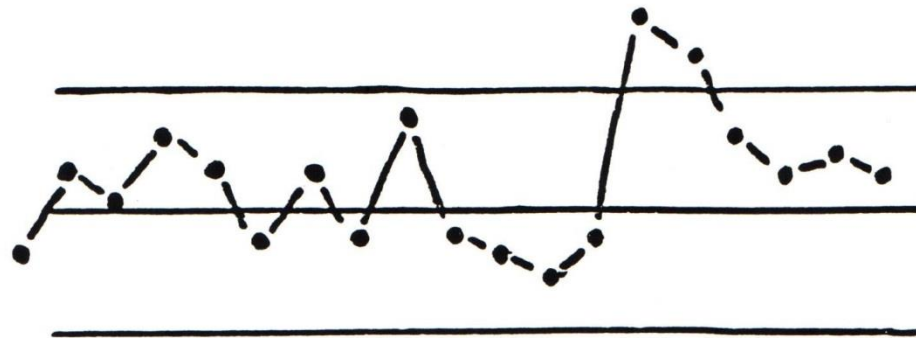
PERCEPTIVE
OBSERVER



Informed Observation

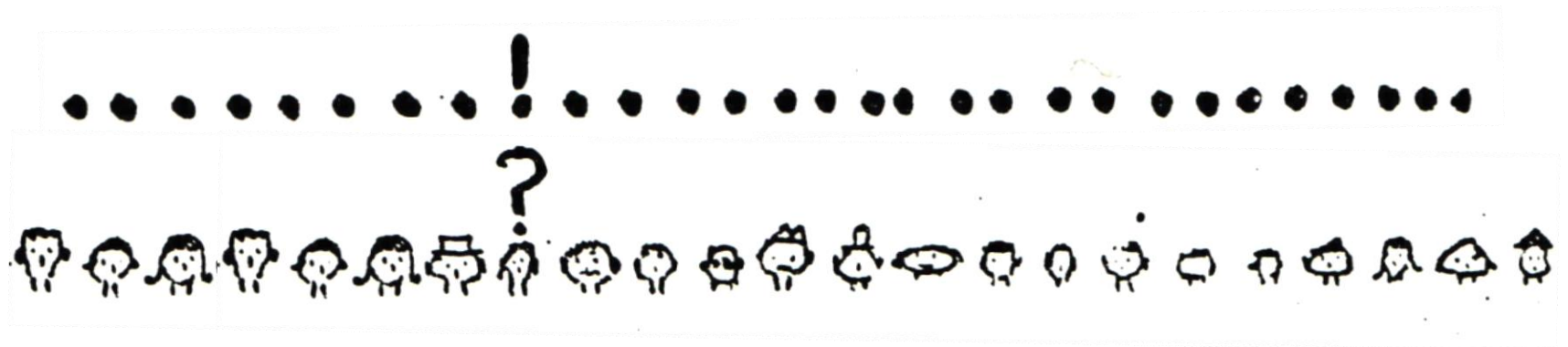


?



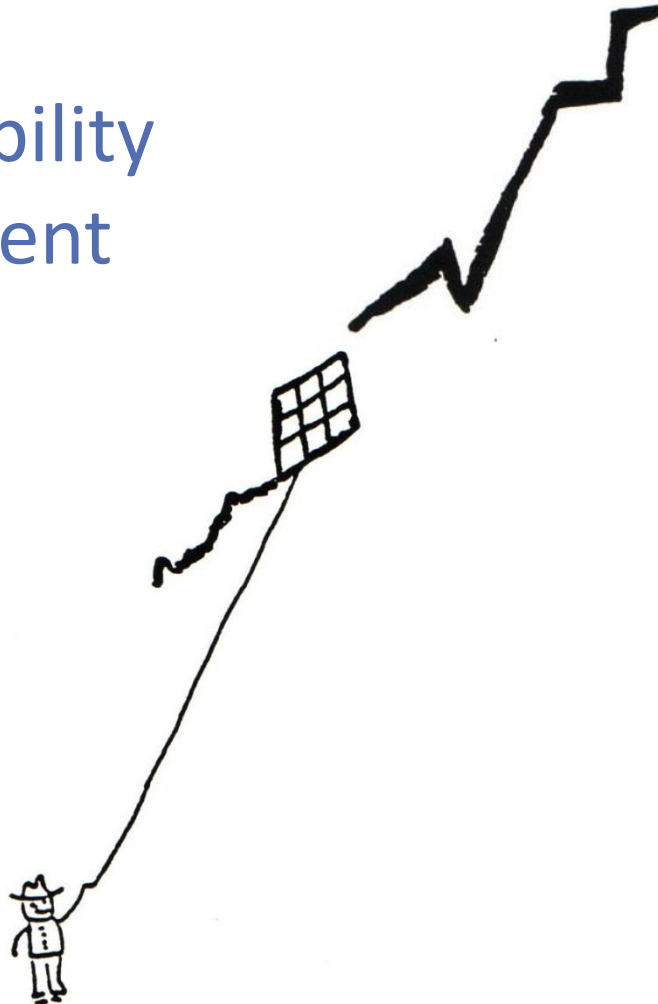
?

Spark to the Tinder



Directed Experimentation

Increases the probability
of an informative event
occurring



Cumulative Learning

$$\begin{array}{ccc} \text{KNOWN} & & \text{UNKNOWN} \\ y = f(\tilde{x}_f) + e(\tilde{x}_e) \\ & \uparrow & \text{.....!} \end{array}$$

“George’s Column” in *Quality Engineering*

- Do interactions matter?
- Must we randomize our experiment?
- Good quality costs less? How come?
- Teaching engineers experimental design with a paper helicopter
- Comparisons, absolute values, and how I got to go to the Folies Bergère

When Murphy Speaks – Listen

When Murphy Speaks—Listen

Can it be? Is
Murphy's Law
good news,
not bad?

MURPHY'S LAW SAYS THAT "ANYTHING THAT CAN GO WRONG WILL GO WRONG." This is usually regarded as bad news. If you think about it though, it's not really bad news, it's good news. Nothing is perfect and Murphy's Law says that the day-to-day operation of a system itself can help tell us what's wrong with it. The catch is that it will only tell us *if we listen*. If we don't listen, the bug that's in the system will cause the same glitch to happen again and again. The inevitable result is frustration and poor quality. But if we listen and take appropriate action, we might get rid of the bug forever.

Another way of saying this is that every operating system supplies information on how it can be improved. If we use that information, it can be a source for continuous improvement. Figure 1 illustrates this with a system radiating information much like a radio transmitter radiates electromagnetic waves.

What is a system?

But what do we mean by a system? The important example that immediately comes to mind is an industrial manufacturing process. But even in industry many systems are not concerned with manufacture but with invoicing, dispatching, billing, etc. It is important to remember that it is just as essential to get the bugs out of *those* systems as out of the process of manufacturing itself. But the kind of a system I'm talking about might have nothing to do with manufacturing. It might be baggage handling in an airline, student registration in a college, or health care delivery in a hospital. If one listens to Murphy, all these systems can be improved.

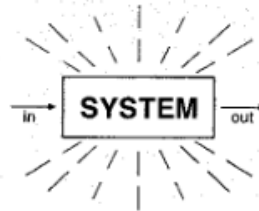
Is anyone listening?

So why isn't Murphy listened to more often? Or, to put it the other way, why isn't the message that is trying to tell us how to improve the system put to use? A common difficulty is that the people closest to the system might have:

- no expectation that the system could be better.
- no understanding of how to make it better.
- no power to change it.

When I discussed this with some friends recently, some stories came up:

Figure 1.
Every operating system supplies information on how it can be improved.



• Friend A told me how he had recently had a medical check-up at the hospital. This involved, in particular, a number of lab tests. Two weeks later in the follow-up appointment with his doctor, it turned out that the results of the tests had so far not been entered in his records. His doctor smiled and said rather triumphantly, "Don't worry. *They* often do this, so I keep a duplicate of the test results myself, although I'm not supposed to."

• Friend B told how, the previous week when he had flown across the country, the airline had lost his baggage, which contained the slides he needed to make an important presentation. They were not returned to him for two days. When he reported the loss at the airline's lost baggage claim, the clerk said, "Oh yes, you were on that flight where you have to change planes in St. Louis. People are always losing luggage on that connection because the time between planes is so short and the gates in St. Louis are such a long way apart. I'm sorry but that's the way *they* schedule them."

• Friend C, who works for a university, mentioned that his department secretary had told him "Don't ever circulate stuff through the campus mail. *They* take forever to get it delivered."

In these stories, the physician, the baggage clerk, and the department secretary and indeed my three friends themselves all recognized that the systems they were dealing with were faulty but believed they were totally powerless to do anything about them.

The author's
name is missing
from the title
page in the
published
version.

Nurturing

- *Howards End:*

“Life is indeed dangerous, but not in the way morality would have us believe. It is indeed unmanageable, but the essence of it is not a battle. It is unmanageable because it is a romance, and its essence is romantic beauty.

Margaret hoped that for the future she would be less cautious, not more cautious, than she had been in the past.”

Among the many quotes George had taped to his wall over the years

A Lighthearted Memoir

George's memoir was a lifetime in the making, but the effort accelerated in the later years as he collected and organized his memories with the help of Claire and Judy Allen.

George found joy in his friends and treasured them dearly. Reminiscing gave him much pleasure, and the collection is now his gift to us.

Alice in Wonderland

- ... was George's favorite book.
- Every chapter in his memoir begins with a quote from *Alice*.

Among his papers after his passing was a handwritten note with one further quote (not in the memoir) ...

Never imagine yourself not to be
otherwise

