

Warum Prüfen oft 50 mal länger dauert als Lesen

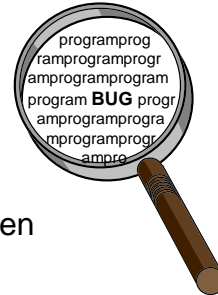
1

... und andere Überraschungen aus der Welt der Software-Reviews

Zusammenfassung:

In Schulungen zu Software-Reviews steht der Trainer immer dann vor einer didaktischen Herausforderung, wenn es darum geht,

- die von Fachleuten genannten Zahlenangaben zu begründen,
- die von den Teilnehmern intuitiv in ganz anderer Größenordnung eingeschätzt werden.



Software-Reviews
GI-TAV 23
18.11.2005

Peter Rösler
www.reviewtechnik.de

Zusammenfassung (2)

2

Zwei dieser Zahlenangaben, die den Teilnehmern üblicherweise besonders unplausibel vorkommen, sind:

- Die optimale Inspektionsrate für Textdokumente beträgt nur ca. **1 Seite pro Stunde** (und liegt damit um ca. den **Faktor 50** unter der reinen Lesegeschwindigkeit).
- Das durchschnittliche Review findet nur ca. **5%** der im Dokument vorhandenen Fehler.

Software-Reviews
GI-TAV 23
18.11.2005

Peter Rösler
www.reviewtechnik.de

Zusammenfassung (3)

3

- Im Beitrag wird gezeigt, mit welchen Kurzexperimenten und Schätzungen, die von den Teilnehmern selbst durchgeführt werden, obige Zahlenangaben zumindest in der Größenordnung als durchaus plausibel dargestellt werden können.
- An den Kurzexperimenten und Schätzungen beteiligten sich bisher ca. 90 Teilnehmer von 13 Reviewtechnik-Seminaren im Zeitraum von Oktober 2004 bis September 2005.

Software-Reviews
GI-TAV 23
18.11.2005

Peter Rösler
www.reviewtechnik.de

1. Einführung

4

- Die in diesem Beitrag betrachteten Software-Reviews werden oft auch als „Software-Inspektionen“, „Fagan/Gilb style inspections“ oder „Peer Reviews“ bezeichnet.
- Definitionen:
„major defect“ (im Gegensatz zu „minor defect“) :
Fehler, der möglicherweise erheblich höhere Kosten verursacht, wenn er später gefunden wird als jetzt.
1 Seite = 300 Wörter [3]

Software-Reviews
GI-TAV 23
18.11.2005

Peter Rösler
www.reviewtechnik.de

Capability Maturity Model (CMM)

5

Level	Focus	Key Process Areas
Level 5 Optimizing	Continuous improvement	Process Change Management Technology Change Management Defect Prevention
Level 4 Managed	Product and process quality	Software Quality Management Quantitative Process Management
Level 3 Defined	Engineering process	Organization Process Focus, Org. Process Definition Peer Reviews , Training Program Intergroup Coordination, SW Product Engineering Integrated Software Management
Level 2 Repeatable	Project management	Requirements Management, SW Project Planning SW Project Tracking and Oversight SW Subcontract Management, SW Quality Assurance SW Configuration Management
Level 1 Initial	Heroes	No KPAs at this time

Software-Reviews
GI-TAV 23
18.11.2005

Source: [www.software.org/
quagmire/descriptions/sw-cmm.asp](http://www.software.org/quagmire/descriptions/sw-cmm.asp)

Peter Rösler
www.reviewtechnik.de

Types of Review of Documents

6

Walkthrough: [[Presentation Review](#)] Activity: understanding

- author guides the group through a document and his or her thought processes, so all understand the same thing, consensus on changes to make

Review: [[Management-Review/Projektstatus-Review](#)] Activity: decision making

- group discusses document and makes a decision about the content, e.g. how something should be done, go or no-go decision

Inspection: Main activity: find defects

- formal individual and group checking, using sources and standards, according to detailed and specific rules

Software-Reviews
GI-TAV 23
18.11.2005

Source: Dorothy Graham

Peter Rösler
www.reviewtechnik.de

Rollen der Teilnehmer

7

- Moderator
- Autor
- Protokollführer
- Reviewer
- Vorleser/Reader (nur wenn „double checking“ gemacht wird)



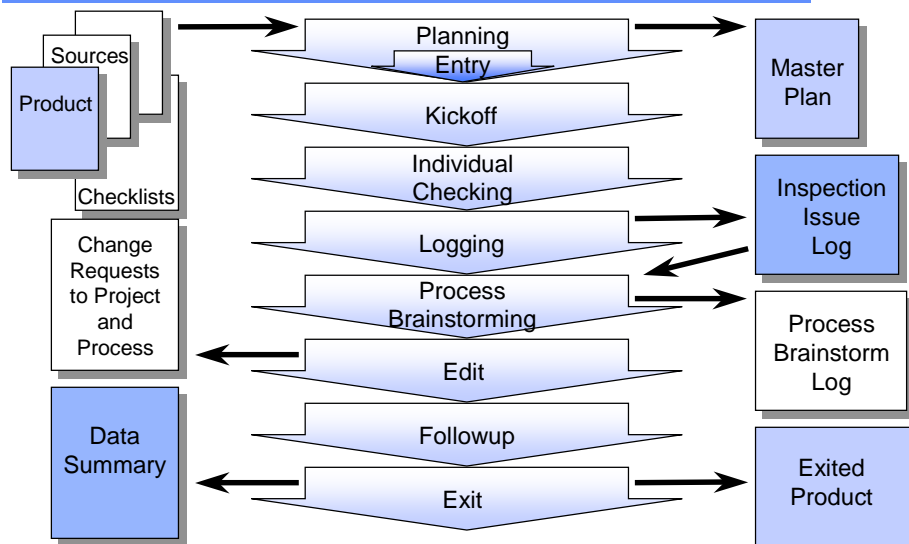
Ein Teilnehmer kann mehrere Rollen übernehmen.
Einzige Einschränkung: der Autor darf zusätzlich höchstens die Rolle eines Reviewers übernehmen.

Software-Reviews
GI-TAV 23
18.11.2005

Peter Rösler
www.reviewtechnik.de

Overall Process Map

8



Software-Reviews
GI-TAV 23
18.11.2005

Source: Tom Gilb,
Team Leader Course

Peter Rösler
www.reviewtechnik.de

Phase „Individual Checking“

9

- Von den Fehlern, die das Reviewteam insgesamt entdeckt, werden **80%** im „Individual Checking“ gefunden und **20%** im „Logging meeting“, sofern dort mit „Double Checking“ erneut Fehlersuche betrieben wird. [1]
- Da in den meisten durchgeführten Reviews auf „Double Checking“ verzichtet wird, werden in der Praxis eher **95%** der durch das Reviewteam entdeckten Fehler im „Individual Checking“ gefunden. [2]



Software-Reviews
GI-TAV 23
18.11.2005

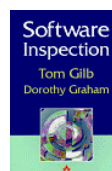
Peter Rösler
www.reviewtechnik.de

2. Lese- und Prüfgeschwindigkeiten im Vergleich 2.1 Die optimale Inspektionsrate

10

- In der Phase „Individual Checking“ wendet ein Reviewer alle ihm zur Verfügung stehenden sinnvollen Prüfstrategien an.
- Für Textdokumente dauert das erfahrungsgemäß **1 Seite pro Stunde**. ([1], Gilb/Graham 1993)
- Bandbreiten für die optimale Inspektionsrate:
 $1 \pm 0,8$ Seiten pro Stunde [3]

„Software Inspection“ von Tom Gilb
und Dorothy Graham, 1993



Software-Reviews
GI-TAV 23
18.11.2005

Peter Rösler
www.reviewtechnik.de

Kurzexperiment

11

- Mit folgenden Kurzexperimenten und Schätzungen kann der unplausibel klingende Wert „1 Seite pro Stunde“ nachvollzogen werden:
- In einem ersten Kurzexperiment prüfen die Teilnehmer einen typischen Spezifikationstext auf „minor defects“.

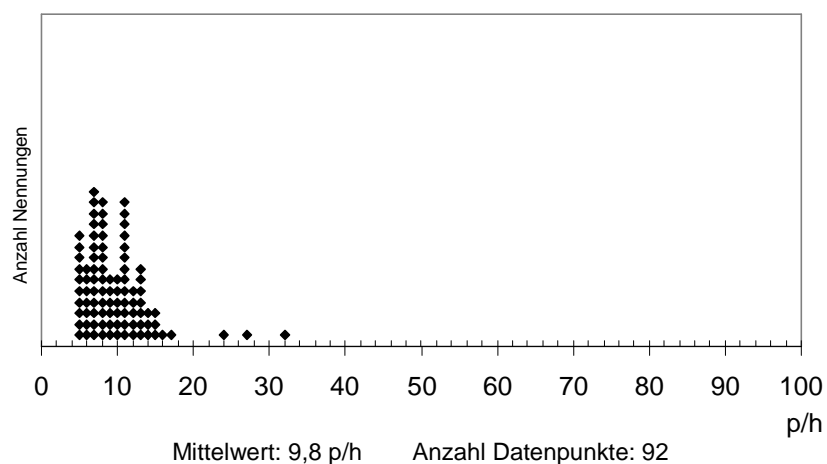
Spezifikationstext: „...For each flight in the database, this configuration file will be examined line by line. The assignment part of the first rule for which the condition part matches the carrier and trip number will be used as the new load control value. ...“).

Software-Reviews
GI-TAV 23
18.11.2005

Peter Rösler
www.reviewtechnik.de

„minor defects“-Inspektionsraten

12



Software-Reviews
GI-TAV 23
18.11.2005

Source: www.reviewtechnik.de

Peter Rösler
www.reviewtechnik.de

Schätzung für Major defects

13

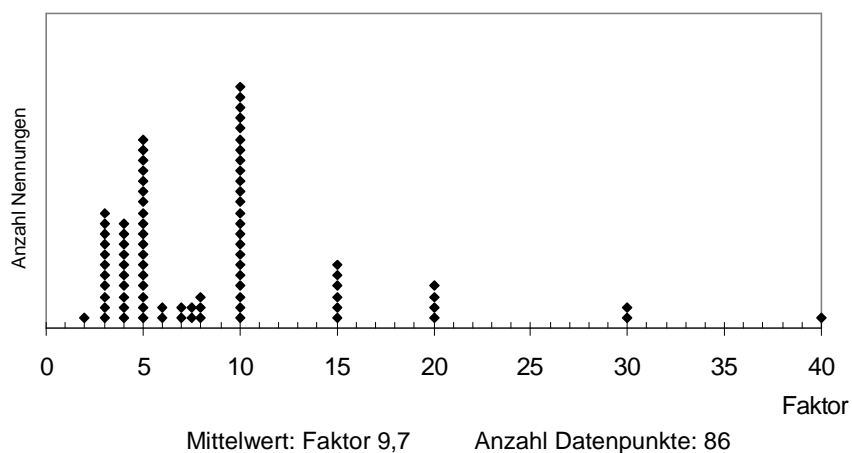
- Nach dem Kurzexperiment mit der „minor defects“-Inspektionsrate werden die Teilnehmer gefragt:
„Wie viel mal mehr Zeit (zusätzlich) würden Sie schätzungsweise benötigen, wenn Sie auch die inhaltlichen Fehler, also die **Major defects**, finden wollten?“

Software-Reviews
GI-TAV 23
18.11.2005

Peter Rösler
www.reviewtechnik.de

Geschätzte Zusatzzeit für „Major defects“-Suche

14



Software-Reviews
GI-TAV 23
18.11.2005

Source: www.reviewtechnik.de

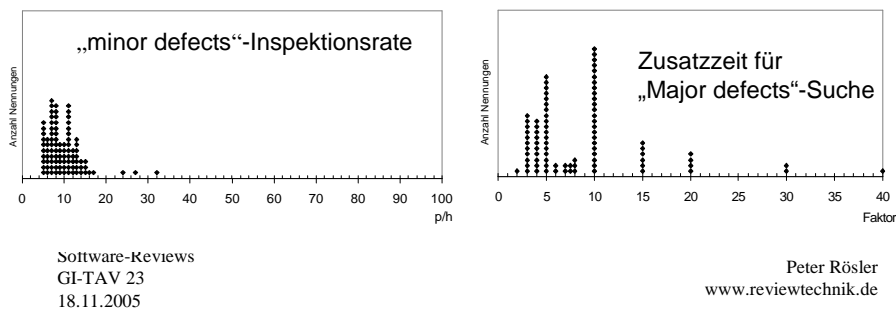
Peter Rösler
www.reviewtechnik.de

Die optimale Inspektionsrate

15

- Der Trainer kann jetzt die optimale Inspektionsrate vorrechnen:

$$\begin{aligned} \text{optimale Inspektionsrate} &= \\ \text{„minor defects“-Inspektionsrate} / (1 + \text{Zusatzzeitfaktor}) &= \\ 9,8 \text{ p/h} / (1 + 9,7) &= \mathbf{0,9 \text{ p/h}} \end{aligned}$$



2.2 Warum Prüfen oft 50 mal länger dauert als Lesen

16

- Das „Naturtempo“ beim Lesen von (deutschsprachigen) Texten beträgt ca. 240 Wörter pro Minute. [4]
- Das sind umgerechnet 48 Seiten pro Stunde.
- In einem Kurzexperiment ermitteln die Teilnehmer ihre eigene Lesegeschwindigkeit an einem Übungstext.

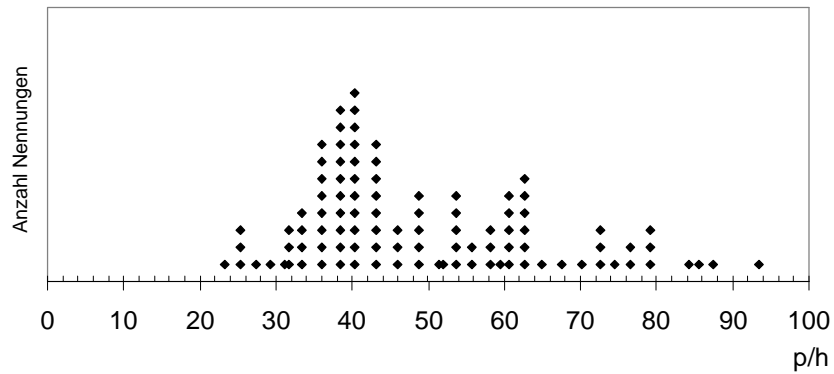
Textausschnitt: „...Ein skeptischer Autor beschloss, Magliabechis Bekanntheitsgrad wegen seines Schnellesens und Gedächtnisses auf die Probe zu stellen und gab ihm ein neues Manuskript, das er vorher nie gesehen haben konnte. ...“ [5])

Software-Reviews
GI-TAV 23
18.11.2005

Peter Rösler
www.reviewtechnik.de

Lesegeschwindigkeiten

17



Mittelwert: 49,4 p/h Anzahl Datenpunkte: 99

Software-Reviews
GI-TAV 23
18.11.2005

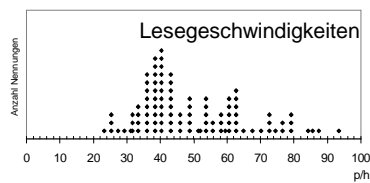
Source: www.reviewtechnik.de

Peter Rösler
www.reviewtechnik.de

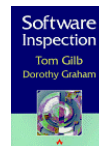
Prüfen dauert 50 mal länger als Lesen

18

- Lesegeschwindigkeit:
ca. 50 Seiten pro Stunde



- optimale Inspektionsrate:
ca. 1 Seite pro Stunde



„Man kann zwar ca. 50 Seiten pro Stunde,
aber nur ca. 1 Seite pro Stunde gründlich prüfen.“

Software-Reviews
GI-TAV 23
18.11.2005

Peter Rösler
www.reviewtechnik.de

3. Effektivität von Reviews

19

- Das Reviewteam findet ca. **50%** der im Dokument vorhandenen Fehler [1][2], wenn die optimale Inspektionsrate eingehalten wird, ansonsten:
- Tom Gilb / Kai Gilb: "We have found Inspection to operate in the area of about **3% (the normal for inspections not optimized with optimal checking rates etc.)** to about 88%." [6]
- Mit den im Folgenden gezeigten Kurzexperimenten und Schätzungen kommt man auf einen ähnlichen Wert: nur ca. **5%** der Fehler werden in typischen Reviews entdeckt.

Software-Reviews
GI-TAV 23
18.11.2005

Peter Rösler
www.reviewtechnik.de

3.1 Typische Inspektionsraten beim Review eines Textdokuments

20

- In einer Umfrage gaben die Teilnehmer an, mit welchen Inspektionsraten die Reviews in ihrer Firma typischerweise durchgeführt werden.

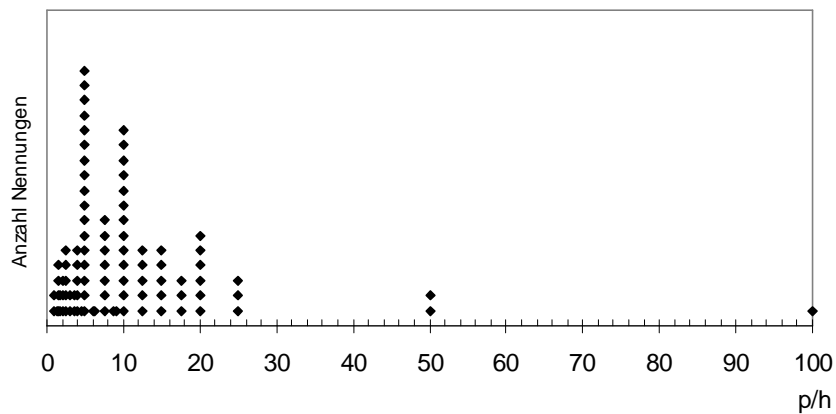
„In Ihrer Firma soll ein Review eines Textdokuments stattfinden. Das Dokument ist so groß, dass es in einzelnen Paketen geprüft wird, jedes Paket in einer eigenen Reviewsitzung. Angenommen, jeder Reviewer kann ca. 2 Stunden Aufwand in die Vorbereitung für eine Reviewsitzung investieren. Versuchen Sie abzuschätzen, aus wie vielen Seiten ein solches Paket in Ihrer Firma typischerweise besteht! ...“

Software-Reviews
GI-TAV 23
18.11.2005

Peter Rösler
www.reviewtechnik.de

Typische Inspektionsraten beim Review eines Textdokuments

21



Mittelwert: 10,2 p/h Anzahl Datenpunkte: 93

Software-Reviews
GI-TAV 23
18.11.2005

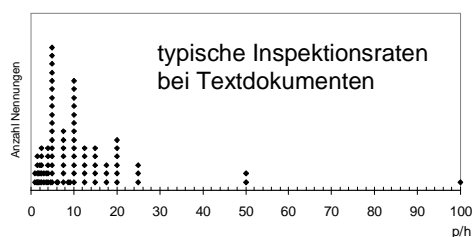
Source: www.reviewtechnik.de

Peter Rösler
www.reviewtechnik.de

Typische Reviews: um Faktor 10 zu schnell!

22

- Die optimale Inspektionsrate (1 Seite pro Stunde) wird in typischen Reviews um den Faktor 10 verfehlt.



- Die Werte aus der Umfrage entsprechen den publizierten Erfahrungen:

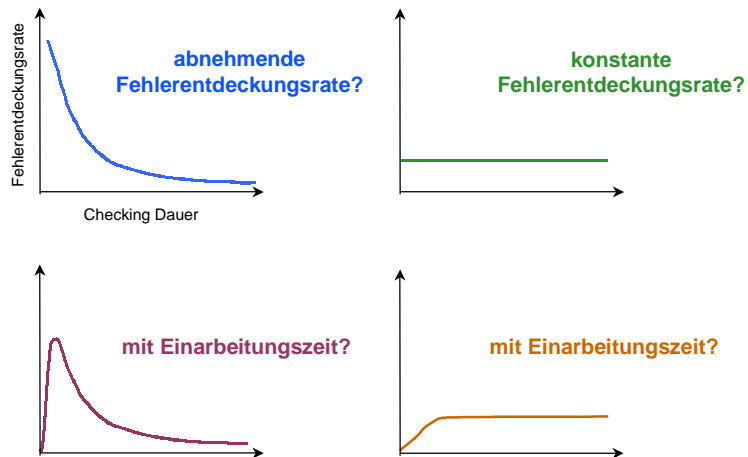
„we typically review documents at **five** to **twenty** pages per hour, when systematic calculation and application of optimum rates are not applied.“ [1]

Software-Reviews
GI-TAV 23
18.11.2005

Peter Rösler
www.reviewtechnik.de

Wann findet ein Reviewer die Fehler in der Phase „Individual Checking“?

23



Software-Reviews
GI-TAV 23
18.11.2005

Source: www.reviewtechnik.de

Peter Rösler
www.reviewtechnik.de

3.2 Hypothese: Im „Individual Checking“ werden die Fehler zeitlich ungefähr gleichverteilt gefunden

24

- 10 Reviewer wurden gebeten, bei jedem entdeckten Fehler den Zeitpunkt minutengenau zu protokollieren.

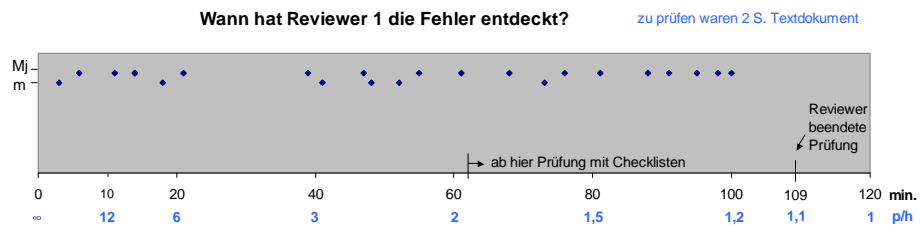


Software-Reviews
GI-TAV 23
18.11.2005

Peter Rösler
www.reviewtechnik.de

Zeitpunkt der Fehlerentdeckung in Phase „Individual Checking“, Reviewer 1

25



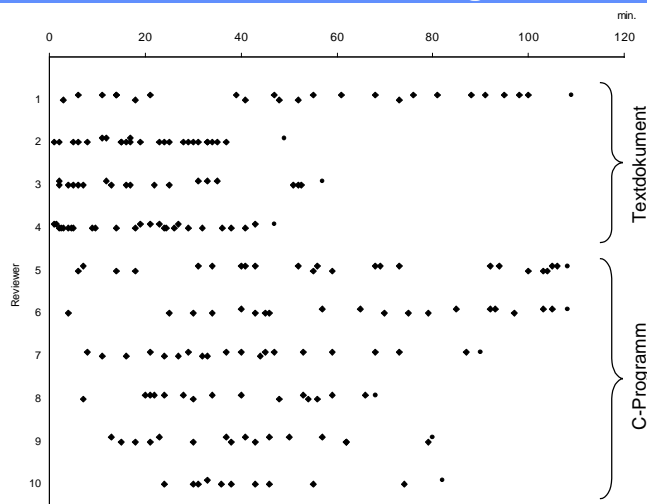
Software-Reviews
GI-TAV 23
18.11.2005

Source: www.reviewtechnik.de

Peter Rösler
www.reviewtechnik.de

Zeitpunkt der Fehlerentdeckung in Phase „Individual Checking“

26



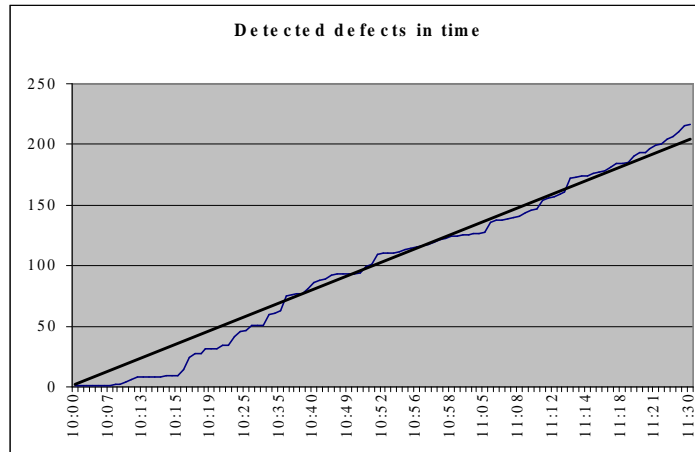
Source: www.reviewtechnik.de

Software-Reviews
GI-TAV 23
18.11.2005

Peter Rösler
www.reviewtechnik.de

Weiteres Indiz für konstante Fehlerentdeckungsrate

27



Source: Górski / Jarzębówicz: Development and validation of a HAZOP-based inspection of UML models, in: Proceedings of the 3rd World Congress for Software Quality, Erlangen, 2005

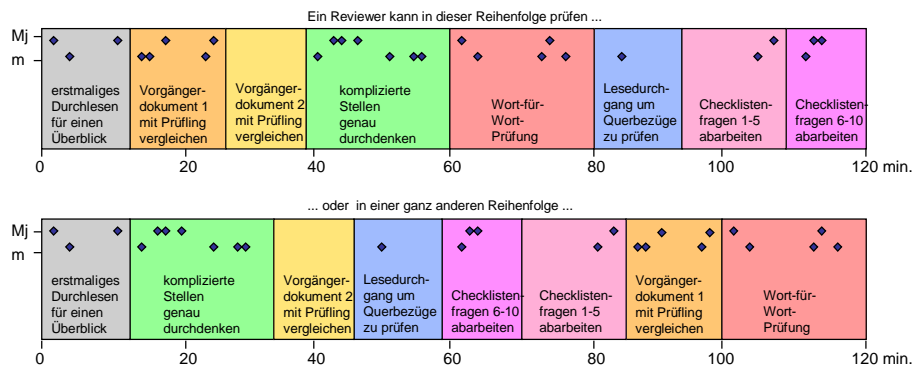
Software-Reviews
GI-TAV 23
18.11.2005

Peter Rösler
www.reviewtechnik.de

Noch ein Indiz: die Reihenfolge der Prüfstrategien ist weitgehend beliebig

28

Ein Reviewer kann vorab nicht wissen, welche Prüfstrategie zu welchem Erfolg führt. Es bleibt reiner Zufall, wann während der Prüfzeit welcher Fehler entdeckt wird.



Software-Reviews
GI-TAV 23
18.11.2005

Source: www.reviewtechnik.de

Peter Rösler
www.reviewtechnik.de

3.3 Schlussfolgerung: Im typischen Review werden nur 5% der vorhandenen Fehler entdeckt

29

- Aus der Hypothese der Gleichverteilung folgt:
„Das Verfehlen der optimalen Inspektionsrate um Faktor x bewirkt ein Sinken der Effektivität um Faktor x “ ($x \geq 1$).
- Da in den realen SW-Projekten ca. 10 Seiten pro Stunde geprüft werden und damit die optimale Inspektionsrate um ca. Faktor 10 verfehlt wird, werden diese Reviews anstelle der möglichen 50% **nur 5%** der vorhandenen Fehler entdecken!

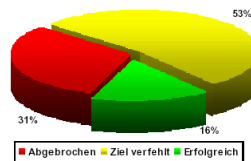
Software-Reviews
GI-TAV 23
18.11.2005

Peter Rösler
www.reviewtechnik.de

Schlussbemerkung

30

- Wenn obige Hypothese richtig ist, dann wirft das ein schlechtes Licht auf den Stand der Qualitätssicherung in den meisten SW-Projekten:
- Ein Projekt, das scheitern würde, weil in den Dokumenten zu viele Fehler stecken, scheitert wahrscheinlich auch dann noch, wenn die „Qualitätssicherung“ mit ihren Reviews 95% dieser Fehler übersieht!



Software-Reviews
GI-TAV 23
18.11.2005

Abbildung 1: Erfolgsstatistik (Quelle „The Standish Group“)

Peter Rösler
www.reviewtechnik.de

Literatur

31

1. Gilb, Tom / Graham, Dorothy: Software Inspection, Addison-Wesley, 1993,
2. Radice, Ronald A.: High Quality Low Cost Software Inspections, Paradoxicon Publishing,
3. www.gilb.com (Download Center), "Optimizing Inspection" von Tom Gilb
4. Michelmann, Rotraut / Michelmann Walter U.: Effizient lesen, Gabler Verlag, Wiesbaden, 1. Auflage 1995
5. Buzan, Tony: Speed Reading, mvg-verlag, Landsberg-München, 2002
6. www.gilb.com, Topic "Insp., Quality Control", Thread "Completeness of candidate doc.", Posting von Kai Gilb vom 17.09.2004
7. Gorski, J. / Jarzbowicz A.: Development and validation of a HAZOP-based inspection of UML models, in: Proc. of the 3rd World Congress for Software Quality, Erlangen, 2005



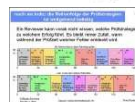
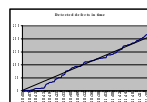
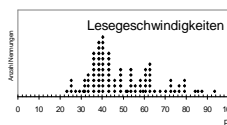
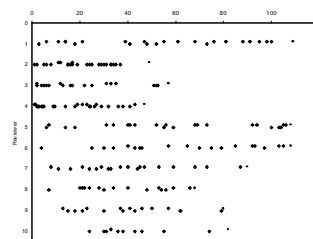
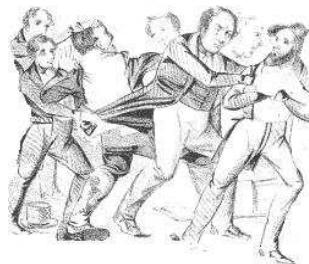
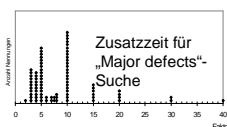
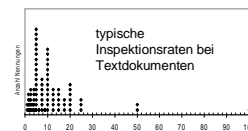
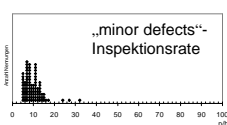
Software-Reviews
GI-TAV 23
18.11.2005

PowerPoint-Folien und Textfassung des Vortrags:
www.reviewtechnik.de/vortraege.html

Peter Rösler
www.reviewtechnik.de

Diskussion

32



Software-Reviews
GI-TAV 23
18.11.2005

Peter Rösler
www.reviewtechnik.de