The Darker Side of Metrics

PNSQC ´00

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

Copyright © 2000, Software Quality Methods, LLC. No part of these graphic overhead slides may be reproduced, or used in any form by any electronic or mechanical duplication, or stored in a computer system, without written permission of the author.
Software Metrics

• Assignment of values according to rules

• Cem Kaner’s ten factors

• Observation versus Control

• Examples on the darker side
Example: A Race

- 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
Kaner’s Measurement Factors

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 variation* of the attribute.

6) The *instrument* that measures the attribute.

   7) The *scale* of the instrument.

   8) The *variation* of measurements made with this instrument.

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

10) The *probable side effects* of using this instrument to measure this attribute.
Observation Versus Control

- Taking measures to learn about a product or process

  or

- Taking measures so corrective action can be taken
Readiness for Release

• Defect find/fix rate

• Percent of tests running/passing

• Complex model based metrics
Defect find/fix rate

• Mechanism
  – Counts of defects
  – Plots to show convergence

• Potential problems
  – Relationship with release readiness
  – Natural variation
  – Difficulties with counting
Defect Rate Side effects

- “Consolidation”
- Unassigned
- Delays in reporting
- Shifting blame
- Reassignment
Percent of Tests

- **Mechanism**
  - Counts of tests planned/run
  - Ratios to show completion

- **Potential problems**
  - Relationship with release readiness
  - Natural variation
  - Difficulties with counting
Percent of Tests

- 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 is seeing it” effect
Model Based Metrics

• Release on faith - the model says so
• Punishment of the innocent
• Proliferation of questionable reports
• “Dry labeling”
Other Effects of Metrics

- 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
What Can We Do?

• Use metrics to observe
• Select metrics scientifically
• Understand the models
• Weigh the costs and benefits
• Watch out for side effects
Acknowledgments

• Kaner, C. “Yes, But What Are We Measuring?,” 1999 PNSQC
• Kaner, C. “Measurement of the Extent of Testing,” 2000 PNSQC
• 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.