

eXtreme Measurement: recognizing, understanding and avoiding measurement dysfunction

(or 'Why software measurement fails, and what to do about it')

UKSMA London, October 2009

C. C. Shelley
OXFORD SOFTWARE ENGINEERING Ltd
9 Spinners Court, 53 West End,
Witney,
Oxfordshire
OX28 1NH
www.osel.co.uk
shelley@osel.netkonect.co.uk
Tel. +44 (0) 1993 700878

Objective

- **I believe that good, sound measurement has the potential to deliver useful information. This can be critical to decision making and direction actions.**
- **But I also know that unsound measurement can mislead, and distort decisions and activities.**
- **The objective of this presentation is to explain a model of unsound (dysfunctional) measurement and describe how you can use it to recognize and eliminate it.**

Contents

- 1. Austin's model of measurement dysfunction**
- 2. Recognizing dysfunction**
- 3. Eliminating dysfunction**

'You can't control* what you can't measure'

Tom DeMarco

'In my experience absolutely everyone who does it screws it up. So I think the concept's wrong. And I think it's kind of pointless to think, "Well, if they did it right, it would be OK." They don't do it right. The people who want to do it are inclined to do it wrong.'

Tom DeMarco

Two measurement stories...

- **Bethlehem Steel**
 - 78, 80, 85...

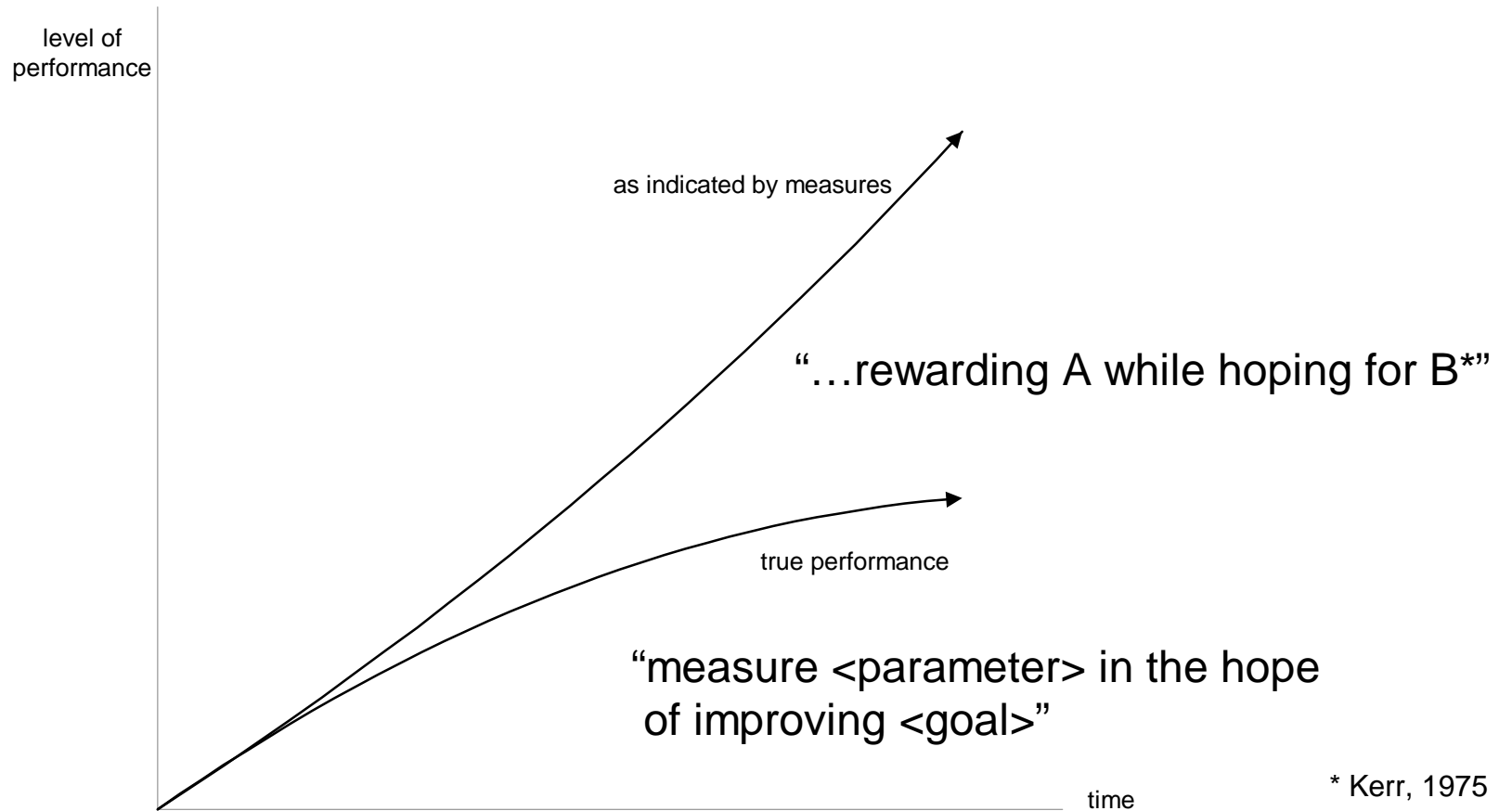
- **TWA.**
 - from loss to profit

– *From R. D. Austin: see www.osel.co.uk/library.htm, # 48*

Austin's* model of measurement dysfunction...

- **Describes the human aspect of software measurement**
- **Think how people (including you) usually react as we work through this model...**

* Ref in paper



- **This is usually treated with assumed pragmatism:**
 - **proxy measures are cost effective**
 - **they are a reasonable compromise**
 - **it's the best we can do**
 - **what alternative is there?**

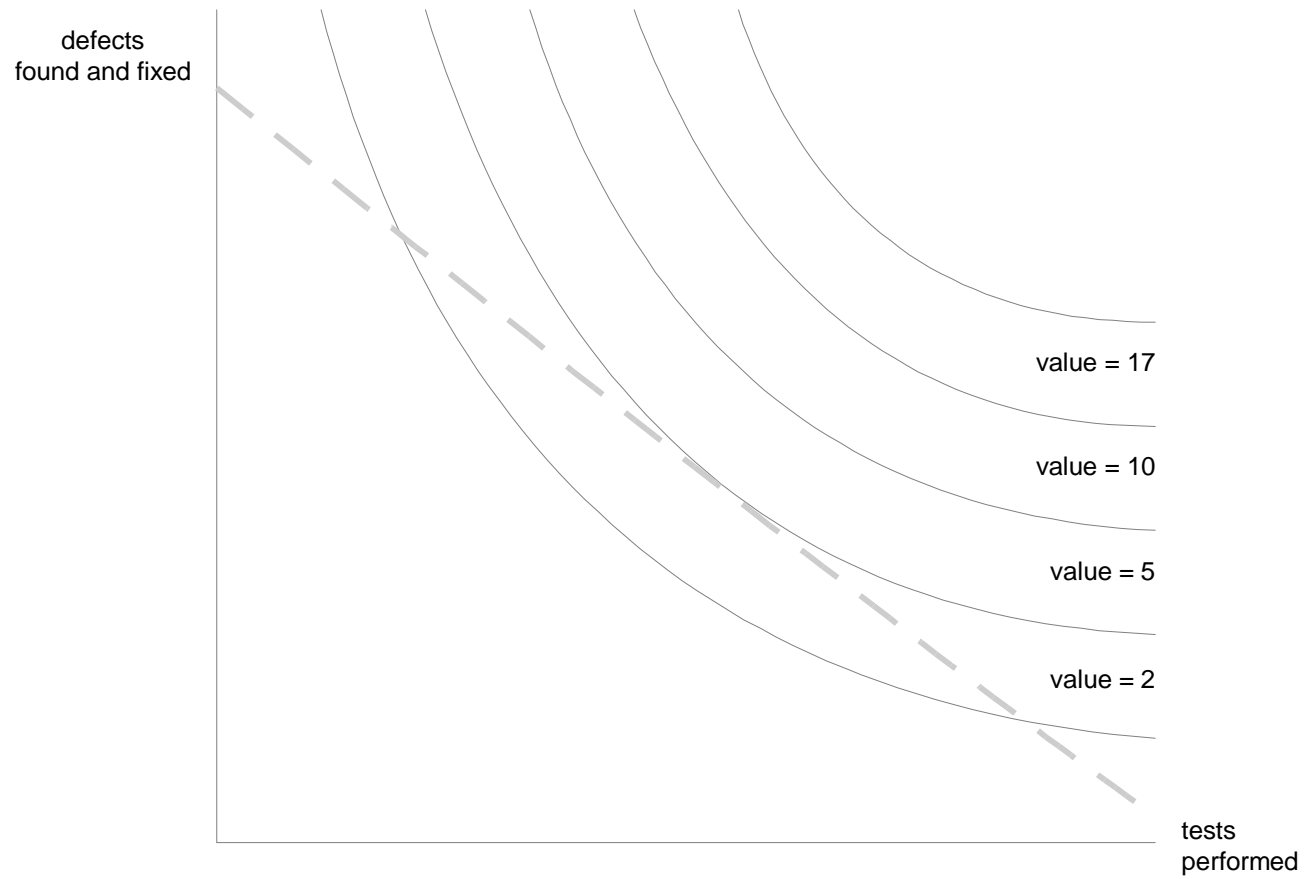
- **And will trigger dysfunctional measurement and behaviours...**

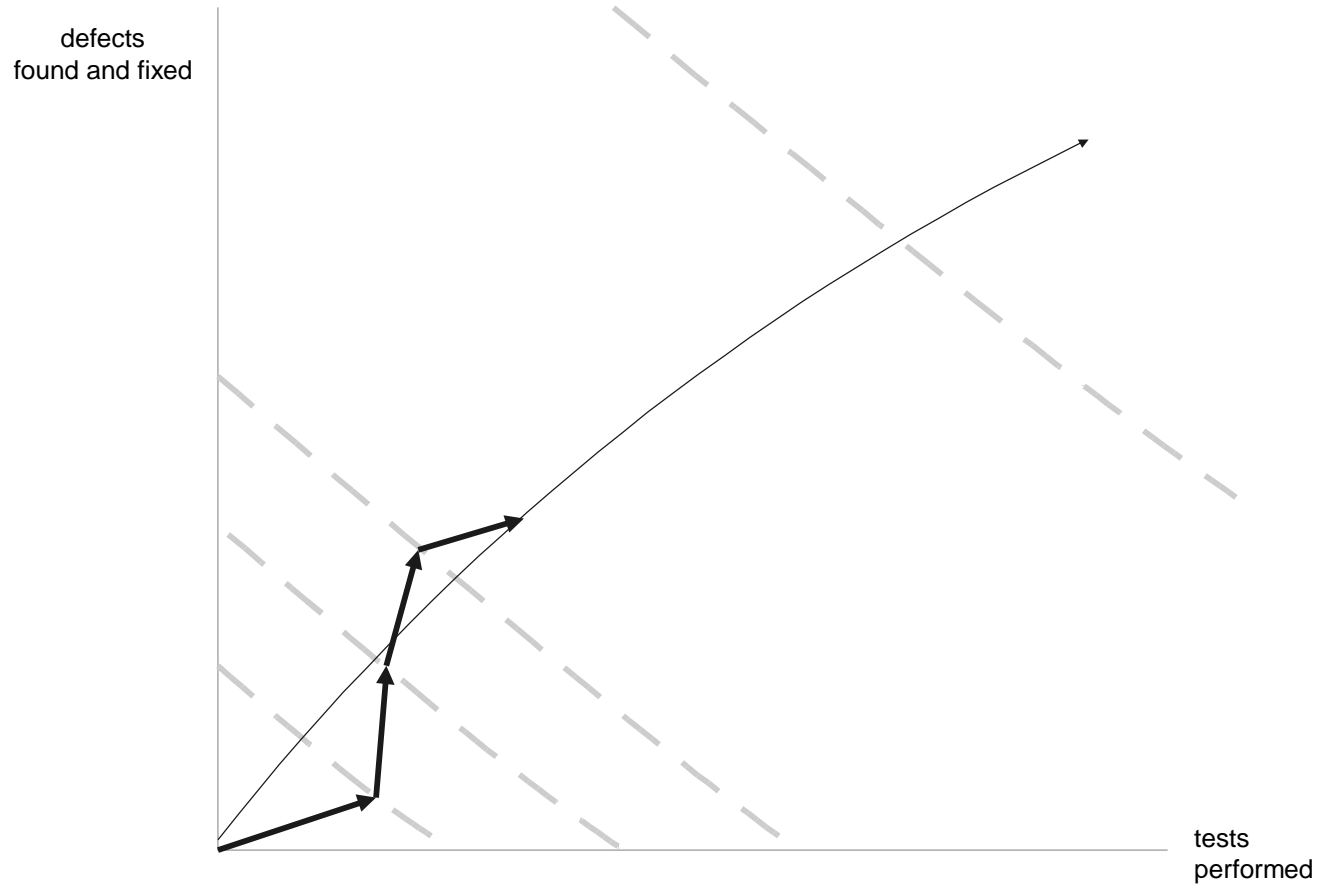
The two categories of measurement

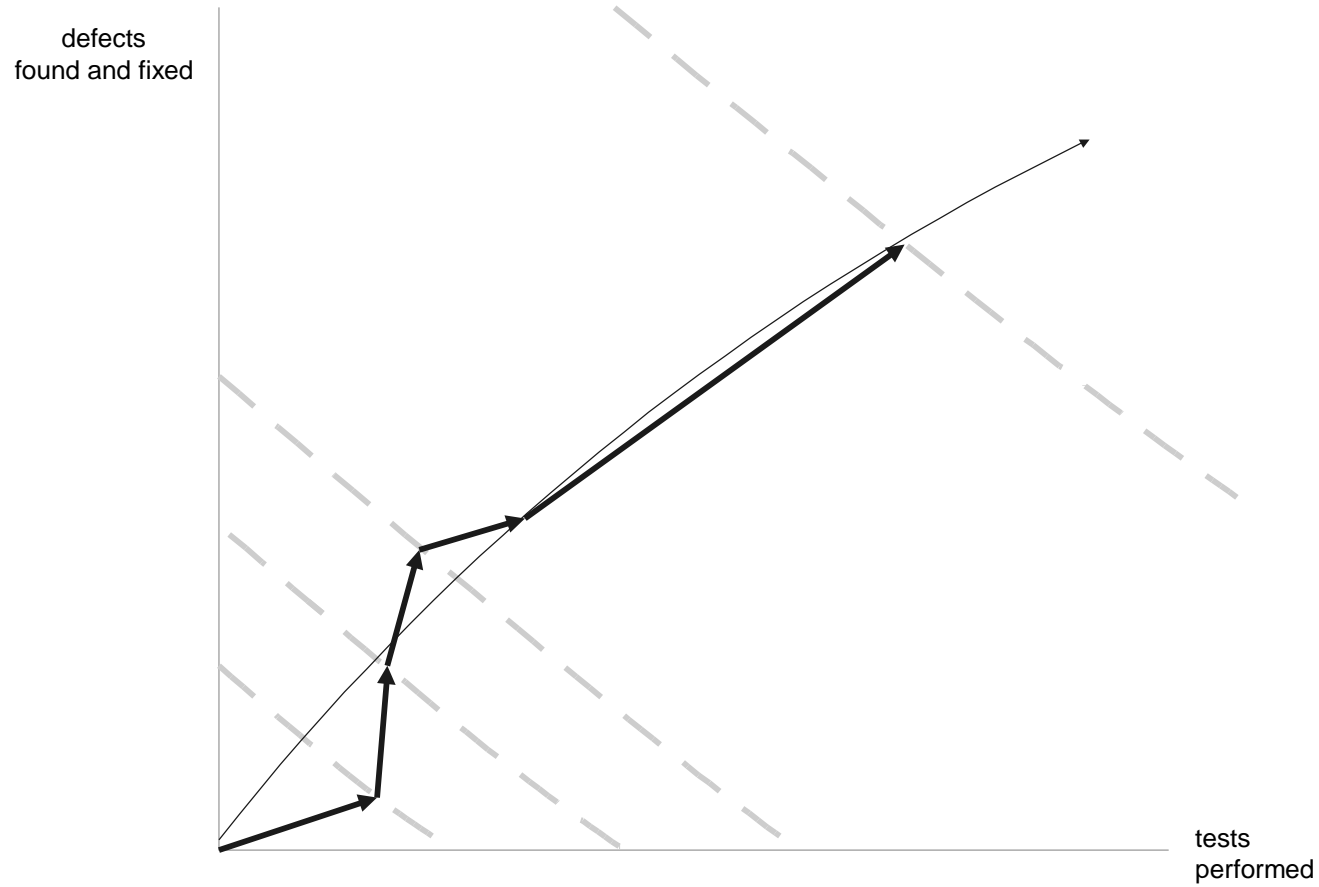
- **Motivational (measure to demonstrate)**
 - Measured outcomes receive rewards or bonuses
 - Theoretically encourages compliance with plans
 - Specifically intended to change that being measured
- **Informational (measure to understand)**
 - Day to day tactical, operational use
 - For long term process refinement
 - Should not change that being measured (need to know how things really are)
- **These should be distinct**
 - but difficult to segregate
 - category is determined by the role of the user

How does dysfunction happen?

- **Consider a team testing a system and fixing defects...**
- **...logging tests and fixes**







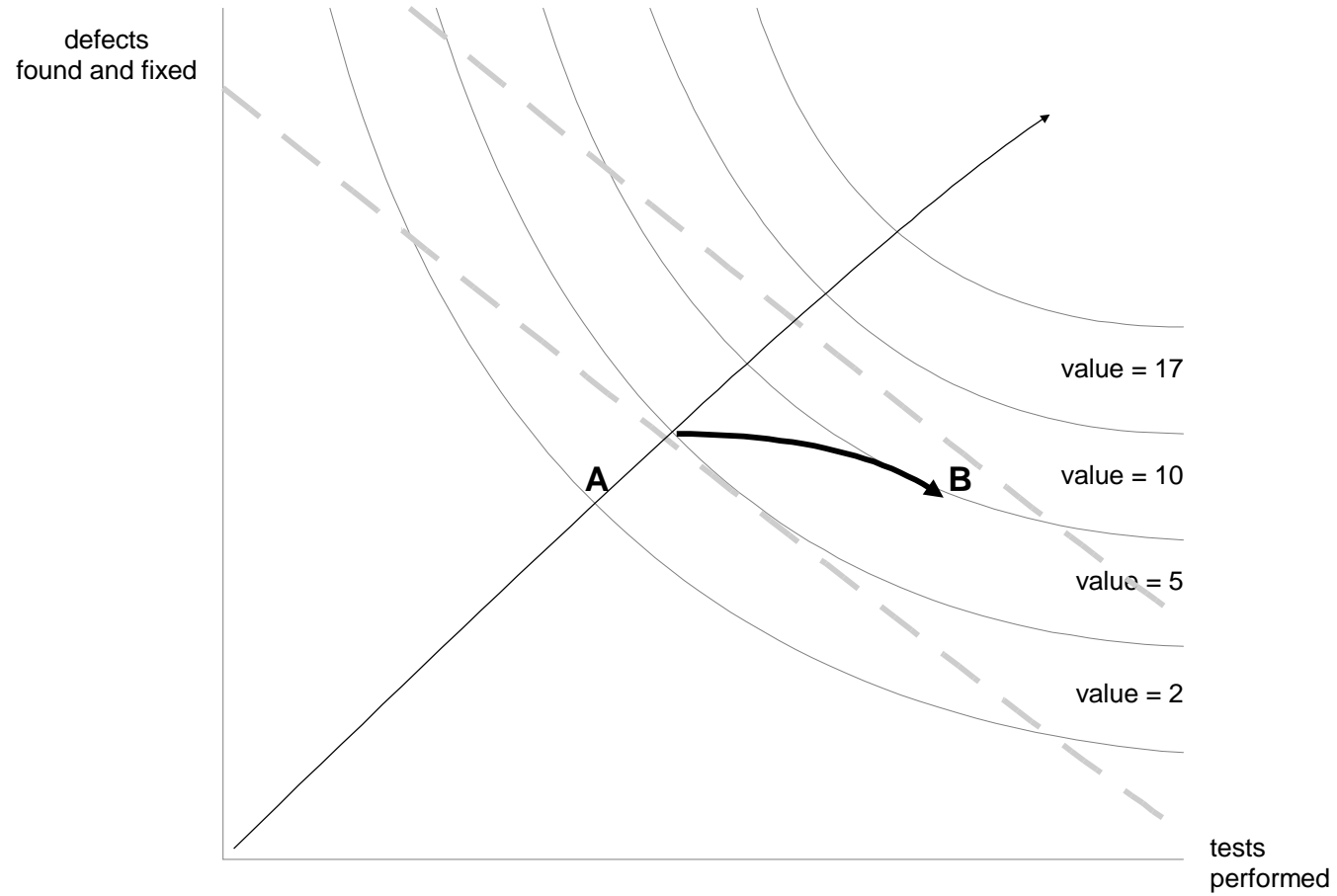
Measurement can range from...

- **Unsupervised**
- to
- **Fully Supervised**

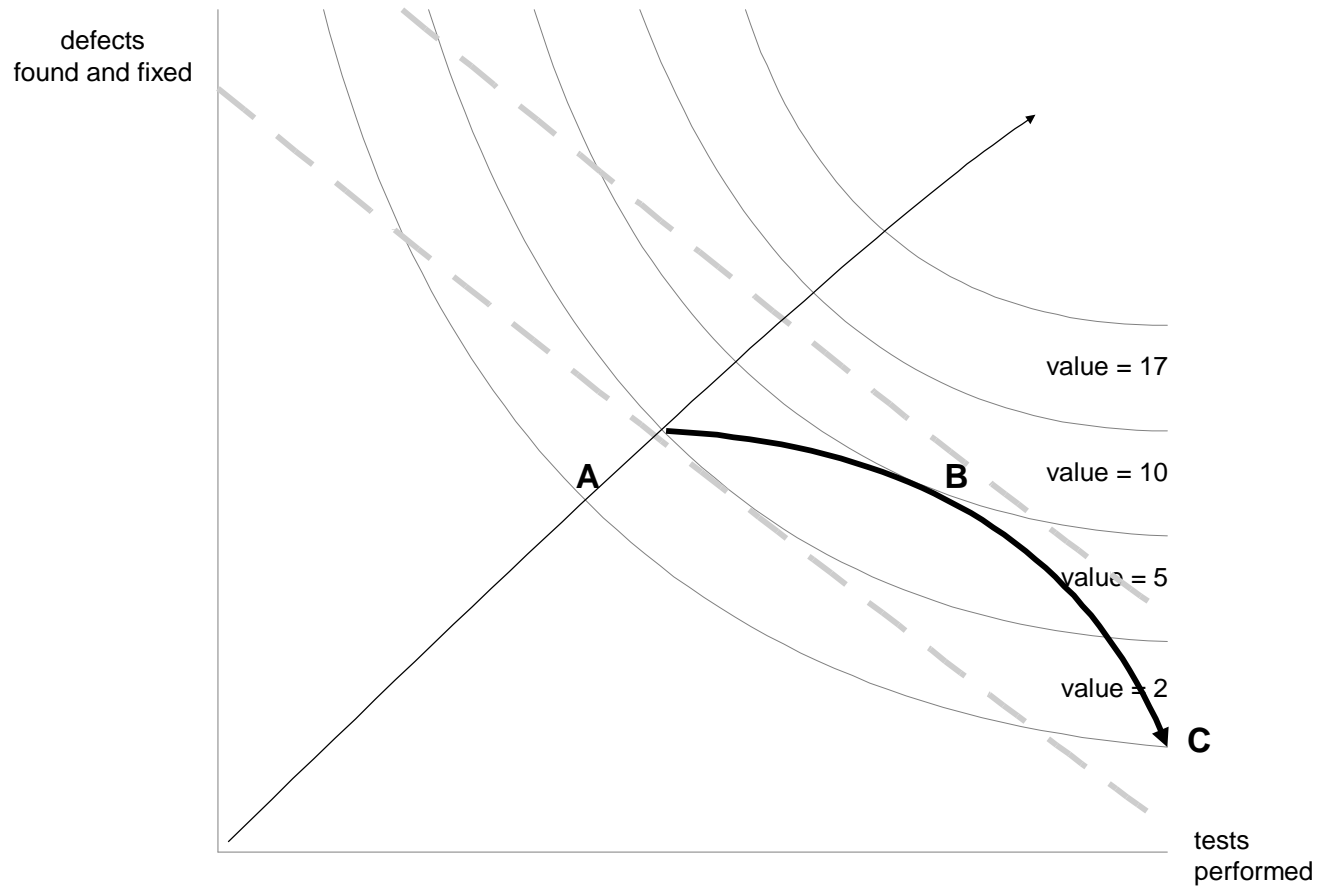
The more usual, intermediate situation...

- **Partially supervised...**
- **...for example, look at the critical element - tests performed (the test schedule)**
 - **and defects will be found and fixed (won't they?)**
- **What will happen?**

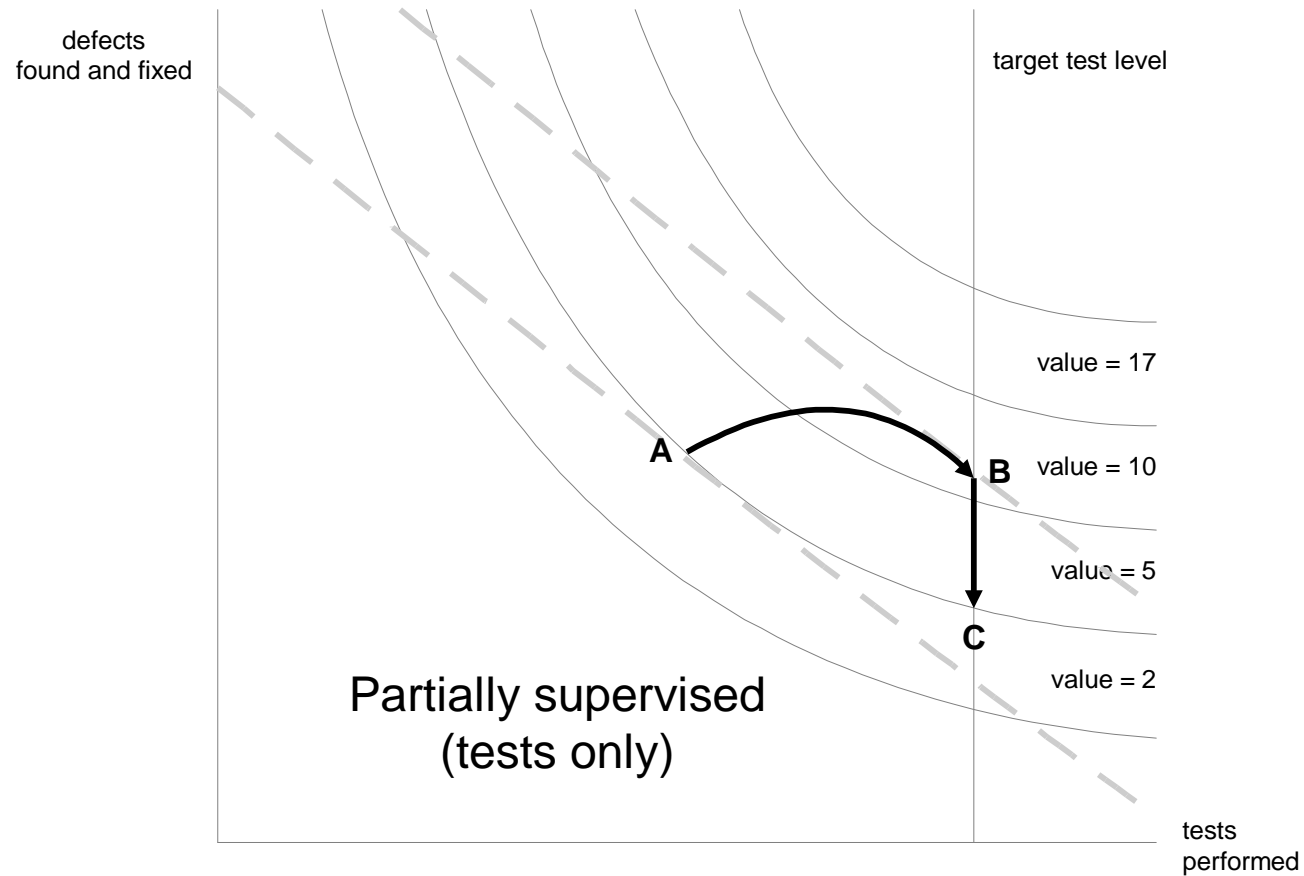
Still functional...



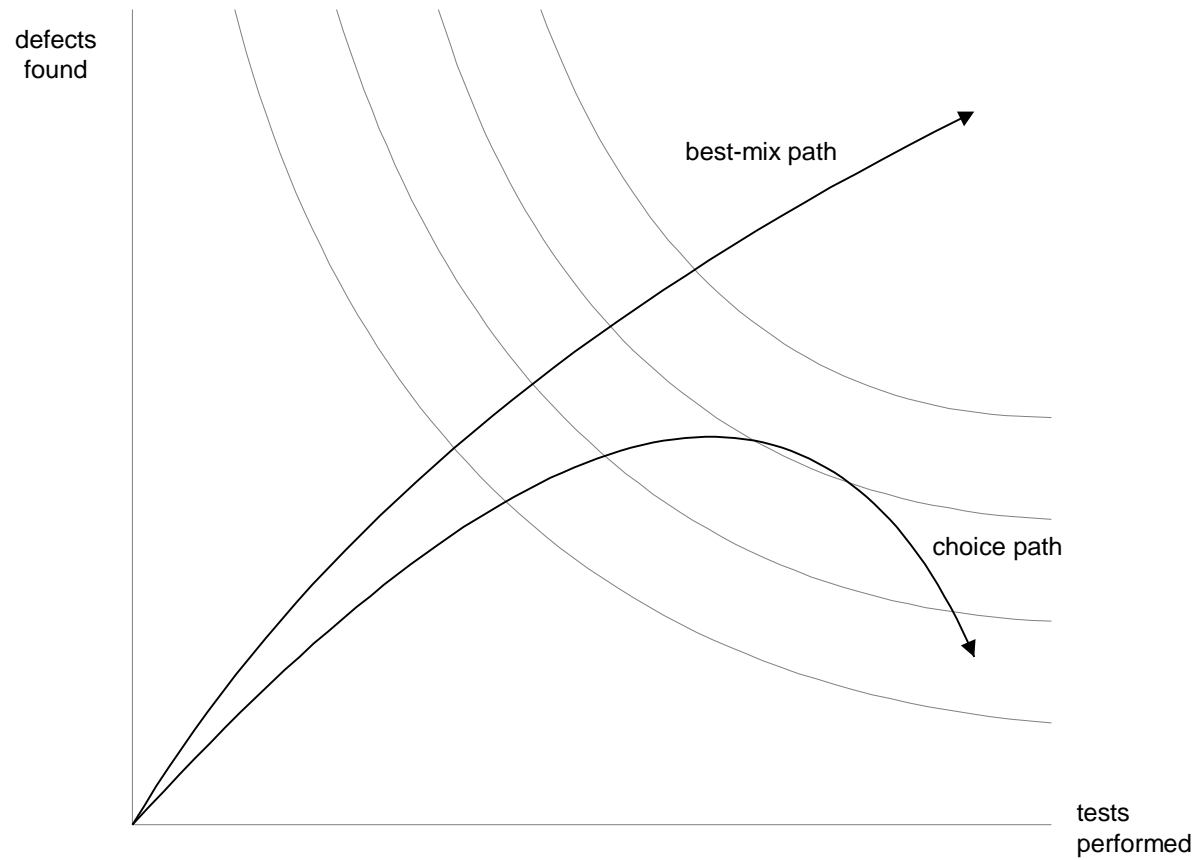
...Now dysfunctional



With the encouragement of performance targets...



...with encouragement



- **We have looked at partial supervision - measuring tests only**
- **Consider partial supervision if only defects fixed were measured...**

	Unsupervised	Partially supervised	Fully supervised
Motivational		No	Yes
Informational	Yes		

Full supervision

- **Is this feasible in complex self aware organizations?**
- **Consider the test/defect model...**
- **...other factors influencing value delivered**
 - **Effort to test**
 - **Time to test**
 - **Test effectiveness**
 - **Effort to fix**
 - **Time to fix**
 - **Definition of a defect ('it's not a defect, it's a feature')**
 - **Defect severity**
 - **Defect priority**
 - **Quality of Fix**
 - **...**

11a “Eliminate numerical quotas for the work force... .. A quota is a fortress against improvement of quality and productivity. I have yet to see a quota that includes any trace of a system by which to help anyone do a better job. A quota is totally incompatible with never ending improvement. There are better ways.”

W. Edwards Deming

11b “Eliminate numerical goals for people in management. Internal goals set in management of a company, without a method, are a burlesque... .. Focus on outcome is not an effective way to improve process or an activity... .. management by numerical goal is an attempt to manage without knowledge of what to do, and is in fact usually management by fear.”

W. Edwards Deming

“Ownership of measures should be with the people who do the work.”

“Managers tend to think they ‘own’ the numbers. They should not.”

“People do what you count, not necessarily what counts.”

“Attention to output can increase costs.”

“Using measures for improvement starts with thinking differently.”

John Seddon ‘I Want You To Cheat’

The nature of measurement

- **Informational measurement is fragile**
- **Motivational measurement is very prone to dysfunction**
- **There is a limited ‘information horizon’ in software development (and other complex knowledge working environments)**
- **Recognition and exploitation of the information horizon can enable s/w development (and test) optimization**

Two steps to good measurement (one easy, one hard)

- 1. Recognize dysfunctional measurement**
- 2. Eliminate it**

1. Recognizing dysfunctional measurement...

- **Investigate**
- **Keep the specifics confidential**
- **Work with developers, testers, team leads, PMs, line managers and board to find answers* to the following questions:**

...1. Recognizing dysfunctional measurement...

- 1. What measurement data do you collect?**
- 2. Why?**
- 3. How do you use it?**
- 4. What decisions or actions do you take based on your analysis of the data?**

...1. Recognizing dysfunctional measurement

And some technical questions:

- 1. What are the <data type or count> definitions**
- 2. How accurate does the data need to be?**
- 3. How accurate is it?**

2. Eliminating dysfunctional measurement...

Note: This is difficult

- 1. Establish, and monitor the move to and adherence to 'policies' for good, functional measurement**
 - that establish local ownership of data**
 - and recognize and value data privacy (i.e. *reduce* 'visibility' to enable informational measurement and acknowledge the 'information horizon')**
 - and/or support aggregation (to avoid focusing on the specific, and to avoid sub-optimization)**

...2. Eliminating dysfunctional measurement...

2. Support technical correctness

- **GQM**
- **'M' definition, and access to definitions by collectors and users**
- **Simple analyses (good graphics and minimal arithmetic, esp. division)**
- **V&V**

...2. Eliminating dysfunctional measurement

3. Periodically evaluate the information need and value delivered by measurements

- **At all levels**
- **Assess the degree of dysfunction (see previous questions)**
- **Attempt to stop collecting dysfunctional (and low value) data**

‘If Software Metrics leads to something like Taylorian “scientific management” in software development, I for one, will bow my head in shame.’

Gerald M. Weinberg

writing in the foreword of Toms Gilb’s ‘Software Metrics’, the first s/w metrics book

‘The good news is that you *can* succeed in producing a culture conducive to measurement. There are organizations in which people have given themselves completely to the pursuit of organizational goals... ... organizations in which members hunger for measurement as a tool that helps get the job done...
... To use measurement inappropriately would betray a sacred trust, and no one would consider such a betrayal.’

Robert D. Austin

**OXFORD
SOFTWARE ENGINEERING
LIMITED**

9 Spinners Court, 53 West End,
Witney,
Oxfordshire
OX28 1NH

www.osel.co.uk
shelley@osel.netkonect.co.uk
Tel. +44 (0) 1993 700878

Some real measurement policies

1. *All data will be unambiguously specified to meet the measurement needs. These specifications will be available and accessible to those collecting, analysing and using the measurement data.*
2. *All collected, analysed and stored data will be traceable back to measurement goals, which are, in turn, traceable to needs*
3. ***Data and any resultant measures will remain private to the individual, project or department concerned unless they are agreed to be helpful in improving the performance of <org> as a whole***
4. *Measures will be routinely reviewed to determine their benefit. Where the cost of collection outweighs the benefit collection will stop*