

# The Darker Side of Metrics

**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

## Abstract

There sometimes is a decidedly dark side to software metrics that many of us have observed, but few have openly discussed. It is clear to me that we often get what we ask for with software metrics and we sometimes get side effects from the metrics that overshadow any value we might derive from the metrics information. Whether or not our models are correct, and regardless of how well or poorly we collect and compute software metrics, people's behaviors change in predictable ways to provide the answers management asks for when metrics are applied. Don't take me wrong; I believe most people in this field are hard working and well intentioned, and although some of the behaviors cause by metrics may seem funny, quaint, or even silly, they are serious responses created in organizations because of the use of metrics. Some of these responses seriously hamper productivity and can actually reduce quality.

The presentation focuses on a metric that I've seen used in many organizations (readiness for release) and some of the disruptive results in those organizations. I've focused on three examples of different metrics that have been used and a few examples of the behaviors elicited by using the metrics. For obvious reasons, the examples have been subtly altered to protect the innocent (or guilty). The three metrics are:

1. Defect find/fix rate
2. Percent of tests running/Percent of tests passing
3. Complex model based metrics (e.g., COCOMO)

Some of the observed behaviors include:

- testers withholding defect reports
- punishment of test groups (and individual testers) for not finding defects sooner
- use of "Pocket lists" of defects by developers and testers

- blocks of unrelated defects being marked as duplicates of one new consolidated defect (to reduce the defect count)
- artificial shifting of defects to other projects or temporary “unassigning” of defects to reduce the defect count
- changing definitions of what a test or test case is to change the count of tests
- shipping of products with known missing features because 100% testing was achieved
- routine changing of expected results to known incorrect results so the test would pass
- lowered ranking of testers because they weren’t finding defects as quickly as the model showed they should
- holding back on defect reporting and testing because the model showed they shouldn’t be found yet

## Biography

Douglas Hoffman is an independent consultant with Software Quality Methods, LLC. He has been in the software engineering and quality assurance fields for over 25 years and now is a management consultant in strategic and tactical planning for software quality. He is Section Chairman for the Santa Clara Valley Section of the American Society for Quality (ASQ) and is past Chairman of the Silicon Valley Software Quality Association (SSQA), a 750 member Task Group of the ASQ. He has been a speaker at dozens of software quality conferences including PNSQC and has been Program Chairman for several international conferences on software quality. He is also a member of the ACM and IEEE. Doug is certificated by ASQ in Software Quality Engineering and has been a registered ISO 9000 Lead Auditor. He has earned an MBA as well as a MS in Electrical Engineering and BA in Computer Science.

His experience includes extensive consulting, teaching, managing, and engineering in the computer systems and software industries. He has over fifteen years experience in creating and transforming software quality and development groups, and twenty years of management experience. His work in corporate, quality assurance, development, manufacturing, and support organizations makes him very well versed in technical and managerial issues in the computer industry.