Measuring Effectiveness of Test Design

Plenary Session @ QUEST 2009
Pradeep Chennavajhula, CEO, Edista Testing Institute

April 2009 | Chicago
The World of Software Development

How the customer explained it
How the Project Leader understood it
How the Analyst designed it
How the Programmer wrote it
How the Business Consultant described it

How the project was documented
What operations installed
How the customer was billed
How it was supported
What the customer really needed
Effectiveness...?
NFR catalog for “Test Effectiveness”

CATALOG for Rigor of Testing

CATALOG for Speed of Testing

CATALOG for ROI of Testing

CATALOG for Defects in Testing

Metrics in later slides
Exploding the “Rigor Of Testing”

It is important to define metrics for attributes of Test Rigor & study impact of influencers on them.
Exploding the “Speed Of Testing”

As delayed defect detection increases the cost of correction, speed of testing directly impacts ROI.
Exploding the “ROI of Testing” Sub-Goal

ROI of testing

Cost of Testing

Cost of Bug Fixing

Benefit of Testing

Cost of Human Resources

Cost of Test Tools

Cost of Test Infrastructure

Detect Reduction Target

Detect Test Automation

Detect Detection Potential

Quality Improvement Potential

Profit Potential

Satisfaction Improvement Potential

Testing contributes to immediate, long term, tangible and soft benefits
Exploding the “Defects in Testing”

Impacts of testing defects may range from mere overheads to major lapses
### Indicative Metrics for Test Effectiveness

**RIGOR**
- Test Case Density
- Test Case Granularity
- % Code Coverage
- Static Analysis Reports
- Structural Analysis Reports
- Risk Mitigated & Risk Balance
- Rounds of Testing
- Test Hit rate
- Defect Density Reduction Scores
- Degree of MTBF improvement

**SPEED**
- Testing Productivity
- Test Agility Quotient
  - Defect Hibernation duration
  - Time to first defect
  - DI-DD delay
  - DD-DR delay
  - (Age of Open Bugs)
  - Open-to-Closed bug ratio
- Time To Test Completion

**DEFECTS**
- Field defects Ratio
- Bug Slippage Ratio
- False Bug Ratio
- Poor Bug Report Ratio
  - “not –reproducible”
  - “need-more-info”
  - “duplicate”
- Reopened Bug Ratio

**COMMON**
- Test Tools Usage Index
- % Test Automation
- Management Tools Usage Index
- Collaboration Tools Usage Index
- SDLC model Fitment Index
- Test Team Competency Index

**ROI**
- Cost of Testing
- Overall Defect Reduction cost
- Differential Defect Reduction cost
- ROI of Test Automation
- Defect detection potential
- Defect reduction potential
- Differential defect reduction potential
- Saving potential due to test
- Profit potential due to test

An assorted combination of these metrics, with appropriate representation for each dimension, would do a good measurement of Test effectiveness
Test Hit Rate

Test Hit Rate: % failed test cases vs. total test cases
or
defects per unit test effort (or cost) per FP

- Hit Rate measures the combined contributions from various aspects of coverage
- To be measured against a reasonable sized group of related test cases
  or
- To be measured at appropriate points of effort/cost consumption
- Hit Rate tends to drop in later rounds of testing as the product matures
- Dropping hit rate indicates one of the two: “stop testing” or “innovate”
- Innovations could be ...
  - Test suite refactoring
  - Use of OA
  - Test Strategy change (Ex: brief exploratory test, risk based test, new tools)
- Monitoring & managing reasonable hit rate all through testing is important
- A genuine drop in hit rate after sufficient testing is an indication to “STOP TESTING”

While requirement & code coverage based testing yield a somewhat uniform hit rate over time, exploratory & risk based testing can yield bursts of high hit rates
Test Agility Quotient

- Test Agility Quotient is an Important measure of Speed of Testing
- It is a collection of a bunch of metrics as below
- Defect Hibernation Duration
  - Time to first defect (from the project start)
  - DI-DD delay (min:avg:max delay between defect introduction & defect detection)
  - DI-DR delay (min:avg:max delay between defect introduction & defect removal)
  - Age of Open defects (min:avg:max duration of bug fixing)
  - Open-to-Closed bug ration, at any point in time
- Rate of Test Effort Consumption in time
- Rate of Defect Detection in time
- Rate of Defect Detection vs.
- Rate of Test Effort consumption

Example: First 30% defects detected @ 5% of test effort spend point
Next 30% defects detected @ 25% of test effort spend point
Next 30% defects detected @ 65% of test effort spend point
Last 10% defects detected @ 100% of test effort spend point

- Speed of testing is less about how many test cases can be executed per unit time, but more about how many defects can be detected per unit test effort.

- Going one step further, it is about how soon the product can be debugged, stabilized & released to the market.

- Test agility quotient is a good measure of how closely the defect removal trend line chases the defect introduction trend line
Defect Reduction Cost & Potential

- **Defect Reduction Potential**: total defects removed / $ spent on testing

- **Differential Defect Reduction Potential**:
  - additional defects removed (beyond what is already removed) Vs.
  - additional $ spent on testing (beyond what is already spent)

- **(Estimated) Defect Density reduction potential**:
  - defects/FP@ start of testing – defects/FP@ end of testing vs. total $ spent on testing

- **Differential Defect reduction Cost**:
  - additional testing $ needed to detect & remove new defects from system under test

**Example**: $/FP, for First 30% defects detection & reduction
  - $/FP for Next 30% defects detection & reduction
  - $/FP for Next 30% defects detection & reduction
  - $/FP for Last 10% defects detection & reduction

Differential defect reduction cost increases exponentially. Test Suite optimization & test strategy innovation become important ROI drivers, particularly towards later stages of testing
Rigor or Speed or ROI – Which One?

All are important. However, relative importance of each might depend on some of the following:

- Nature of the product/project (mission critical or not)
- Non-Functional requirements of the product
- Expected user community & usage pattern for the product
- Certification requirements for the product
- Quality Promise made to the customers of the product
- SLAs signed for services provided using the product
- Penalty clauses against failure in meeting quality Promise

It is useful to set the “When to Stop Testing” criteria before strategizing testing.

“When to Stop Testing” criteria could tilt the balance towards any of the three sub-goals.

It can help decide the right balance.
Some “When To Stop Testing” criteria

Ex 1: Testing for a pre-defined duration: Something like ..
“Testing stops after 6 weeks irrespective of how many defects get detected”

Ex 2: Testing with a pre-fixed budget: Something like ..
“Testing stops when the testing budget of $XXX is exhausted”

Ex 3: Executing a pre-defined set of test cases: Something like ..
“Testing stops when the 2700 standard test cases are executed”.

Ex 4: Testing till average failure rate drops below a certain level: Something like .. “no matter what the cost/effort/time be, testing & debugging must continue till average failure rate drops below N defects per one execution hour”.

Examples above are all practical & useful strategies. They make the activity of testing that much more deterministic.
“Most dinosaurs were vegetarians and they never smoked tobacco or drank alcohol — and where are they now?!”
Thank you for Listening...

Pradeep Chennavajhula
pradeep.c@edistatesting.com | +91 98868 06317