

The Darker Side of Metrics



CAST 2006

June 7, 2006

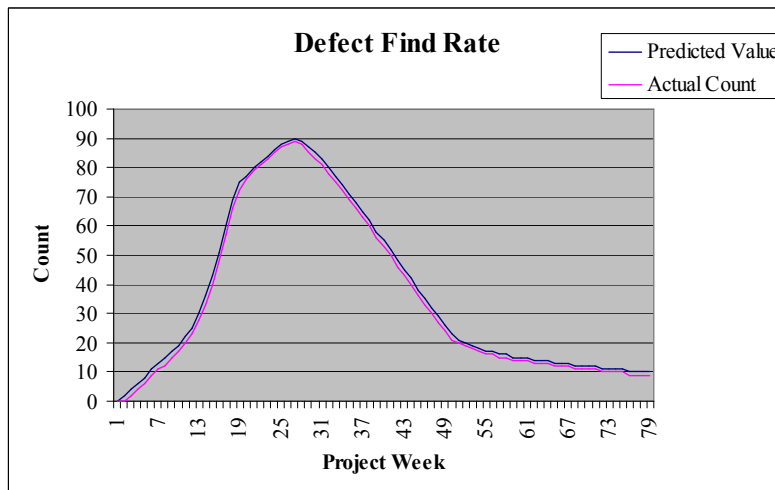
Douglas Hoffman
Software Quality Methods, LLC.
24646 Heather Heights Place
Saratoga, California 95070-9710
Phone 408-741-4830
Fax 408-867-4550
doug.hoffman@acm.org

Douglas Hoffman

Copyright © 2000-2006, SQM, LLC.

1

What Does This Chart Show?



Douglas Hoffman

Copyright © 2000-2006, SQM, LLC.

2

Kaner's Measurement Factors²

- 1) The **purpose** of the measure.
- 2) The **scope of the measurement**.
- 3) The **attribute to be measured**.
 - 4) The appropriate **scale for the attribute**.
 - 5) The natural **variability of the attribute**.
- 6) The **instrument that measures the attribute**.
 - 7) The **scale of the instrument**.
 - 8) The **variability of measurements** made with this instrument.
- 9) The **relationship** between the **attribute** and the **instrument**.
- 10) The natural, foreseeable [*probable*] **side effects** of using the instrument to measure this attribute.

²Kaner, C. "Measurement Issues and Software Testing," QUEST Conference, March 2001

Fundamental Issues

The **relationship** between the **attribute** and the **instrument**.

- What are we really trying to find out?
- Does the **measurement** really show us the **attribute**?
- Does the **measurement always vary** when the **attribute varies**? (And always *in the same direction*?)

The natural, foreseeable **side effects** of using the **instrument** to measure this **attribute**.

- What behavior changes are we likely to see to improve the measurements?
- Can measured values be changed without the attribute changing?

Readiness for Release

- Defect Find/Fix Rates
- Percent of Tests Run/Passed
- Complex Model Based Metrics (e.g., COCOMO II)

Defect Find/Fix Rate

- Mechanism
 - Counts of defects
 - Plots to show convergence
- Potential problems
 - Relationship to release readiness
 - Natural variation
 - Difficulties with counting

Defect Rate: Examples of Dysfunction

- “Consolidation”
- Unassigned
- Delays in reporting
- Shifting blame
- Reassignment

Douglas Hoffman

Copyright © 2000-2006, SQM, LLC.

7

Percent of Tests Run/Passed

- Mechanism
 - Counts of tests planned/run
 - Ratios to show completion
- Potential problems
 - Relationship with release readiness
 - Natural variation
 - Difficulties with counting

Douglas Hoffman

Copyright © 2000-2006, SQM, LLC.

8

Percent of Tests: Examples of Dysfunction

- Redefining what a test is
- Not counting tests that can't run
- Redefining "Pass"
- Updating expected results

Model Based Metrics

- Mechanism
 - Several measurements combined
 - Equation used to describe progress
- Potential problems
 - Relationship to project status
 - Natural variation
 - Difficulties with measures
 - "*Believing it is seeing it*" effect

Model Based Metrics: Examples of Dysfunction

- Release on faith - the model says so
- Punishment of the innocent
- Proliferation of questionable reports
- “Dry-Labing”

Defect Report Counts: Examples of Dysfunction

- Management changes the rules
 - No deferral
 - No assignment to other projects
 - No cloning of defects
- “Go to the movies” report reduction
- Questionable resolutions
- Un-assignment of defects

Observation Versus Control

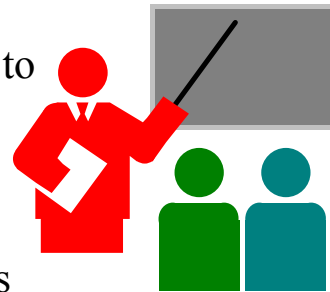
- Taking measures to learn about a product or process

or

- Taking measures so corrective action can be taken

What Do I Recommend?

- Understand your models
- Weigh the costs and benefits
- Select a range of metrics to provide balance
- Use metrics to observe
- Watch out for side effects



Partial Bibliography

- Hoffman, D, "The Darker Side of Metrics," 2000 PNSQC
- Kaner, C and Bond, W, "Software Engineering Metrics: What Do They Measure and How Do We Know?," *10th International Software Metrics Symposium (Metrics 2004)*
- Kaner, C. "Yes, But What Are We Measuring?," 1999 PNSQC
- Kaner, C. "Rethinking Software Metrics," *Software Testing and Quality Engineering* vol. 2, no. 2 (2000)
- Kaner, C. "Measurement of the Extent of Testing," 2000 PNSQC
- Lawrence, Brian "Measuring Up," *Software Testing and Quality Engineering* vol. 2, no. 2 (2000)
- Weinberg, G. "How Good Is Your Process Measurement," *Software Testing and Quality Engineering* vol. 2, no. 1 (2000)
- This information was first generated for presentation and discussion at the *Eighth Los Altos Workshop on Software Testing* in December, 1999. I thank the LAWST attendees, *Chris Agruss, James Bach, Jaya Carl, Rocky Grober, Payson Hall, Elisabeth Hendrickson, Bob Johnson, Mark Johnson, Cem Kaner, Brian Lawrence, Brian Marick, Hung Quoc Nguyen, Bret Pettichord, Melora Svoboda, and Scott Vernon*, for their participation and ideas.

Douglas Hoffman

Copyright © 2000-2006, SQM, LLC.

15

Example: A Race¹

- Sandy, Joe and Susan run in a race. Sandy comes in first, Joe second, and Susan third.
 - We assign Sandy the number 1 for first place and give her \$10,000
 - We assign Joe the number 2 and give him \$1,000
 - We assign Susan the number 3 and give her \$100

We assigned the numbers according to a rule.

- Questions
 - Is Sandy twice as fast as Joe and three times as fast as Susan?
 - Is Sandy 10 times as fast as Joe and 100 times as fast as Susan?
 - Isn't the assignment of the numbers based on their speed?

Did we measure their speed or not?

¹Kaner, C. "Yes, But What Are We Measuring?," 1999 PNSQC

Douglas Hoffman

Copyright © 2000-2006, SQM, LLC.

16