

No quality mobile apps without testing in production (TiP)

Marc van 't Veer

Content

- Test environments
- Relevance of (Big)up-front testing
- Difference between upfront and TiP
- My definition of TiP
- Strategy and methodologies
- Examples
- Problems with introducing TiP
- Conclusion

Representative test environments?

Number of users

Age of the device

Providers

OS and version

Browsers

Manufacturer

And the other ...
factors

Number of apps on
device

Devices models

Resolution

Network type

Location

Language

As test environment is like The Zoo



The wild ... is like production environment



Relevance of (Big)up-front testing?

- Testing on a test environment can give you a false positive result
- The app may look OK but is really not working

False positive
or negative



Why is (Big)up-front less important?

- Assumption

A tester *can* execute test cases that will simulate how the mobile app operates in an environment

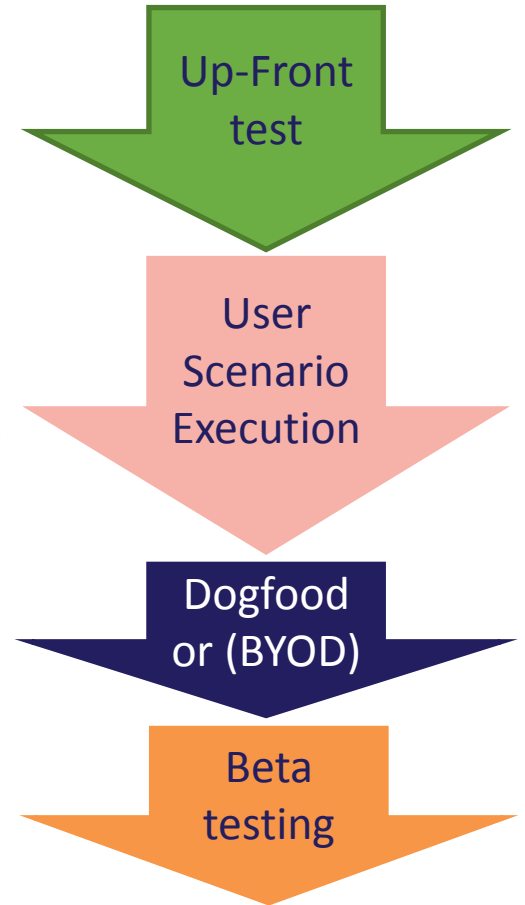
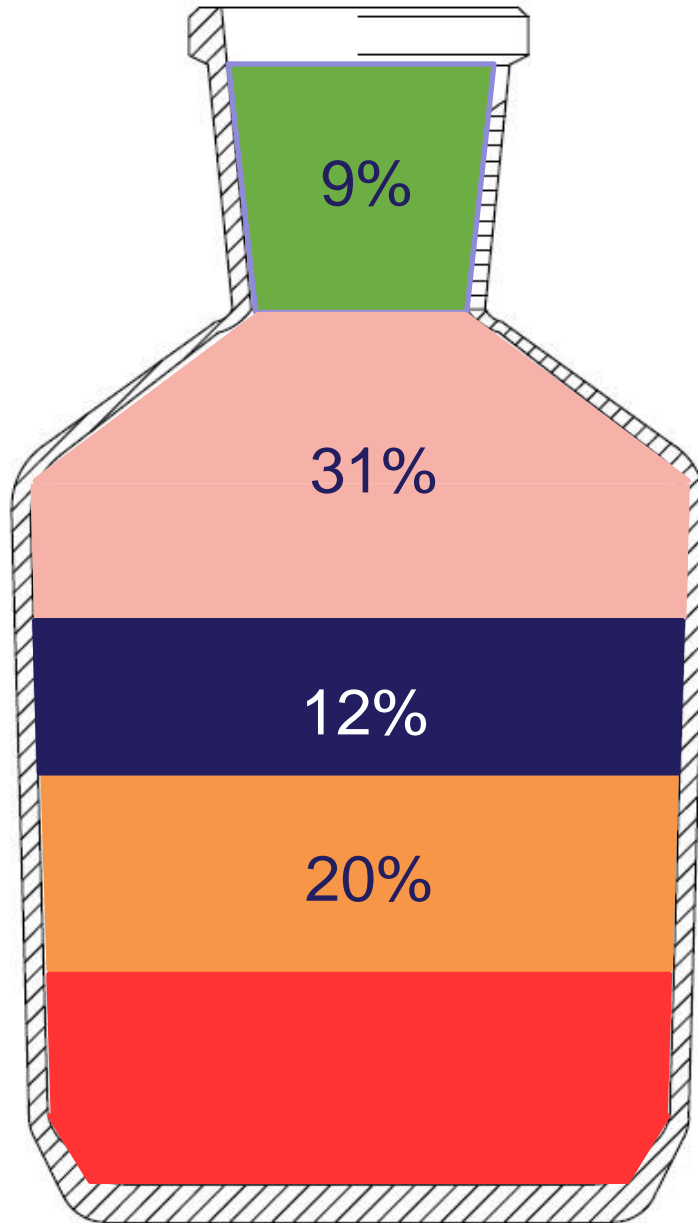
- There are too many variables to test

- The pace of changes is high

- The test environment is not representative

- The environment is *determining* the quality of the app

100%
91%
60%
48%
Residual risk



[Google GTAC 2010]

Difference between upfront and TiP

- Testing on the test environment is validation of the expected behavior
 - **Does an application work as specified?**
- Testing on the production environment is validation of the unexpected behavior
 - **Does an application do other things as well?**

So no quality apps without testing in production

My definition of TiP

“Testing in Production is a combined group of test activities (a test level) that uses the diversity of the production environment and real customer data to test the behavior of the developed service in the live environment and so minimizing the risks for the end customer”

Strategy to introduce TiP for Mobile apps

- Search for management support and awareness
 - Add TiP to the overall test strategy
 - Start with TiP methods without end user impact
 - Analyze current problems in production
 - Create pilot with test and operations (TestOps)
 - Collect learning's and improvement for next step
 - Create a direct feedback loop to current test team
-
- End result: New balance between upfront and TiP

Overview of TiP methodologies

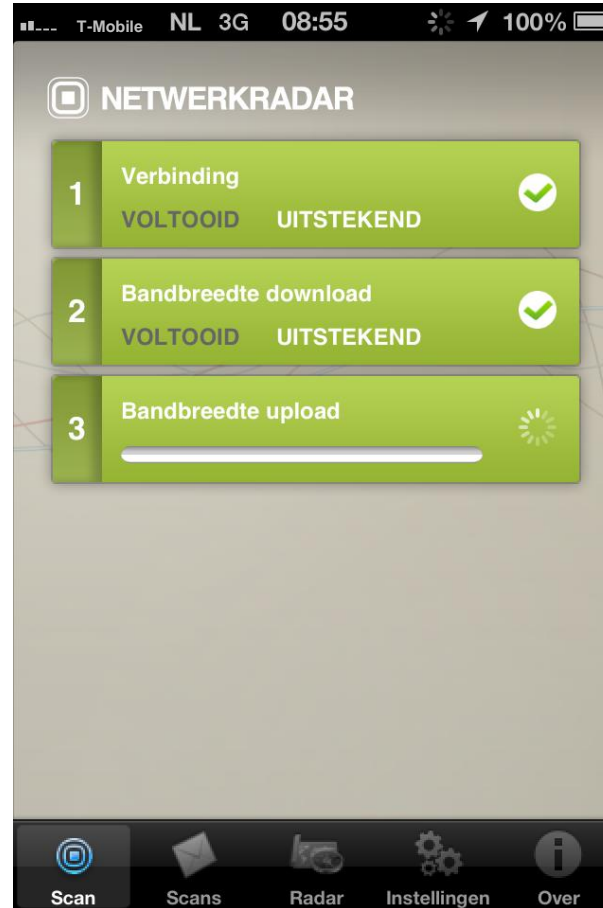
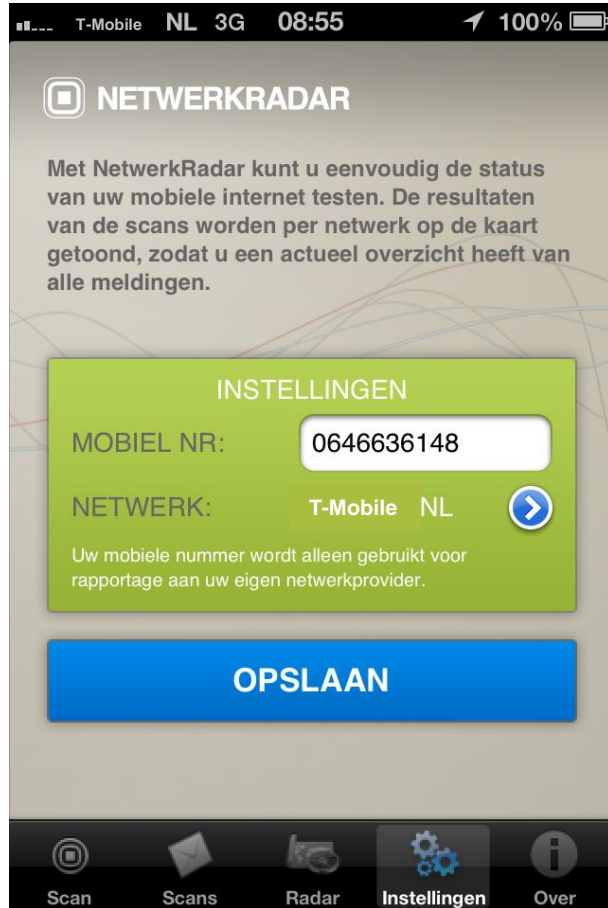
Methodology	Inputs are...	Effect is...	We Observe...
Data Mining	Real User Data	None	User and system behavior
User Performance Testing	Real User Data	None	System Behavior
Environment Validation	Real System Data	None	System Behavior
Experimentation for Design	Real User Data	Experiment with Users	User Behavior
Controlled Test Flight	Real User Data	Experiment with Users	System Behavior
Dogfood/beta	Real User Data	Experiment with Users	User and system behavior
Synthetic Tests in Production	Synthetic User Data	Acting on System	System Behavior
User Scenario Execution	Synthetic User Data	Acting on System	System Behavior
Load Testing in Production	Synthetic User Data	Stress System	System Behavior
Destructive Testing	Synthetic System Data	Stress System	System Behavior

Data mining - Coverage measurement

- Customer feedback on actual coverage
- If bad coverage is reported then improvement is planned

The screenshot displays the 'T-Mobile netwerk informatie' page. On the left, a navigation menu includes 'Samen Meer Bereiken', 'Dekkingskaart', 'Geplande werkzaamheden', 'Datatransparantie', and 'Datsnelheid verhogen'. The main content area features a search bar for 'Postcode of zoekwoord:' and radio buttons for 'Sprak (2G)', 'Mobiel internet (3G/HSDPA)', and 'Mobiel Internet (3G/HSDPA), verwacht binnen 3 maanden'. A search button labeled 'Zoek' is present. Below the search bar is a map of the Netherlands with a red location marker at '2612 RL Delft'. A tooltip above the marker asks 'Ervaart u een probleem op deze locatie? (inloggen)' and provides a tip: 'Tip: Versleep de marker met de muis naar de exacte locatie (Voorbeeld: plaats op treintraject)'. The map shows three coverage levels: 'Goede dekking smartphone' (dark pink), 'Goede dekking binnenshuis' (medium pink), and 'Goede dekking buitenshuis' (light pink). On the right, the 'T-Mobile Beheer' section lists benefits: 'Altijd grip op uw kosten', 'Facturen analyseren.', and 'Bedrijfsgegevens wijzigen.', along with a 'Inloggen op T-Mobile Beheer' button and links for 'Wachtwoord vergeten?' and 'Wat is T-Mobile Beheer?'. A 'Legenda' section explains the map's color coding, and a 'Veelgestelde vragen' section is also visible.

Data mining - Coverage measurement app



Environment validation - app in it life cycle

- What is the most dangerous period for an app in its life cycle?
 - First minute after installation or update
 - Fine tune analytic reports to look at this period
- Which gesture/action performed on pages, device OS, app version, orientation (portrait/landscape), battery status, temperature, memory use, disk space, geo-location, network type, time stamps, other apps running, crashes and exceptions
- But:
 - Server logs and mobile analytics need to be aligned
 - Analysis is hard work to understand what happens if installation or usage goes wrong

General preparation – lower risk for end users

- Combine tester, developers and operations
- Create continuous integration process
- Automate building and deployment process
- Have rollback/backup scenario (also tested)
- Have thorough version control
- Make test data recognizable (tags)
- Introduce real time monitoring (server / API / app)

Risks with introducing TiP

Department	Risks
Operations	More uncertainty (don't muddle on production)
Testers	New role, bigger impact of test errors, less room to play
Opinion	Fixes in a later stadium are more expensive and take more time
Organizations	Test/Dev/Ops are different companies
Development method	<ul style="list-style-type: none">• To (fully) test in production also the development process needs to change (Continue Integration)• Architectural changes, more work
End-users	End-users don't report incidents like testers (tools/process improvement needed)
Customers	Impact on normal business (e.g. outages)
Marketing	Possible bad PR by end-user
Business	<ul style="list-style-type: none">• Mixing production and test data• Unclear who is paying to get defects fixed• Less costs for testing en environments (replacement)

Conclusion: Quality apps with TiP

- Actual validation of the app quality with TiP methods
- Most important

~~False positive
or negative~~





Questions?

Marc van 't Veer
Test consultant

Polteq Amersfoort (Netherlands)
+31 (0) 6 46 63 61 48 (mob)
<http://www.polteq.com>
marc.vantveer@polteq.com



Polteq supports its customers:

- with executing and managing test projects
- with implementation and optimization of test processes
- with practical training and certification courses

Amersfoort



Amsterdam



Dordrecht



Groningen



Leuven