



Prof. Dr. Andreas Spillner

Hochschule Bremen
 (E-Mail: andreas.spillner@hs-bremen.de)
 Andreas Spillner ist Hochschullehrer für Software Engineering mit Fokus auf Qualitätssicherung an der Hochschule Bremen. Er war Gründungsmitglied und ist nun Ehrenmitglied des German Testing Boards e. V. (GTB). Er war Gründungssprecher der Fachgruppe »Test, Analyse und Verifikation von Software (TAV)« der Gesellschaft für Informatik e.V. (GI).



Prof. Dr. Karin Vosseberg

Hochschule Bremerhaven
 (E-Mail: karin.vosseberg@hs-bremerhaven.de)
 Karin Vosseberg ist Hochschullehrerin für Informatik und Wirtschaftsinformatik an der Hochschule Bremerhaven mit den Schwerpunkten Systemintegration und Qualitätssicherung.



Prof. Dr. Mario Winter

Fachhochschule Köln
 (E-Mail: mario.winter@fh-koeln.de)
 Mario Winter ist Hochschullehrer an der Fachhochschule Köln für Software Engineering, Projektmanagement und Qualitätssicherung. Er ist Mitglied im GTB und war Sprecher der Fachgruppe TAV der GI e.V.



Peter Haberl

ANECON Software Design und Beratung G.m.b.H.
 (E-Mail: peter.haberl@anecon.com)
 Peter Haberl hat Wirtschaftsinformatik studiert und verantwortet bei der ANECON Software Design & Beratung G.m.b.H. als Geschäftsführer die Dienstleistungen und Kundenbeziehungen in Deutschland.

Wie wird in der Praxis getestet? Online-Umfrage in Deutschland, Schweiz und Österreich

In den letzten Jahren gab es im Bereich Qualitätssicherung und Testen viele neue Trends und Überlegungen: Test Driven Development, agiles Testen, exploratives Testen, modellbasiertes Testen, um nur einige zu nennen. Was davon hat Einzug in den Testalltag gehalten? Hat sich in den letzten Jahren viel verändert? Wie sieht der Testalltag überhaupt aus? Im Mai 2011 wurde eine anonyme Online-Umfrage zum Software-Test in der Praxis durchgeführt. In diesem Beitrag wird der Fokus der Auswertung auf die Durchführung von Qualitätssicherungsmaßnahmen gelegt: Wir betrachten, welche der Vorgehensweisen und Methoden in der Praxis angewendet werden.

»In den letzten Jahren hat das Testen von Software an Bedeutung zugenommen, es wird viel früher und systematischer getestet.« Stimmt diese Aussage? Oder ist es mehr ein Bauchgefühl oder ist hier eher der Wunsch der Vater des Gedankens? Wir waren uns nicht sicher, inwieweit diese Aussage von der derzeitigen Realität der Softwareentwicklung entfernt ist. Daher haben wir vom 1. bis zum 31. Mai 2011 eine anonyme Online-Umfrage durchgeführt, welche den aktuellen Stand des Software-Tests in der Praxis zum Gegenstand hatte. Die wissenschaftliche Leitung übernahmen die Fachhochschulen in Bremen, Bremerhaven und Köln. Als Förderer und Unterstützer der Umfrage konnten wir folgende Firmen und Organisationen gewinnen: ANECON Software Design und Beratung G.m.b.H., German Testing Board e.V. und Schweizer Testing Board, sowie ASQF e.V., Austrian Testing Board, dpunkt

Verlag GmbH, Softwareforen Leipzig GmbH und die Fachgruppe TAV der Gesellschaft für Informatik. Näheres zu den Unterstützern und Förderern sowie Links zu den einzelnen Institutionen sind zu finden unter [1].

Die Umfrage ist rollenspezifisch aufgebaut. Es gibt die folgenden drei Gruppen:

Projektleiter, QS-Beauftragte, Testmanager und Tester
 (Gruppe Tester)

Business Analyst, Entwickler, Mitarbeiter aus Betrieb&Support und andere Mitarbeiter
 (Gruppe Entwickler)

Executive und mittleres Management
 (Gruppe Management)

Mit der Umfrage haben wir Mitarbeiter aus allen Branchen mit einem leichten Schwerpunkt aus den Bereichen Automotiv, Telekommunikation und Banken erreicht. Es sind sowohl klein- und mittelständische

Unternehmen als auch Großunternehmen vertreten, wobei die Antworten aus den mittelständischen Unternehmen (101-1000 Mitarbeiter) ein wenig überwiegen. Der Bereich Forschung ist mit circa 10 % relativ gering vertreten. Insgesamt haben 1623 Personen an der Umfrage teilgenommen. Die Gruppe Tester ist mit 1008 Teilnehmenden vertreten, aber auch in der Gruppe Entwickler mit 394 und der Gruppe Management mit 221 Fragebogen gab es eine gute Resonanz. Details zur Anzahl der Teilnehmenden, zur Größe der Unternehmen, zur Zuordnung zu Branchen und der Art der Projekte sind auf der Internetseite der Umfrage [1] zu finden.

Einige Fragen wurden an alle drei Gruppen gestellt, andere wiederum waren auf die jeweiligen Rollen bezogen. Die Beantwortung der Fragen war optional und mussten nicht unbedingt beantwortet werden, Über alle Gruppen gleich verteilt

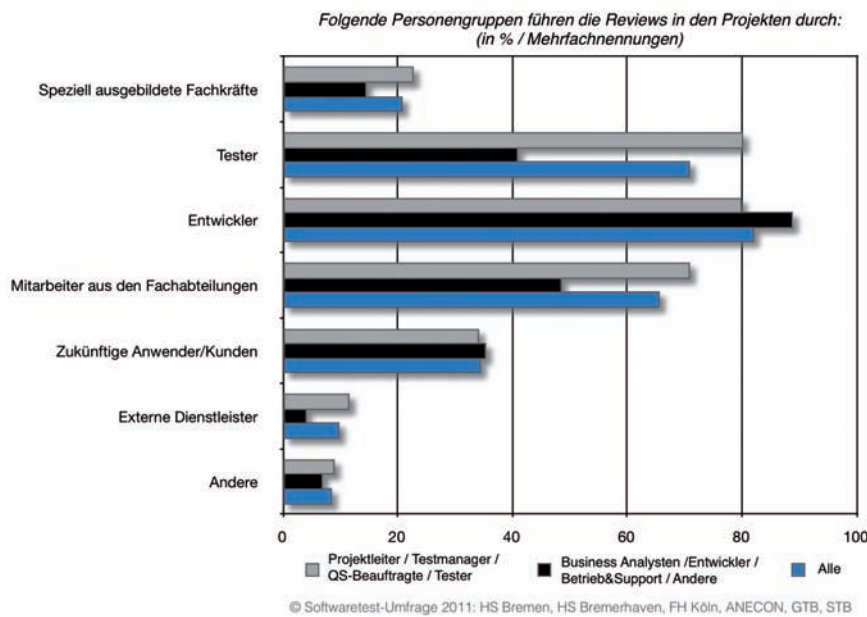


Abb. 1: Verantwortlichkeiten in Reviews

haben 50 % den sehr umfangreichen Fragebogen vollständig ausgefüllt. Damit haben wir im Wesentlichen unsere Zielgruppe, Praktiker mit unterschiedlichen Blickwinkeln auf die Qualitätssicherung, erreicht und es liegt eine gute Datenbasis für die Auswertungen vor. Um eine Verfälschung der Ergebnisse durch unterschiedliche Interpretation der Fachbegriffe zu verhindern, wurden bei fast jeder Frage die verwendeten Fachbegriffe mit einer Definition hinterlegt. Verwendet wurde das ISTQB®-Glossar [2].

Weil wir in diesem Beitrag auf die Durchführung von Qualitätssicherungsmaßnahmen blicken, betrachten wir nur die Auswertungen der beiden erstgenannten Gruppen. Das Management wurde nicht zur konkreten Durchführung von Qualitätssicherungsmaßnahmen befragt.

Werden Reviews eingesetzt?

Dynamisches Testen alleine genügt nicht, um die Qualität eines Softwaresystems zu sichern. Dokumente müssen geprüft und ggf. verbessert werden. Eine Frage lautete daher: Werden in ihren Projekten Reviews durchgeführt? Über 80 % der Gruppe Tester und 60 % der Gruppe Entwickler bejahen die Frage. Dies bedeutet aber auch, dass 20 % bzw. 40 % nicht die Notwendigkeit der Durchführung von Reviews sehen. Der vermeintliche Aufwand wird gescheut oder der Nutzen wird nicht gesehen.

Diejenigen, die Reviews einsetzen, wurden gefragt, wer diese durchführt.

Tester, Entwickler und die Mitarbeiter der Fachabteilungen sind an den Reviews beteiligt (siehe Abbildung 1). Deutlich wird, dass die Entwickler sich am meisten involviert sehen, die Tester sehen sich und die Entwickler eher gleich beteiligt. Speziell ausgebildetes Personal hat mit knapp über 20 % einen erfreulich hohen, aber noch ausbaufähigen Anteil.

Wird in Teststufen getestet?

Die aktuellen Diskussionen zum Testen weisen darauf hin, dass mit „kleinen“

Testobjekten begonnen wird, damit ein intensiver Test durchgeführt werden kann. Es wird daher angeregt, den Testprozess in Teststufen zu zerlegen. Wir fragten danach, ob die Tests auf unterschiedlichen Teststufen durchgeführt werden. Für knapp 90 % der Gruppe Tester ist dies der Fall, zwei Drittel der Gruppe der Entwickler sehen ebenfalls die Sinnhaftigkeit der Teststufen. Für die Tester ist es also klar, dass Teststufen sinnvoll und notwendig sind, die Entwickler sind eher etwas zurückhaltender.

Werden die Antworten detaillierter analysiert, wird der Unterschied noch deutlicher. Bei der Gruppe der Entwickler, die agil unterwegs sind, sehen 35 % keine Notwendigkeit für Teststufen. Wird ohne ein Vorgehensmodell (weder phasenorientiert noch agil) entwickelt, dann sehen 50 % keinen Sinn in Teststufen. Bei der Gruppe der agilen Tester kommen nur 12,5 % ohne Teststufen aus. Zwei Drittel der Tester testen in Teststufen, obwohl die Software nach keinem Vorgehensmodell entwickelt wird. Selbst ohne vorgeschriebenes Vorgehen strukturieren die Tester ihre Tätigkeit.

Bei der Betrachtung der einzelnen Teststufen gab es abgestufte Antwortmöglichkeiten. Wir haben Antworten wie „immer“ und „meist“ als zustimmende und „selten“ und „nie“ eher als ablehnende Haltung interpretiert, während „teils/teils“ eine neutrale Haltung beschreibt. Um die

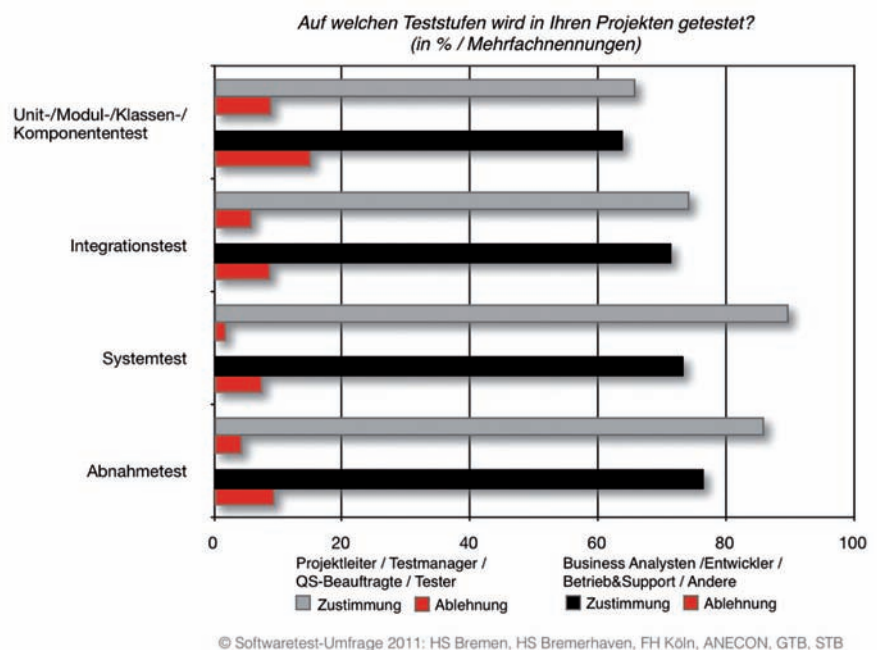


Abb. 2: Aufteilung in Teststufen

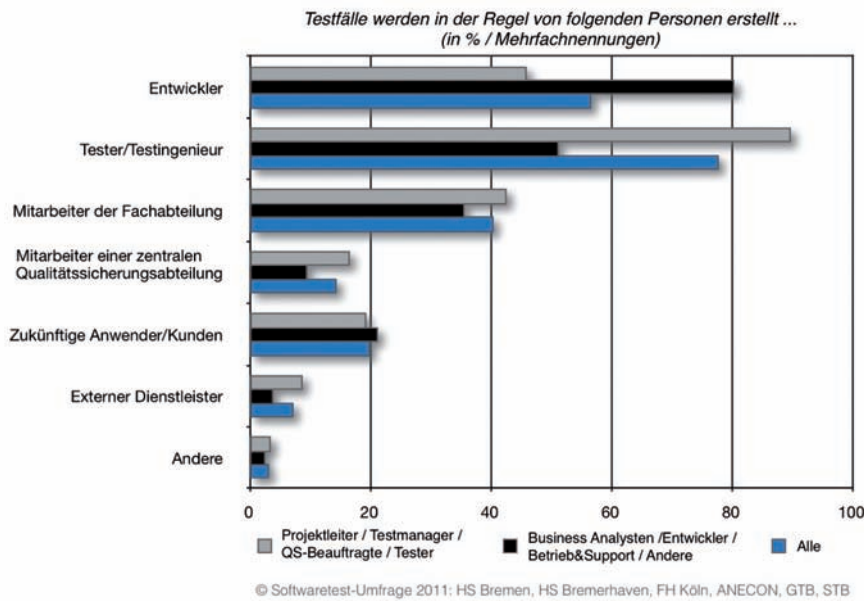


Abb. 3: Verantwortlichkeiten für die Erstellung der Testfälle

Polarisierung zu verdeutlichen, haben wir die Zustimmung und Ablehnung jeweils rollenbezogen dargestellt. Für diejenigen, die Teststufen einsetzen, ergibt sich damit ein ziemlich einheitliches Bild. Den verschiedenen Teststufen wird auf allen Ebenen eine relativ hohe Bedeutung zugeordnet, wobei besonders bei der Gruppe

der Tester der Systemtest und der Abnahmetest den höchsten Stellenwert erhalten. Was erstaunt, ist die ablehnende Haltung (knapp 20 %) gegenüber dem Komponententest bei der Gruppe Entwickler, wobei der Komponententest ja in der Praxis sehr oft in der Verantwortung der Entwickler liegt (siehe Abbildung 2).

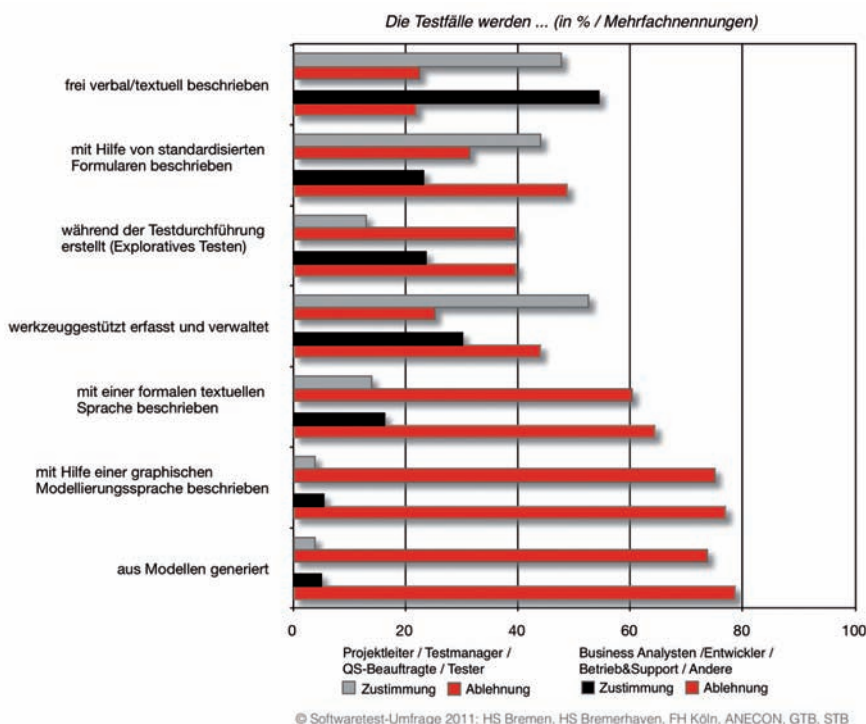


Abb. 4: Erstellung von Testfällen

Wie erfolgt die Testrealisierung und Testdurchführung?

Betrachten wir nun die Durchführung der Tests. Zunächst sind Testfälle zu erstellen. Wer erstellt die Testfälle?

Wer gehofft hat, dass Entwickler die Testfallerstellung den Testern überlassen, sieht sich widerlegt. 80 % der Entwickler schreiben auch Testfälle. Interessant ist, dass beide Gruppen unterschiedliche Sichtweisen haben. Während die Gruppe Tester sich selbst mit knapp 90 % in der Hauptverantwortung sieht, sehen die Entwickler ebenfalls sich selbst vorne (80 %, siehe Abbildung 3). Traut man sich hier gegenseitig nicht oder will man sich nicht in die Karten sehen lassen? Ein kooperatives Miteinander bei der Testfallerstellung scheint nicht gegeben.

Interessant ist, wie Testfälle erstellt und verwaltet werden. Auch hier macht die Zusammenfassung der positiven und negativen Antworten die Polarisierung deutlich (siehe Abbildung 4). Testfälle werden überwiegend frei verbal/textuell beschrieben, insbesondere in der Gruppe Entwickler. In der Gruppe Tester ist die Dokumentation der Testfälle durch standardisierte Formulare gleich auf. Formale Sprachen oder modellbasierte Erstellung von Testfällen haben in der Praxis keinen Stellenwert und werden nur wenig eingesetzt. Hier hat sogar die Gruppe der Entwickler leicht die Nase vorn. Ein Viertel der Entwickler erstellt die Testfälle während der Durchführung, von den Testern sind es knapp über 10 %. Die werkzeuggestützte Erfassung von Testfällen hat sich noch nicht umfassend etabliert. 50 % der Tester und 30 % der Entwickler verwenden eine Werkzeugunterstützung zur Erfassung der Testfälle. Die Entwickler haben erhebliche Ressentiments (45 %) gegenüber den Werkzeugen zur Erfassung und Verwaltung von Testfällen.

Wer testet?

Wie bei den Reviews teilen sich Entwickler, Tester und Mitarbeiter der Fachabteilung die Aufgabe „Testdurchführung“ auf den unterschiedlichen Teststufen. Hinzu kommen die zukünftigen Kunden und Anwender. Auch hier sehen sich die jeweiligen Gruppen (Entwickler, Tester) als Hauptverantwortliche. Externe Dienstleister spielen nur eine untergeordnete Rolle (siehe Abbildung 5).

Dieses verschiebt sich etwas bei der Frage nach den Verantwortlichkeiten für Tests zu

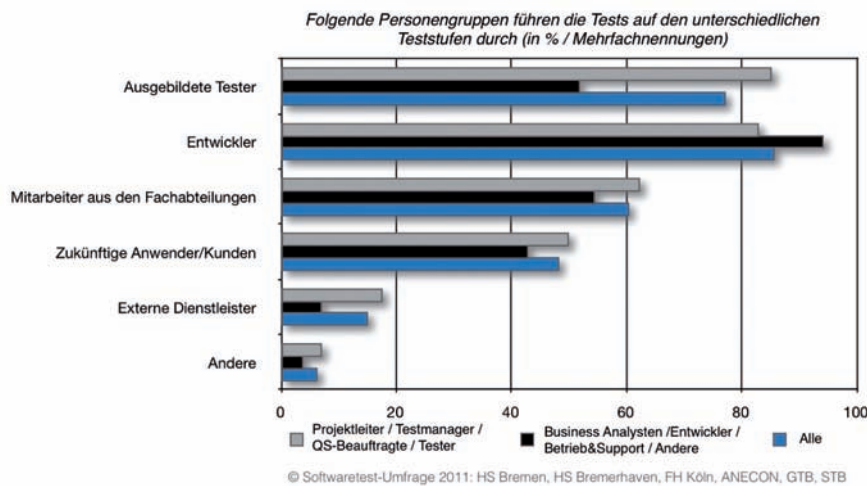


Abb. 5: Verantwortung der Tests in den Teststufen

verschiedenen Testarten, wie beispielsweise Last- und Perfomanztests, Migrationstest oder Sicherheits- und Penetrationstests. Hier sind ausgebildete Tester mit 90 % gefragt und auch Entwickler und Mitarbeiter aus den Fachabteilungen werden mit 70 % bzw. 50 % eingebunden. Circa 25 % der Befragten beauftragen externe Dienstleister mit der Durchführung der Tests für die speziellen Testarten.

Welche Testmethoden werden eingesetzt?
Wie systematisch wird getestet? Welche Testentwurfsverfahren zur Herleitungen der Testfälle werden eingesetzt? Die Gruppe der Tester haben wir zusätzlich nach Details zu ihrem Vorgehen befragt. Spezifikationsorientierte Verfahren (knapp 90 %), gefolgt von den erfahrungsbasierten Verfahren (knapp 80 %) und die strukturorientierten Verfahren (knapp 50 %) haben die weitaus größte Zustimmung. Mit 30 % sei noch die werkzeuggestützte dynamische Fehleranalyse erwähnt. Andere Verfahren spielen keine große Rolle.

Es wird deutlich, dass ein systematisches Vorgehen beim dynamischen Testen einen hohen Stellenwert hat. Nur wenige (unter 5 %) setzen keine konkreten Verfahren ein. Bei näherer Betrachtung lassen sich drei Gruppen von Black-Box-Testentwurfsverfahren ausmachen (siehe Abbildung 6). Der Schwerpunkt (60 bis 80 %) liegt auf Funktionsüberdeckung, anwendungsfallbasierten Tests sowie der Grenzwertanalyse und der Äquivalenzklassenbildung, gefolgt (30 bis 40 %) vom Zufallstest, zustandsbasiertem Test und der Verwendung von Entscheidungstabellen. Mit über 10 % werden Klassifikationsbaumverfahren, paar-

weises Testen und andere spezifikationsorientierte Verfahren in der Praxis angewendet.

Strukturorientierte Testverfahren (White-Box) scheinen für die Gruppe der Tester eher vernachlässigbar. Knapp 40 % wenden keine strukturorientierten Verfahren an. Anweisungs- und Entscheidungs-/Zweigüberdeckung werden von circa 35 bis 40 % eingesetzt, der einfache Bedingungstest und Pfadtest von etwas mehr als 20 %.

In den Antworten zu einer weiteren Frage zum erfahrungsbasierten Testen wird deutlich, dass nach wie vor sehr stark auf freies Testen ohne konkrete Vorgaben oder Ziele gesetzt wird. Die intuitive Testdurchführung (über 80 %) und das explorative Testen ohne konkrete Zielvorgabe (65 %) werden in hohem Maße eingesetzt.

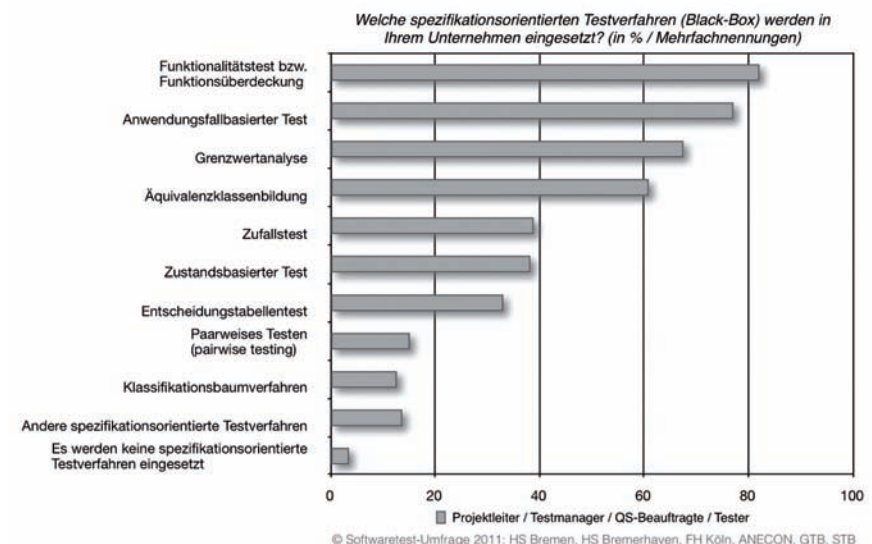


Abb. 6: Spezifikationsorientierte Verfahren

Abbildung 4 zeigt jedoch, dass die Ergebnisse aus dem explorativen Testen nicht für eine explizite Erstellung von Testfällen genutzt werden.

Wie weit ist der Stand der Testautomatisierung?

Das Thema Testautomatisierung beschäftigt viele Unternehmen. Insbesondere durch die Diskussion über agile Vorgehensmodelle in der Softwareentwicklung hat die Testautomatisierung einen neuen Stellenwert erhalten. Mit der Umfrage wollten wir herausfinden, wie der aktuelle Stand der Testautomatisierung ist. Wir fragten nach dem Grad der Testautomatisierung auf den unterschiedlichen Teststufen. 50 % der Gruppe Tester geben an, dass sie 70 bis 100 % ihrer Unittests automatisiert haben, bei den Entwicklern sind es 60 %. Jeder dritte Entwickler automatisiert 100 % der Unittests, bei den Tester ist es nur jeder fünfte.

Auf höheren Teststufen nimmt der Grad der Automatisierung deutlich ab. Beim Integrations- und beim Systemtest geben circa 25 % aus der Gruppe der Tester den Grad ihrer Automatisierung mit 70 % oder mehr an. In der Gruppe der Entwickler liegt dieser Grad für den Integrationstest mit 35 % etwas höher. Dies ist mit dem etwas höheren Anteil von „agilen“ Entwicklern zu erklären. Für sie ist der automatisierte Test der kontinuierlichen Integration selbstverständlich und gehört zu ihrem Entwicklungsmodell. Für den Abnahmetest nimmt der Grad der Automatisierung deutlich ab. Nur 10 %

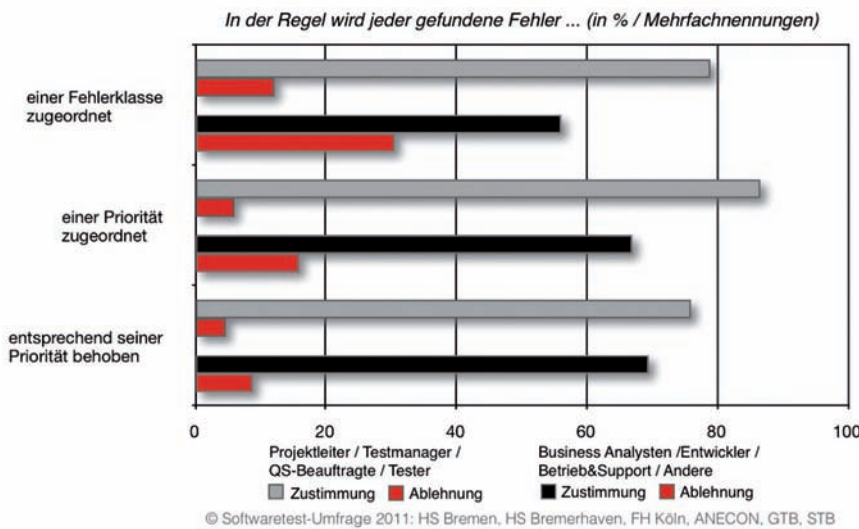


Abb. 7: Einordnung von Fehlern

geben in beiden Gruppen an, Testfälle des Abnahmetests zu automatisieren. Sogar knapp 40 % geben an, keine Automatisierung von Abnahmetests zu haben.

Wenn man den Grad der Testautomatisierung bei den unterschiedlichen Entwicklungsmodellen differenziert betrachtet, wird deutlich, dass in diesem Bereich ein großes Entwicklungspotential liegt. Selbst bei agilen Entwicklungsmodellen, die in der Diskussion eng verbunden sind mit Konzepten der testgetriebenen Softwareentwicklung, geben nur 50 % der Entwickler und sogar nur 30 % der Tester den Grad ihrer Testautomatisierung mit 100 % an.

Wie werden die Testergebnisse dokumentiert?

Das Ergebnis der Umfrage zeigt, dass die gefundenen Fehler systematisch erfasst werden und dies von der überwiegenden Mehrheit durch Werkzeuge wie einer Fehlerdatenbank unterstützt wird (über 80 % bei der Gruppe Tester, über 60 % bei

der Gruppe Entwickler). Formlos dokumentieren knapp 20 % der Entwickler und unter 10 % der Tester. Bei der Verwaltung der Fehler sind die Werkzeuge etabliert, bei der Verwaltung der Testfälle (s.o.) noch nicht.

In beiden Gruppen wird gesehen, dass die Einordnung und Priorisierung von Fehlern notwendig ist, wobei Tester diese Aufgabe viel stärker in ihrem Fokus haben. Wenn es eine Priorisierung gibt, wird diese in der Regel auch bei der Bearbeitung der Fehler berücksichtigt (siehe Abbildung 7).

Mit welchem Erfolg wird getestet?

Die Gruppe der Tester haben wir befragt: Wie schätzen Sie die Effektivität Ihrer Testfälle ein? Ein Großteil der Fehler wird bereits im Test gefunden, so schätzen fast 60 %, und 20 % schätzen die Effektivität der Testfälle immer noch als mittel ein, d. h. einige Fehler werden im Test gefunden.

Wir wollten es noch genauer wissen. Wir fragten zusätzlich ab, ob Fehler nach der Auslieferung der Softwaresysteme beim

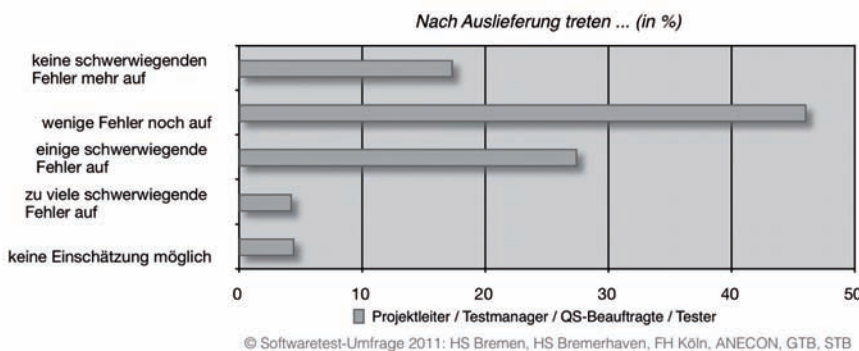


Abb. 8: Effektivität der Tests nach der Auslieferung

Kunden auftreten. Auch hier wird mehrheitlich darauf verwiesen, dass nur wenige Fehler beim Kunden auftreten. Trotzdem rutschen noch einige bis viele der schwerwiegenden Fehler durch. Mehr als insgesamt 30 % geben dies an (siehe Abbildung 8). Dies Ergebnis korrespondiert mit dem oben beschriebenen Ergebnis der Effektivität der Testfälle.

Resümee

Hat sich was in der Durchführung der Qualitätssicherungsmaßnahmen getan? Hat sich unser Bauchgefühl bestätigt? Die hohe Resonanz auf die Umfrage zeigt, dass das Thema in der Praxis angekommen ist, diskutiert und ernst genommen wird. Bereits in der Studie »State of the Practice« der Prüf- und Testprozesse in der Softwareentwicklung von 1997 haben Uwe Müller u. a. [3,4] festgestellt, dass viele der diskutierten Verfahren Einzug in die Praxis gehalten haben. Dies ist mit der aktuellen Umfrage 2011 bestätigt und der Trend konnte verstärkt nachgewiesen werden.

Tester und Entwickler teilen sich die Aufgaben beim Testen in nahezu gleichen Teilen. Die Tester sind für diese Aufgaben gut ausgebildet. So kennen circa 90 % der Gruppe Tester das ISTQB®-Ausbildungsschema Certified Tester, 70 % davon haben das Foundation Level-Zertifikat (d. h. etwa zwei Drittel aus dieser Gruppe sind zertifiziert). Bei den Entwicklern kennen knapp 40 % das Schema und davon haben über die Hälfte das Foundation Level-Zertifikat (d. h. etwa 20 % aller Entwickler sind zertifiziert). Es ist sehr erfreulich, dass auch Entwickler an einer fundierten Ausbildung im Testen interessiert sind.

Unabhängige, speziell ausgebildete Tester bzw. auch im Testen geschulte Entwickler sind aus unserer Sicht ein großes Potenzial zur Systematisierung der Prüf- und Testprozesse und somit zur Qualitätssteigerung insgesamt. In vielen Punkten ist das Vorgehen der Tester bereits jetzt strukturiert und systematisch. Es bleibt aber noch Handlungsbedarf. Die Einschätzung, dass noch einige oder zu viele schwerwiegende Fehler nach der Auslieferung der Software auftreten, macht deutlich, dass an der Effektivität der Tests gearbeitet werden muss. Aktuelle Themen wie modellbasiertes Testen sind noch nicht in der Praxis angekommen.

Wir möchten uns an dieser Stelle bei allen Teilnehmern der Umfrage recht herzlich bedanken. Wir hoffen, dass sie nun die

Möglichkeit haben, ihr Vorgehen beim Testen in Relation zum aktuellen Vorgehen in der Praxis zu setzen und daraus ihre Schlüsse für Verbesserungen zu ziehen.

Hinweis

Auf der Internetseite der Umfrage [1] werden wir alle Ergebnisse präsentieren, sobald die entsprechende Auswertung vorliegt. Hier sind neben den graphischen Darstellungen der Antworten auch Hinweise auf weiteren Veröffentlichungen, die andere Schwerpunkte der Umfrage auswerten.

Referenzen

- [1] Internet-Seite der Umfrage <http://www.softwaretest-umfrage.de/>
- [2] ISTQB®-Glossar als Download des German Testing Board
<http://www.german-testing-board.info/de/downloads.shtm>
- [3] Uwe Müller: Prüf- und Testprozesse in der Softwareentwicklung, Shaker Verlag, 1999
- [4] Uwe Müller, Thomas Wiegmann, Oral Avci: »State of the Practice« der Prüf- und Testprozesse in der Softwareentwicklung, Ergebnisse einer empirischen Untersuchung bei deutschen Softwareunternehmen, W. Mellis, G. Herzwurm, D. Stelzer (Hrsg): Studien zur Systementwicklung, Band 16, 1998 (steht zum Download zur Verfügung unter <http://systementwicklung-archiv.bibliothek.informatik.uni-koeln.de/28/>)