How to Get What You Really Want from Testing

Michael Bolton
DevelopSense
http://www.developsense.com
@michaelbolton
michael@developsense.com

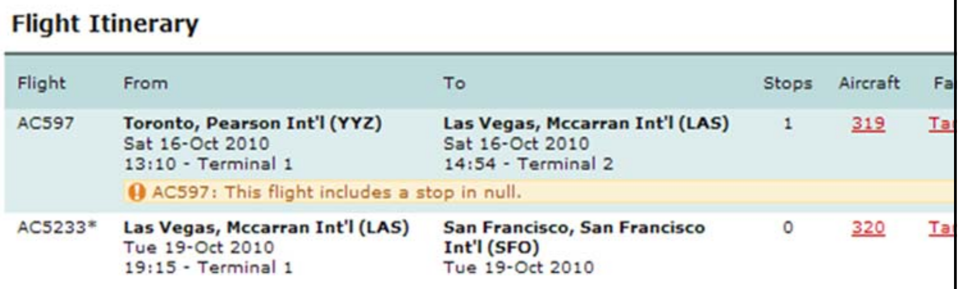
James Bach
Satisfice
http://www.satisfice.com
@jamesmarcusbach
james@satisfice.com

Updates

- This presentation is ALWAYS under construction.
- I probably will skip over some slides.
- Slides for this presentation are available on request.
- All material comes with lifetime free technical support.

THANK YOU!

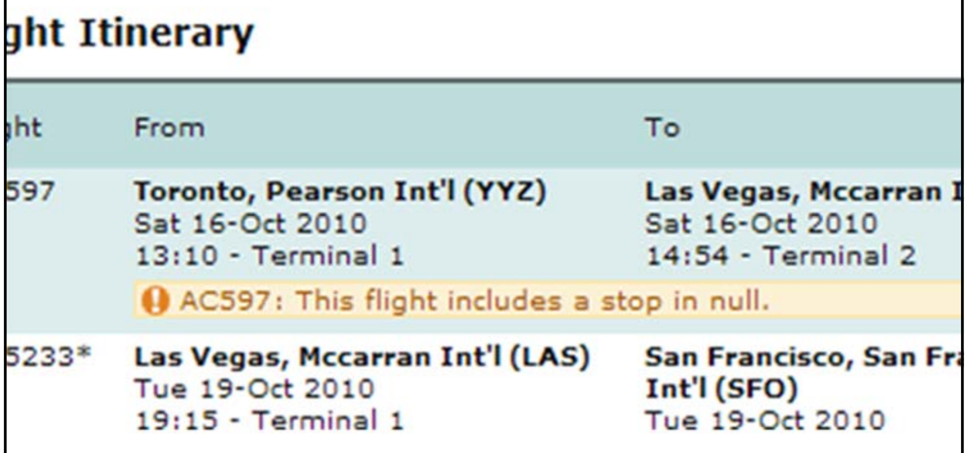
To all of the sponsors
and especially
to the EuroSTAR crew!



Flight Itinerary

Flight	From	To	Stops	Aircraft	Fa
AC597	Toronto, Pearson Int'l (YYZ) Sat 16-Oct 2010 13:10 - Terminal 1	Las Vegas, Mccarran Int'l (LAS) Sat 16-Oct 2010 14:54 - Terminal 2	1	319	Ta
	AC597: This flight includes a stop in null.				
AC5233*	Las Vegas, Mccarran Int'l (LAS) Tue 19-Oct 2010 19:15 - Terminal 1	San Francisco, San Francisco Int'l (SFO) Tue 19-Oct 2010	0	320	Ta

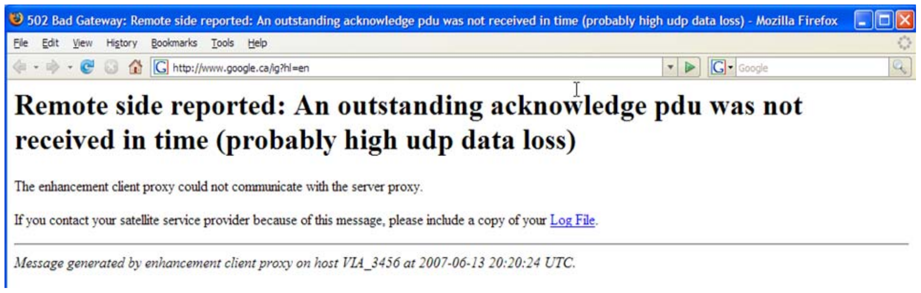
aircanada.com



ght Itinerary

ght	From	To
597	Toronto, Pearson Int'l (YYZ) Sat 16-Oct 2010 13:10 - Terminal 1	Las Vegas, Mccarran I Sat 16-Oct 2010 14:54 - Terminal 2
	AC597: This flight includes a stop in null.	
5233*	Las Vegas, Mccarran Int'l (LAS) Tue 19-Oct 2010 19:15 - Terminal 1	San Francisco, San Fr Int'l (SFO) Tue 19-Oct 2010

aircanada.com


502 Bad Gateway: Remote side reported: An outstanding acknowledge pdu was not received in time (probably high udp data loss) - Mozilla Firefox
File Edit View History Bookmarks Tools Help
http://www.google.ca/ig?hl=en
Google

Remote side reported: An outstanding acknowledge pdu was not received in time (probably high udp data loss)

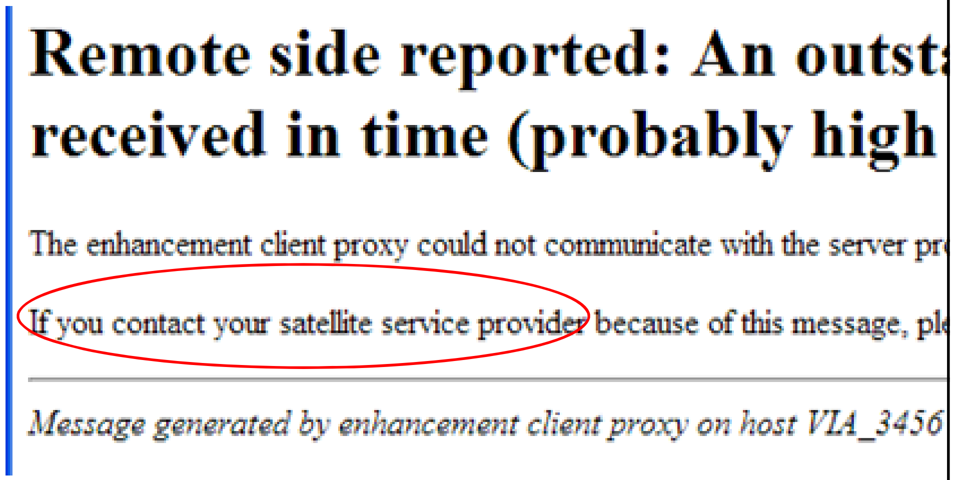
The enhancement client proxy could not communicate with the server proxy.

If you contact your satellite service provider because of this message, please include a copy of your Log File.

Message generated by enhancement client proxy on host VIA_3456 at 2007-06-13 20:20:24 UTC.

Error Message
Via Rail, between Montreal and Toronto, 2007

But I can’t contact my… oh, never mind.



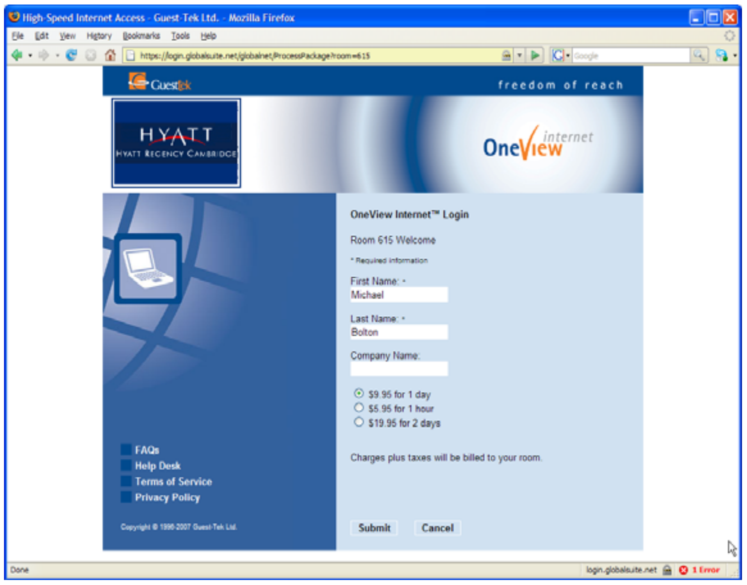
Remote side reported: An outst
received in time (probably high

The enhancement client proxy could not communicate with the server pr

If you contact your satellite service provider because of this message, pl

Message generated by enhancement client proxy on host VIA_3456

Error Message
Via Rail, between Montreal and Toronto, 2007


High-Speed Internet Access - Guest-Tek Ltd. - Mozilla Firefox
File Edit View History Bookmarks Tools Help
https://login.globalsuite.net/globalnet/ProcessPackage?room=615
GuestTek
freedom of reach
HYATT
HYATT REGENCY CAMBRIDGE
OneView internet
FAQs
Help Desk
Terms of Service
Privacy Policy
Copyright © 1996-2007 Guest-Tek Ltd.
Submit
Cancel
Done
login.globalsuite.net
1 Error

OneView Internet™ Login

Room 615 Welcome

* Required information

First Name: *
Michael

Last Name: *
Bolton

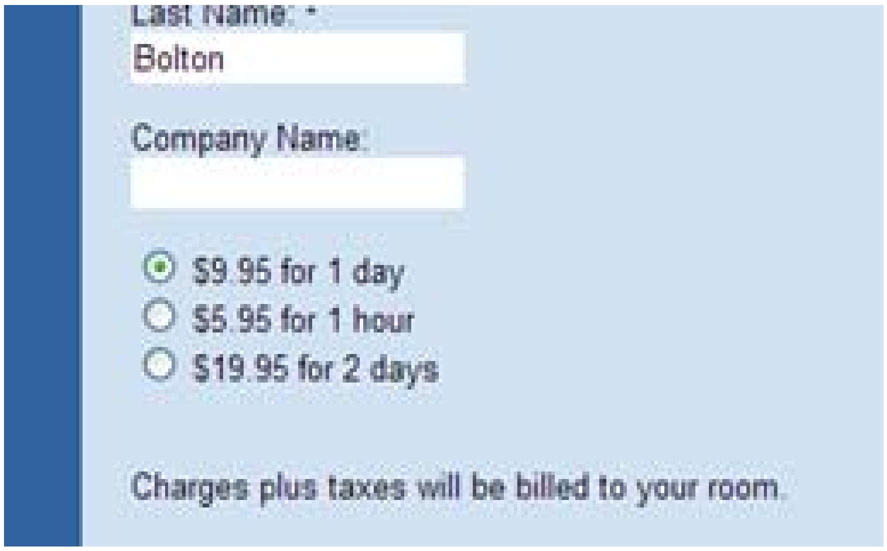
Company Name:

[✓ $9.95 for 1 day]
$5.95 for 1 hour
$19.95 for 2 days

Charges plus taxes will be billed to your room.

Hyatt Regency, 2008

If you can't do math, it's a nickel extra.



Last Name: *
Bolton

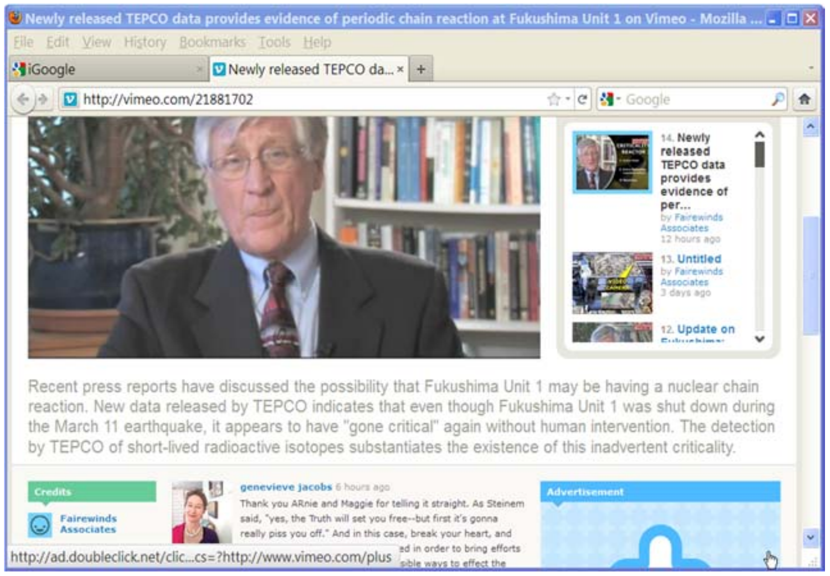
Company Name:

[✓ $9.95 for 1 day]
$5.95 for 1 hour
$19.95 for 2 days

Charges plus taxes will be billed to your room.

Hyatt Regency, 2008

Why you shouldn't let an unsupervised algorithm choose your sponsored links (1).

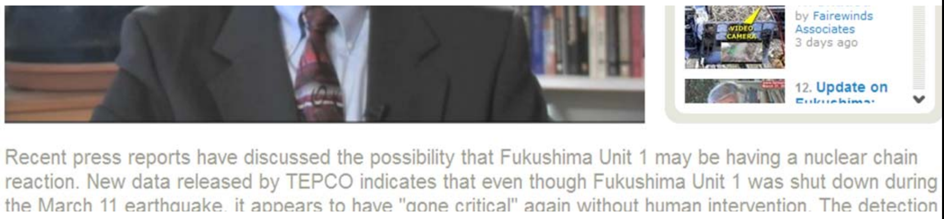

Newly released TEPCO data provides evidence of periodic chain reaction at Fukushima Unit 1 on Vimeo - Mozilla ...
File Edit View History Bookmarks Tools Help
iGoogle
Newly released TEPCO da...
http://vimeo.com/21881702
Google
14. Newly released TEPCO data provides evidence of per...
by Fairewinds Associates
12 hours ago
13. Untitled
by Fairewinds Associates
3 days ago
12. Update on
Credits
Fairewinds Associates
Advertisement
http://ad.doubleclick.net/clic...cs=?http://www.vimeo.com/plus

Recent press reports have discussed the possibility that Fukushima Unit 1 may be having a nuclear chain reaction. New data released by TEPCO indicates that even though Fukushima Unit 1 was shut down during the March 11 earthquake, it appears to have "gone critical" again without human intervention. The detection by TEPCO of short-lived radioactive isotopes substantiates the existence of this inadvertent criticality.

genevieve jacobs 6 hours ago
Thank you ARnie and Maggie for telling it straight. As Steinem said, "yes, the Truth will set you free--but first it's gonna really piss you off." And in this case, break your heart, and
ed in order to bring efforts
sible ways to effect the

Vimeo's Web Page
Spring 2010

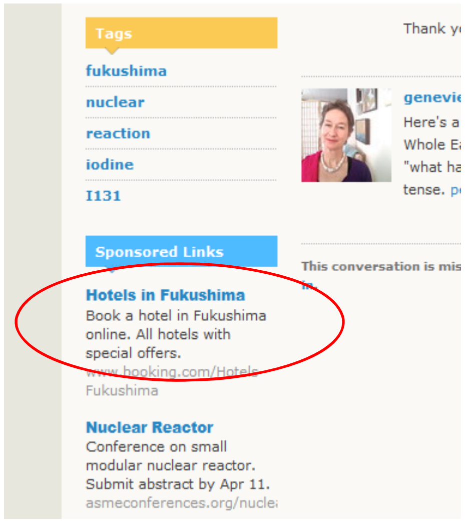
Why you shouldn't let an unsupervised algorithm choose your sponsored links (2).



Recent press reports have discussed the possibility that Fukushima Unit 1 may be having a nuclear chain reaction. New data released by TEPCO indicates that even though Fukushima Unit 1 was shut down during the March 11 earthquake, it appears to have "gone critical" again without human intervention. The detection

Vimeo's Web Page
Spring 2010

Why you shouldn't let an unsupervised algorithm choose your sponsored links (3).


Tags

fukushima

nuclear

reaction

iodine

I131

Sponsored Links

Hotels in Fukushima
Book a hotel in Fukushima online. All hotels with special offers.
www.booking.com/Hotels-Fukushima

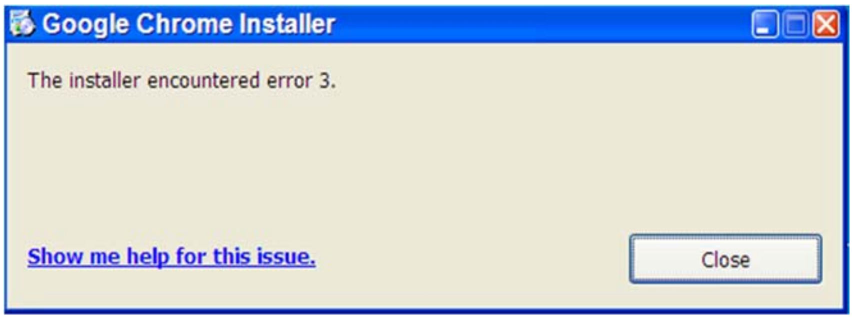
Nuclear Reactor
Conference on small modular nuclear reactor. Submit abstract by Apr 11.
asmeconferences.org/nucle

Thank y

genevi
Here's a
Whole E
"what ha
tense. p

This conversation is mis

Vimeo's Web Page
Spring 2010

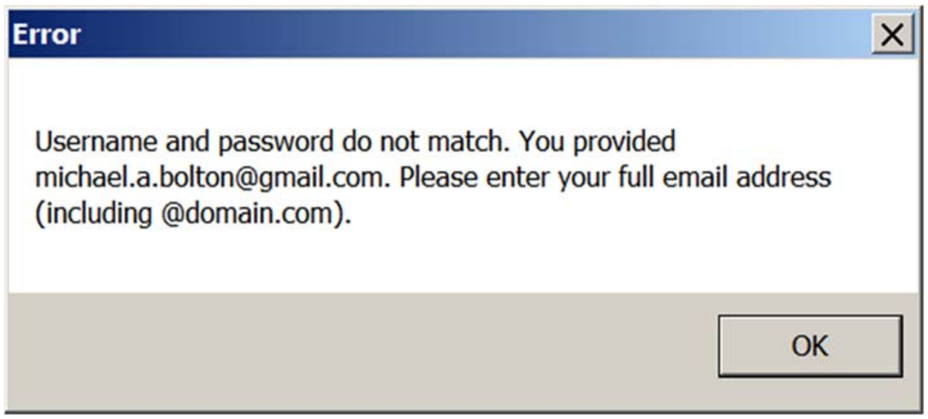

Google Chrome Installer

The installer encountered error 3.

Show me help for this issue.

Close

Google Chrome

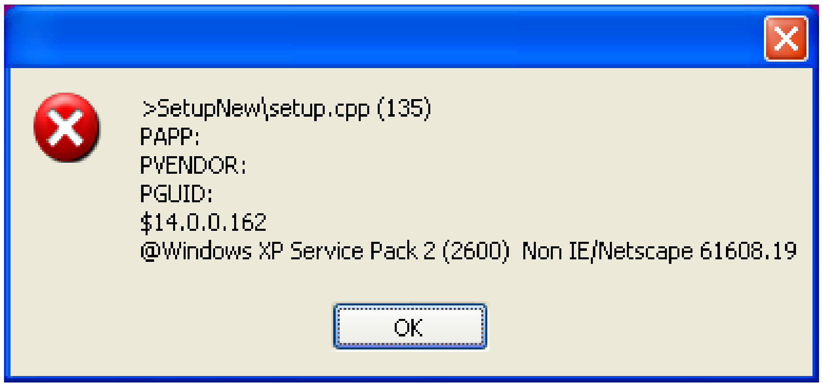


Error

Username and password do not match. You provided michael.a.bolton@gmail.com. Please enter your full email address (including @domain.com).

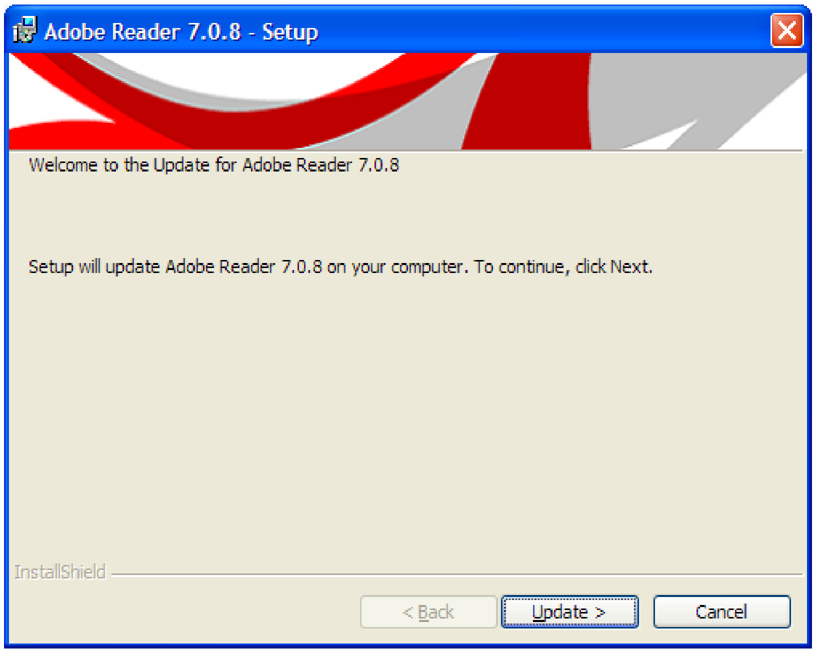
Google Calendar Update

OK, fine.



```
>SetupNew\setup.cpp (135)
PAPP:
PVENDOR:
PGUID:
$14.0.0.162
@Windows XP Service Pack 2 (2600)  Non IE/Netscape 61608.19
```

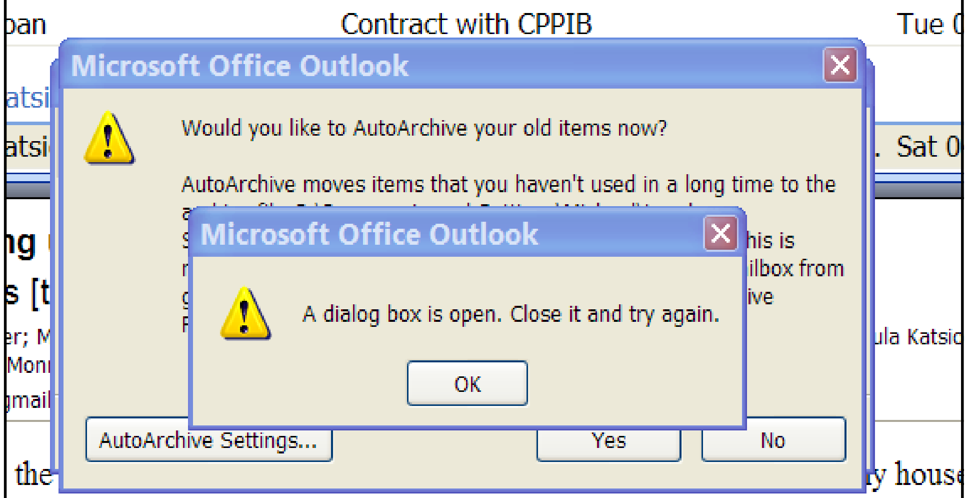
Don't Know Who This Was


Adobe Reader 7.0.8 - Setup

Welcome to the Update for Adobe Reader 7.0.8

Setup will update Adobe Reader 7.0.8 on your computer. To continue, click Next.

Adobe Acrobat


Microsoft Office Outlook

Would you like to AutoArchive your old items now?

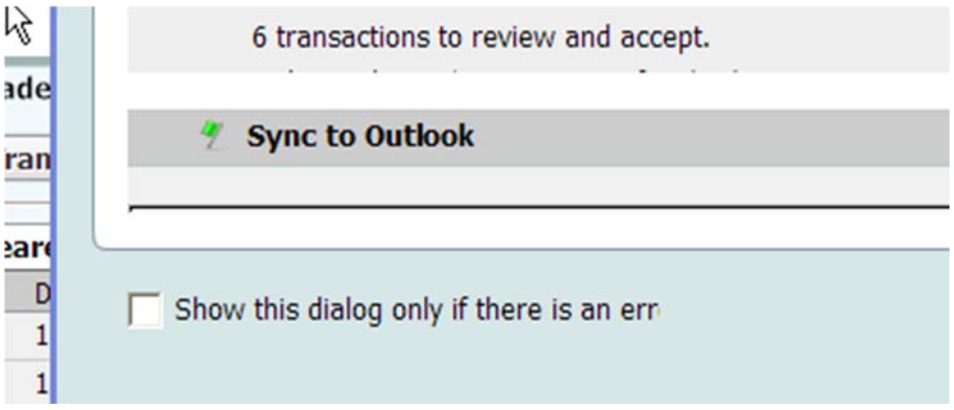
AutoArchive moves items that you haven't used in a long time to the

Microsoft Office Outlook

A dialog box is open. Close it and try again.

Microsoft Outlook

I Already See an Err!

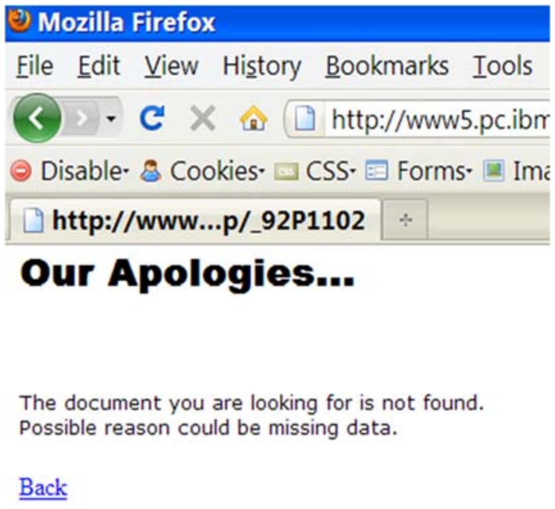


6 transactions to review and accept.

Show this dialog only if there is an err

Intuit Quicken

Hmm… I suppose that's possible.


Mozilla Firefox
File Edit View History Bookmarks Tools
http://www5.pc.ibm
Disable· Cookies· CSS· Forms· Ima
http://www...p/_92P1102

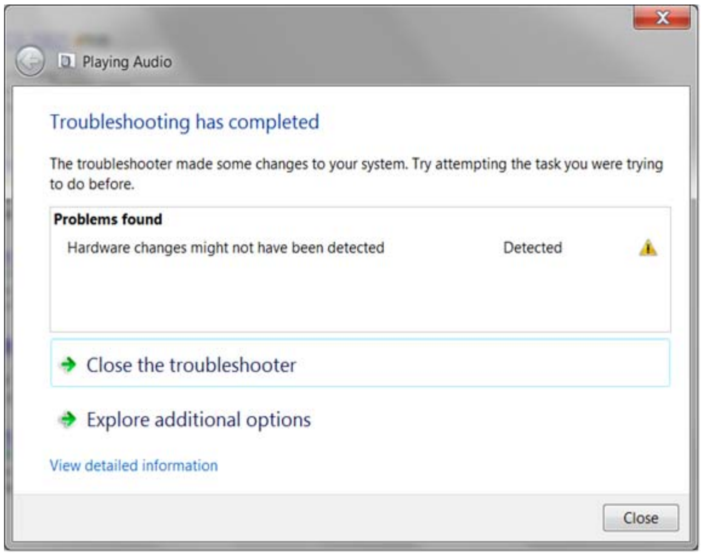
Our Apologies...

The document you are looking for is not found.
Possible reason could be missing data.

Back

ibm.com

Well, were they detected or not?


Playing Audio

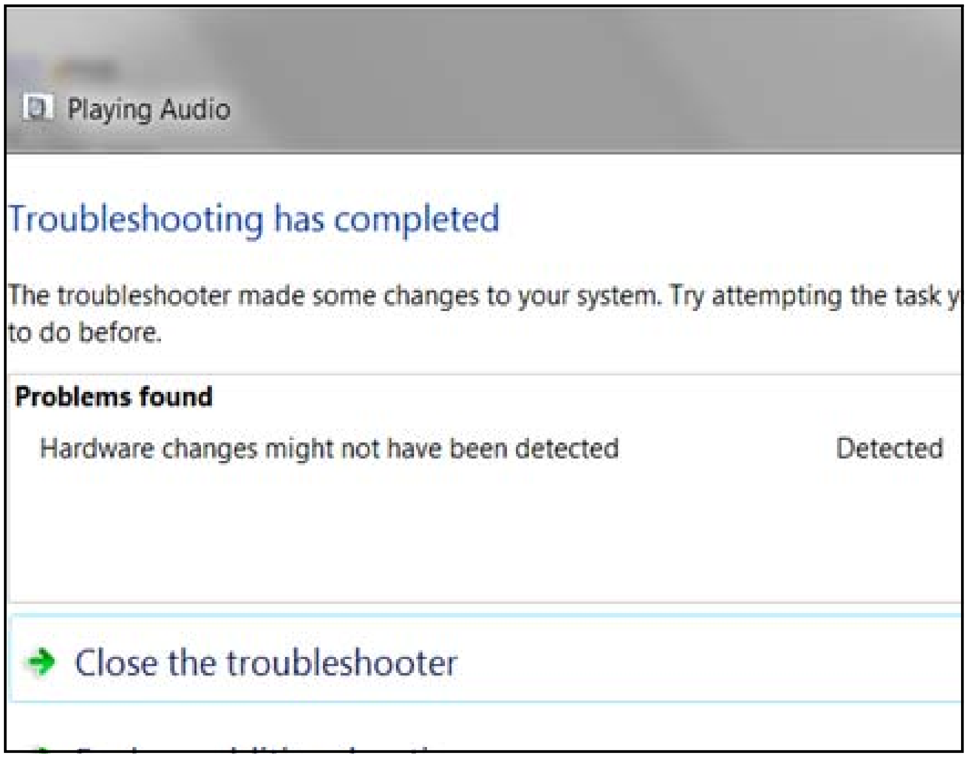
Troubleshooting has completed

The troubleshooter made some changes to your system. Try attempting the task you were trying to do before.

Problems found

Hardware changes might not have been detected Detected

Windows

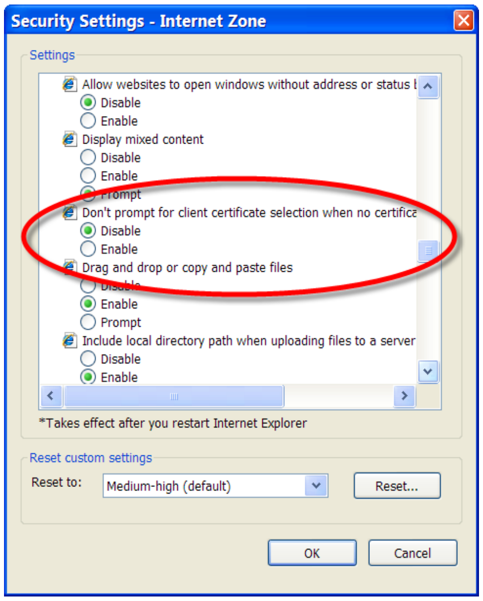

Playing Audio

Troubleshooting has completed

The troubleshooter made some changes to your system. Try attempting the task y
to do before.

Problems found

Hardware changes might not have been detected Detected



Security Settings - Internet Zone

Settings

- Allow websites to open windows without address or status b
  - [✓ Disable]
  - Enable
- Display mixed content
  - Disable
  - Enable
  - Prompt
- Don't prompt for client certificate selection when no certifica
  - [✓ Disable]
  - Enable
- Drag and drop or copy and paste files
  - Disable
  - [✓ Enable]
  - Prompt
- Include local directory path when uploading files to a server
  - Disable
  - [✓ Enable]

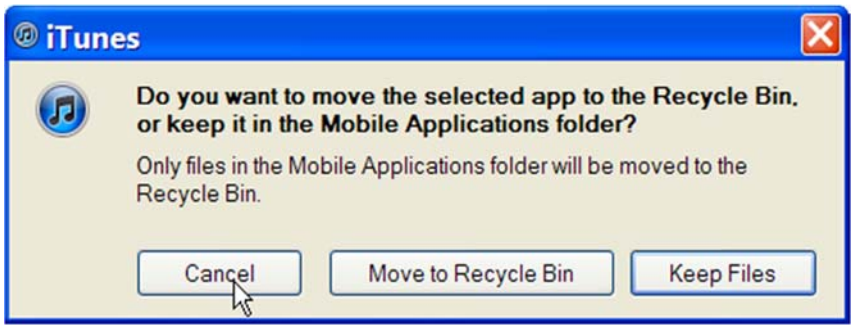
*Takes effect after you restart Internet Explorer

Reset custom settings

Reset to: Medium-high (default)

Windows

Go home, iTunes. You're drunk.



iTunes

Do you want to move the selected app to the Recycle Bin, or keep it in the Mobile Applications folder?

Only files in the Mobile Applications folder will be moved to the Recycle Bin.

iTunes

A Clarification

- "How To Get What You Really Want From Testing" is a potentially misleading title.
- As a member of the Context-Driven School of Software Testing, I maintain there are no *universally* best results available from testing, just as there are no best practices…
- …and, of course, I'm not a mind reader.
- In this presentation, I'll try to help introduce ideas intended to help you choose or assign or develop testing missions that line up with *what managers and other testing clients typically need to know, in my experience*.
- Only you can really decide what's best *for you*.

I'm Biased Towards Rapid Testing

Rapid testing is a *mind-set*
and a *skill-set* **of testing**
focused on how to do testing
more quickly,
less expensively,
with *excellent results*.

This is a general testing methodology. It adapts to any kind of project or product.

Read more about Rapid Testing at http://www.developsense.com

A Computer Program

A set of instructions
for a computer.

See the Association for Software Testing's
Black Box Software Testing Foundations course, Cem Kaner & James Bach

A House

A set of building materials,
arranged in the
"House" design pattern.

A House

Something for people to live in.

Kaner's Definition of a Computer Program

- A computer program is
- a *communication*
- among several people
- and computers
- separated over distance and time
- that contains instructions that can be run on a computer.

The purpose of a computer program is
to provide **value** to **people**

Implications of Kaner’s Definition

- A computer program is **far more** than its code
- A software product is **far more** than the instructions for the device
- Quality is **far more** than the absence of errors in the code.
- Testing is **far more** than writing some code to confirm that other code returns a “correct” result.

Quality is value to some person(s).

—Jerry Weinberg

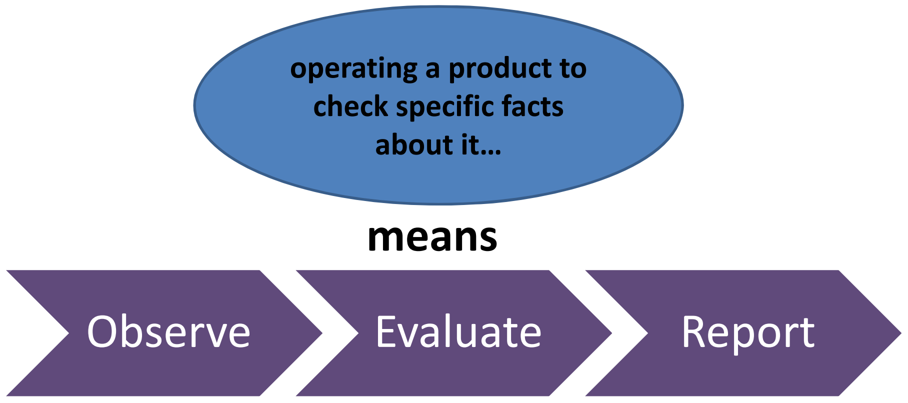
Software testing is the investigation of *systems* consisting of people and their work, computers, programs, and the relationships between them.

How to Test

- Write test cases?
- Play with the produ
- Use the all-pairs test t
- Use equivalence class partitioning?
- Use record and playback automation?
- Read the spec?
- Report bugs? Help design the product?
- Write code for the product?
- Play ping pong in the break room?

Yes, learn these things…
But there is no formula
for DOING them.

Call this "Checking" not Testing



Interact with the product in specific ways to collect specific observations.

Apply algorithmic decision rules to those observations.

Report any failed checks.

A Check Has Three Elements

1. An *observation* linked to…
2. A *decision rule* such that…
3. both observation and decision rule can be applied algorithmically.

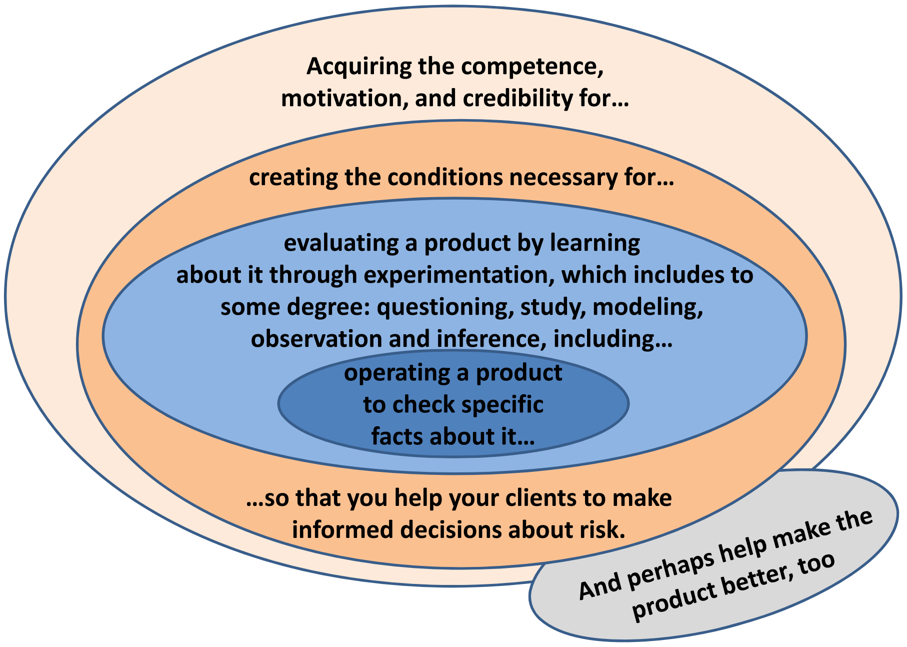
A ***check*** can be performed

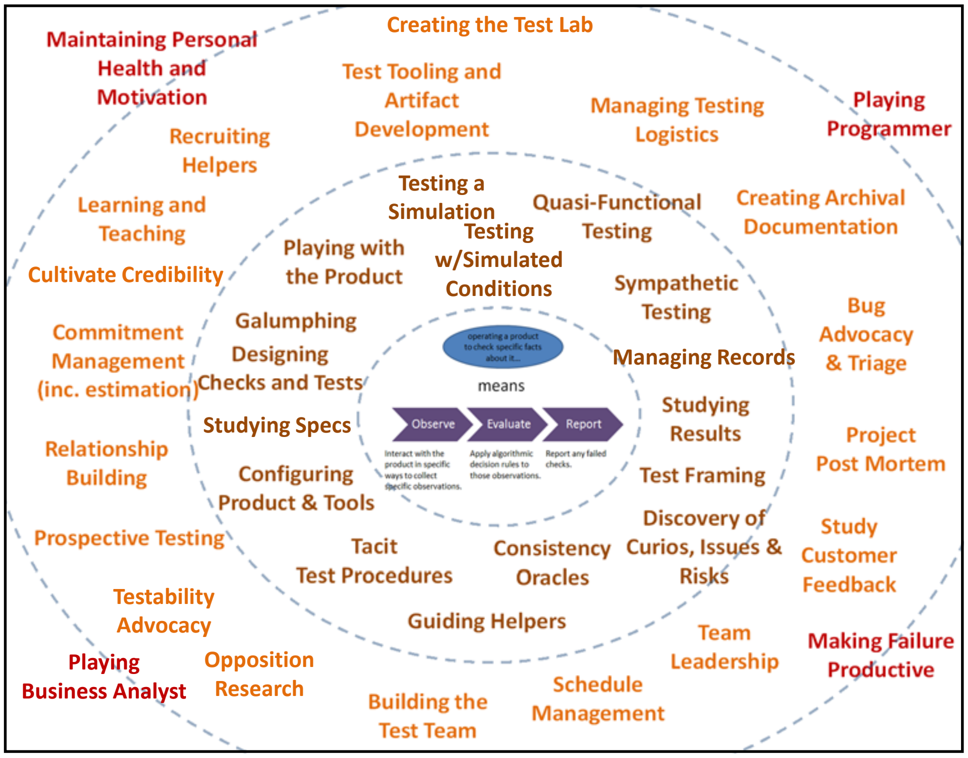
by a machine
that *can't* think
(but that is quick and precise)

by a human who has been
instructed *not* to think
(and who is slow and variable)

See http://www.satisfice.com/blog/archives/856

Testing is…


Acquiring the competence, motivation, and credibility for…
creating the conditions necessary for…
evaluating a product by learning about it through experimentation, which includes to some degree: questioning, study, modeling, observation and inference, including…
operating a product to check specific facts about it…
…so that you help your clients to make informed decisions about risk.
And perhaps help make the product better, too


Maintaining Personal Health and Motivation
Creating the Test Lab
Test Tooling and Artifact Development
Managing Testing Logistics
Playing Programmer
Recruiting Helpers
Learning and Teaching
Testing a Simulation
Quasi-Functional Testing
Creating Archival Documentation
Playing with the Product
Testing w/Simulated Conditions
Cultivate Credibility
Sympathetic Testing
Bug Advocacy & Triage
Galumphing
Commitment Management (inc. estimation)
Designing Checks and Tests
operating a product to check specific facts about it…
means
Observe
Evaluate
Report
Interact with the product in specific ways to collect specific observations.
Apply algorithmic decision rules to those observations.
Report any failed checks.
Managing Records
Studying Results
Studying Specs
Relationship Building
Project Post Mortem
Configuring Product & Tools
Test Framing
Prospective Testing
Discovery of Curios, Issues & Risks
Study Customer Feedback
Tacit Test Procedures
Consistency Oracles
Testability Advocacy
Guiding Helpers
Team Leadership
Making Failure Productive
Playing Business Analyst
Opposition Research
Building the Test Team
Schedule Management

What Is Testing?

- Excellent testing is not merely a branch of computer science
  - testing *includes* computer science, mathematics, technical domains
  - BUT… focus only on programs and functions, and you leave out questions of *value* and other relationships that include people
- To me, excellent testing is more like *anthropology*—interdisciplinary, systems-focused, investigative, storytelling

Biology

Archaeology

Language

Culture

Why is context such a big deal?

Why Is Context A Big Deal?

Because increasingly, software mediates not only mimeomorphic actions, but also polimorphic actions.

There's your answer!

Thanks, and have a great day!!

Behaviour

- the physical counterpart of an action. (The other part of action is intention)
- behaviours are observable and describable *without* reference to context.
- Examples: a blink, a scratch, a clap

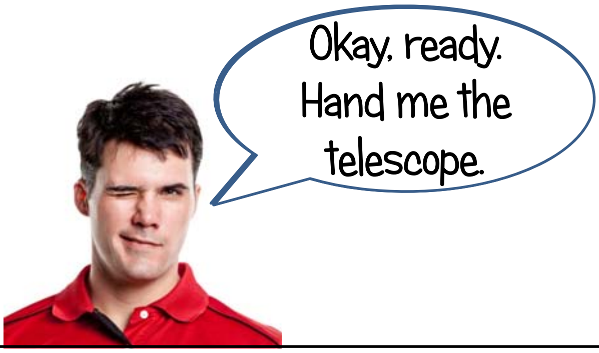
See Harry Collins, *The Shape of Actions* and *Tacit and Explicit Knowledge*

Mimeomorphic Action

- an action that can always be carried out by repeating the behaviour, irrespective of context
- can therefore be mimicked by machines which do not need to understand the context

Polimorphic Action

- An action where the *behaviour* must be changed according to context in order to carry out the *action.* (e.g. a greeting, a love letter)
- note that sometimes the same *behaviour* can instantiate different *actions* according to context.



Harry Collins on Software Testing

“Computers and their software are two things. As collections of interacting cogs they must be ‘checked’ to make sure there are no missing teeth and the wheels spin together nicely. Machines are also ‘social prostheses’, fitting into social life where a human once fitted. It is a characteristic of medical prostheses, like replacement hearts, that they do not do exactly the same job as the thing they replace; the surrounding body compensates.

Abstract, “Machines as Social Prostheses”, EuroSTAR 2013

Harry Collins on Software Testing

“Contemporary computers cannot do just the same thing as humans because they do not fit into society as humans do, so the surrounding society must compensate for the way the computer fails to reproduce what it replaces. This means that a complex judgment is needed to test whether software fits well enough for the surrounding humans to happily ‘repair’ the differences between humans and machines. This is much more than a matter of deciding whether the cogs spin right.”

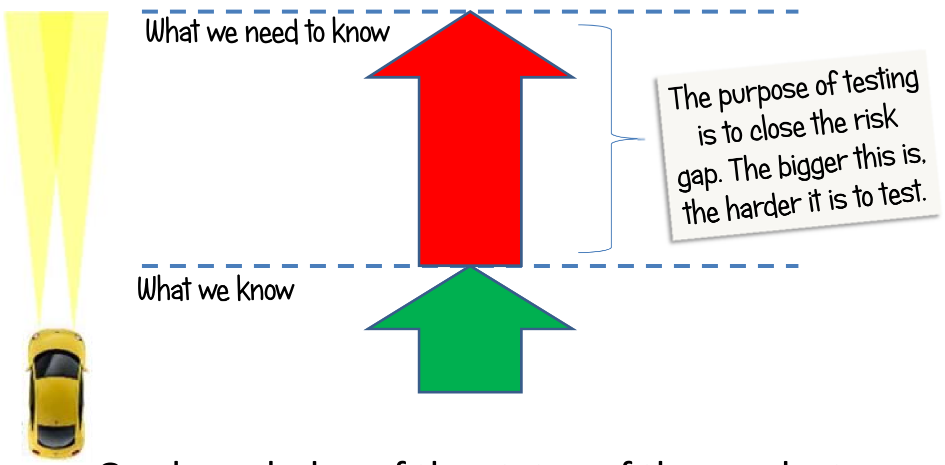
Abstract, “Machines as Social Prostheses”, EuroSTAR 2013

2 + 2 = 4

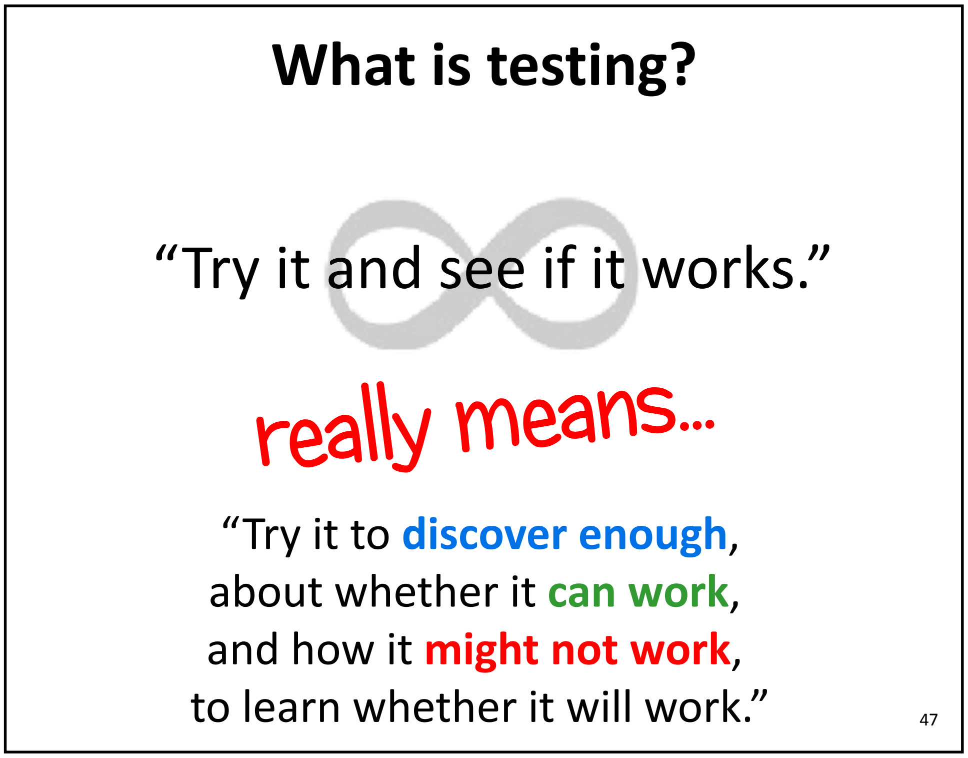
- The automated check was programmed to enter
  2
  and then +
  and then 2
  and then =
- The program returned 4. (YAY!)
- The automated check logged "pass!"
- A customer was still unhappy.

WHY?

Testers Light the Way: *The Risk Gap*


What we need to know
What we know
The purpose of testing is to close the risk gap. The bigger this is, the harder it is to test.

Our knowledge of the status of the product.



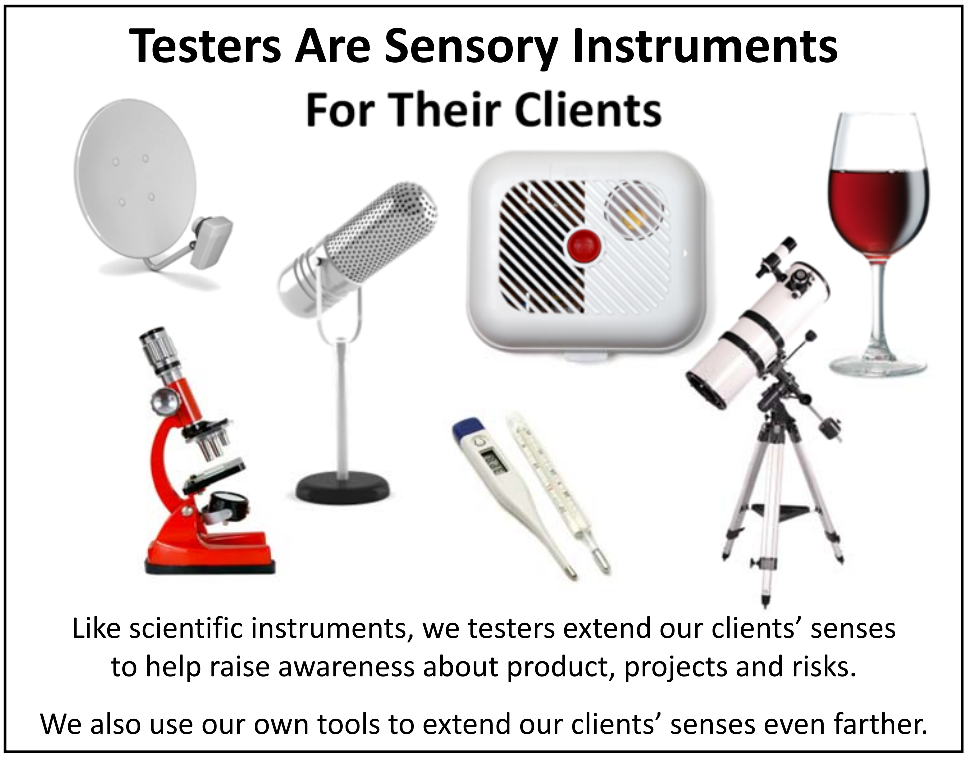
What is testing?

"Try it and see if it works."

really means...

"Try it to **discover enough**,
about whether it **can work**,
and how it **might not work**,
to learn whether it will work."

47



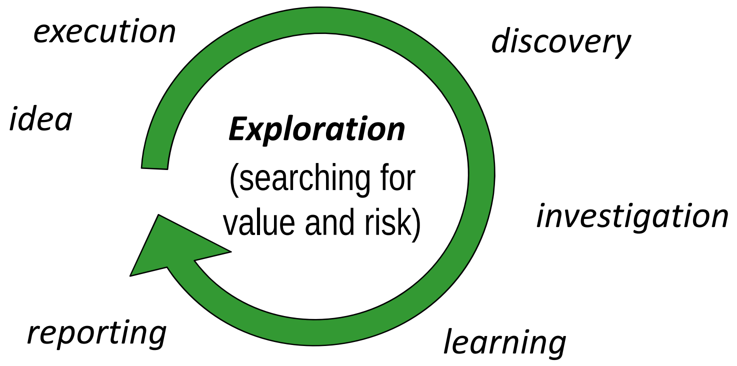
Testers Are Sensory Instruments For Their Clients

Like scientific instruments, we testers extend our clients' senses
to help raise awareness about product, projects and risks.

We also use our own tools to extend our clients' senses even farther.

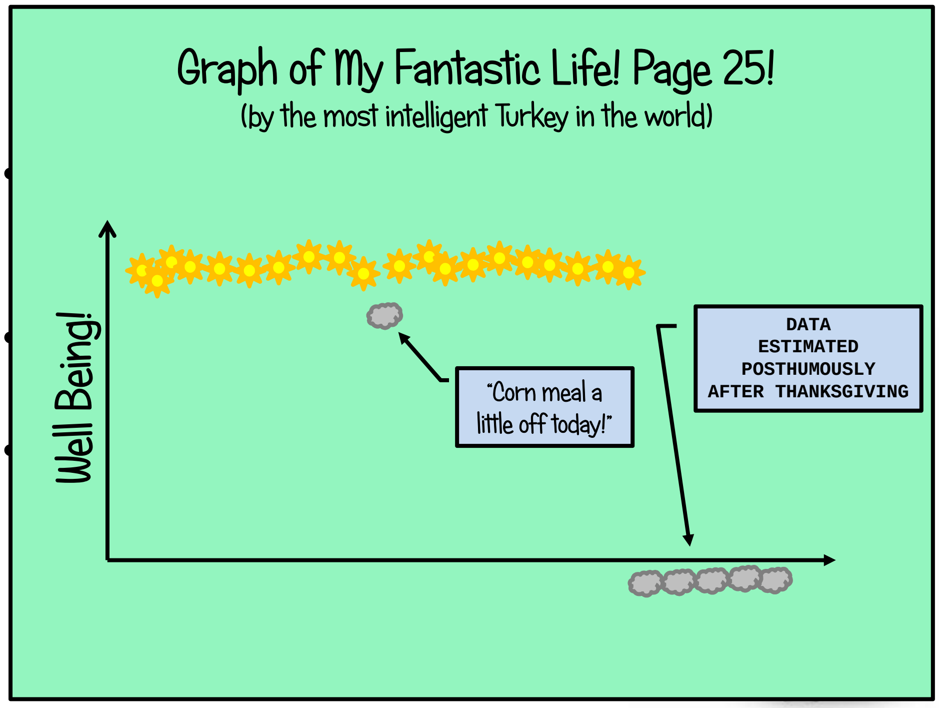
Testing is about *Learning* and *Adaptation*

Skilled testers help to defend the value of the product by *learning* about it on behalf of our clients.



…and there are little loops between all of these things.

Confirmatory testing *feels* positive,
but prevents insight and conceals risk.


Graph of My Fantastic Life! Page 25!
(by the most intelligent Turkey in the world)
Well Being!
"Corn meal a little off today!"
DATA ESTIMATED POSTHUMOUSLY AFTER THANKSGIVING

Don't Be A Turkey!

- No experience of the past can LOGICALLY be projected into the future, because we have no experience OF the future.
- This is no big deal in a world of stable, simple patterns.
- **BUT NEITHER SOFTWARE NOR PROJECTS ARE STABLE OR SIMPLE.**
- **"PASSING" TESTS CANNOT PROVE SOFTWARE GOOD.**

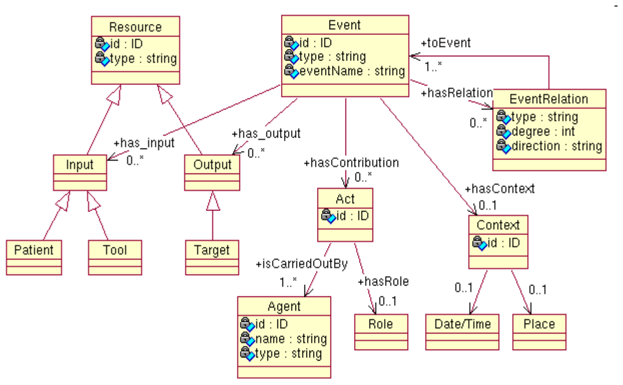
Based on a story told by Nassim Taleb, who stole it from Bertrand Russell, who stole it from David Hume.

Modeling

- How should I test?
- To test well, you must *learn to test*.
  For that, you need good models of testing.
- How should I test *my product*?
- To test a product you need a *diversified strategy*.
  For that, you need to model product-specific risks.
- How do I find important problems?
- To find problems, you must *cover the product*.
  For that, you need to model the product.
- How do I know when something is a problem?
- To recognize problems, you must *apply oracles*.
  For that, you need to learn about oracles.

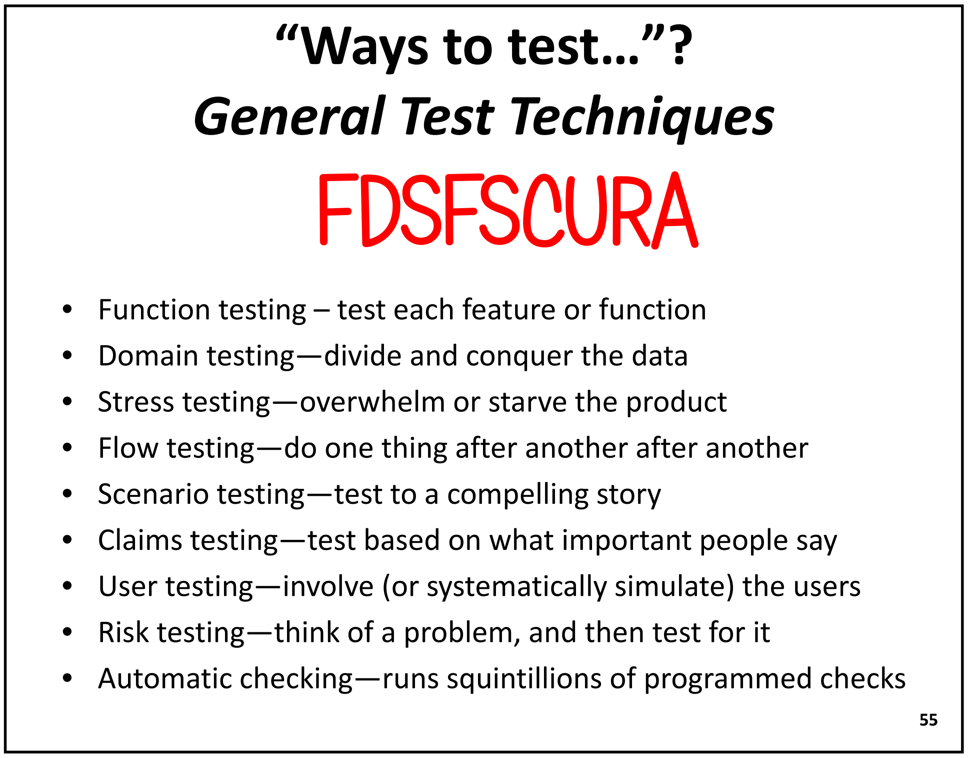
Models

- A model is a simpler representation of a more complex idea, object, or system that helps you to understand, control, observe, or explore it.



Project Environment

- **Mission**
  - The problem we are here to solve for our customer.
- **Information**
  - Information about the product or project that is needed for testing.
- **Developer relations**
  - How you get along with the programmers.
- **Team**
  - Anyone who will perform or support testing.
- **Equipment & tools**
  - Hardware, software, or documents required to administer testing.
- **Schedule**
  - The sequence, duration, and synchronization of project events.
- **Test Items**
  - The product to be tested.
- **Deliverables**
  - The observable products of the test project.

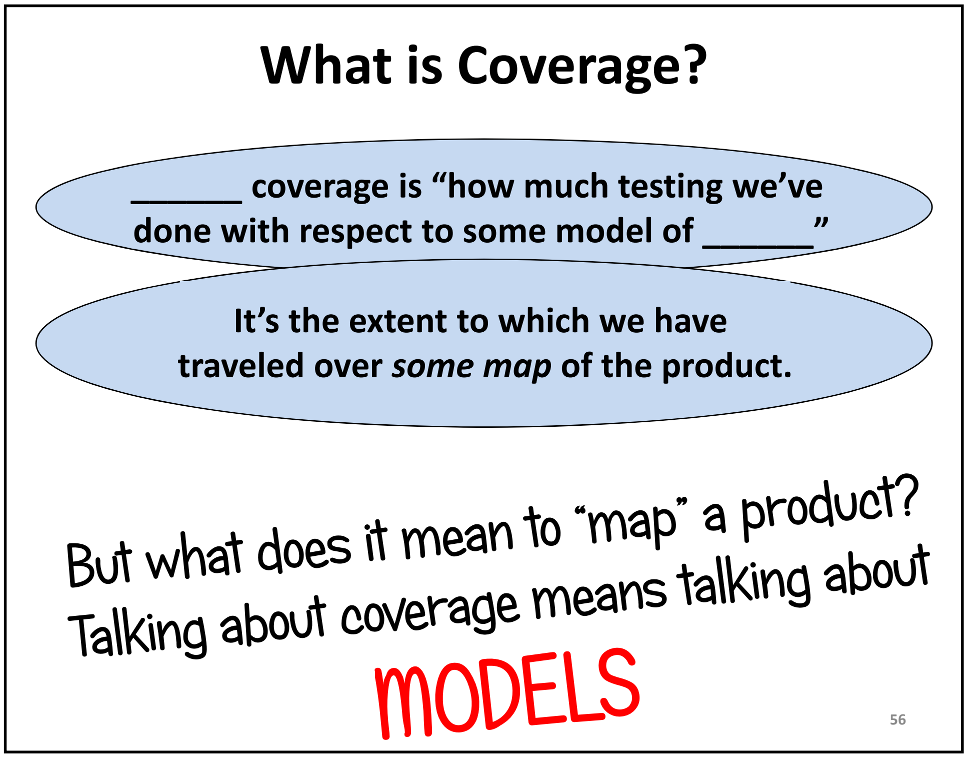


"Ways to test…"? *General Test Techniques*

FDSFSCURA

- Function testing – test each feature or function
- Domain testing—divide and conquer the data
- Stress testing—overwhelm or starve the product
- Flow testing—do one thing after another after another
- Scenario testing—test to a compelling story
- Claims testing—test based on what important people say
- User testing—involve (or systematically simulate) the users
- Risk testing—think of a problem, and then test for it
- Automatic checking—runs squintillions of programmed checks

55



What is Coverage?

______ coverage is "how much testing we've done with respect to some model of ______"

It's the extent to which we have traveled over *some map* of the product.

But what does it mean to "map" a product?
Talking about coverage means talking about

MODELS

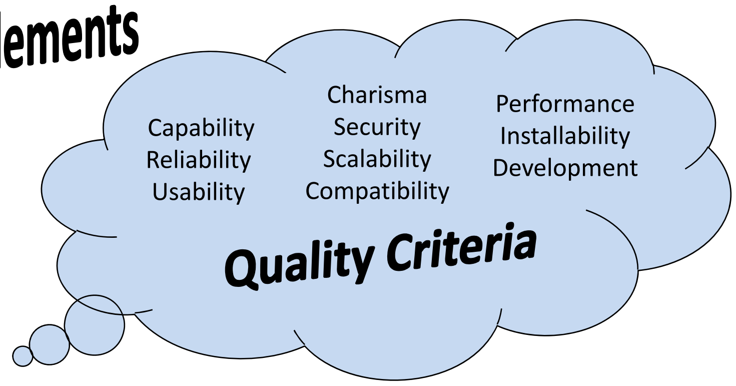
56

One Way to Model Coverage: Product Elements (with Quality Criteria)

SFDIPOT – "San Francisco Depot"

Product Elements

- Structure
- Function
- Data
- Interfaces
- Platform
- Operations
- Time



57

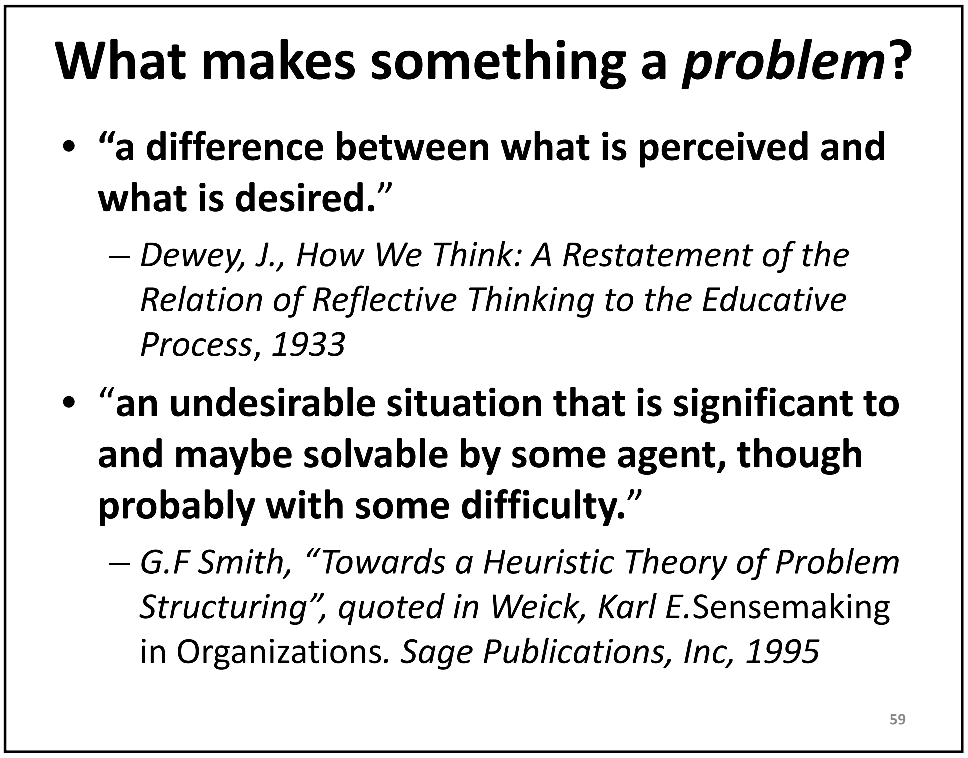
Quality Criteria Where's the Value to People?

CRUCSSCPID

Capability	Scalability
Reliability	Compatibility
Usability	Performance
Charisma	Installability
Security	Development

Our job is to identify value, and threats to value. Many test approaches focus on capability (functionality) and underemphasize the other criteria.

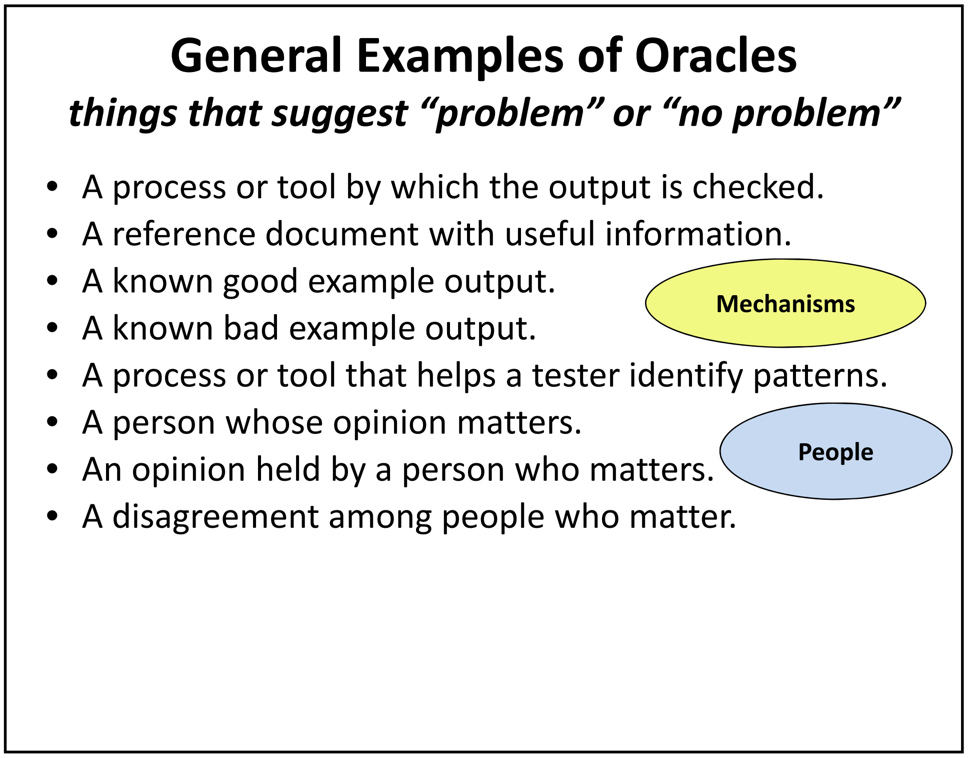
58



What makes something a *problem*?

- **“a difference between what is perceived and what is desired.”**
  - *Dewey, J., How We Think: A Restatement of the Relation of Reflective Thinking to the Educative Process, 1933*
- **“an undesirable situation that is significant to and maybe solvable by some agent, though probably with some difficulty.”**
  - *G.F Smith, “Towards a Heuristic Theory of Problem Structuring”, quoted in Weick, Karl E.*Sensemaking in Organizations. *Sage Publications, Inc, 1995*

59



General Examples of Oracles

things that suggest “problem” or “no problem”

- A process or tool by which the output is checked.
- A reference document with useful information.
- A known good example output.
- A known bad example output.
- A process or tool that helps a tester identify patterns.
- A person whose opinion matters.
- An opinion held by a person who matters.
- A disagreement among people who matter.

Mechanisms

People

How Do People React to Software?

Impatience

Frustration

Amusement

Surprise

Confusion

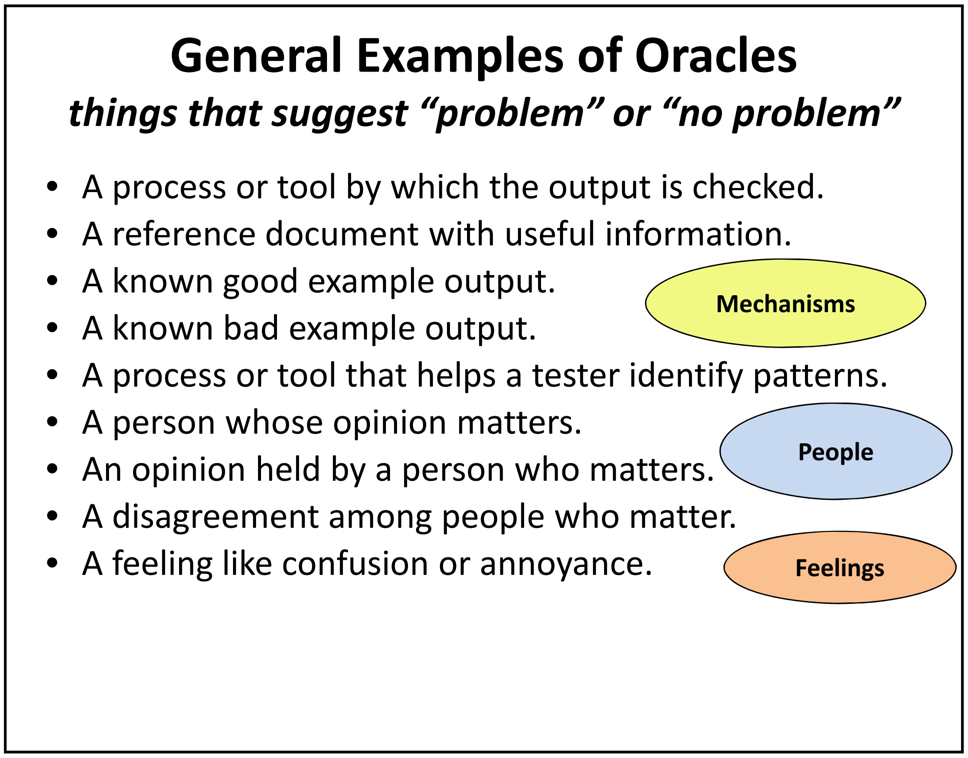
Annoyance

What Might Feelings Tell Us?

Impatience	⇨ a threat to performance?
Frustration	⇨ a threat to capability?
Fear	⇨ a threat to security?
Surprise	⇨ a threat to reliability?
Confusion	⇨ a threat to usability? to testability?
Annoyance	⇨ a threat to charisma?
Boredom	⇨ an insignificant test?
Tiredness	⇨ time for a break?
Anxiety	⇨ a need for a particular skill?
Curiosity	⇨ a pointer to useful investigation?

Bugs!

Issues!



General Examples of Oracles

things that suggest "problem" or "no problem"

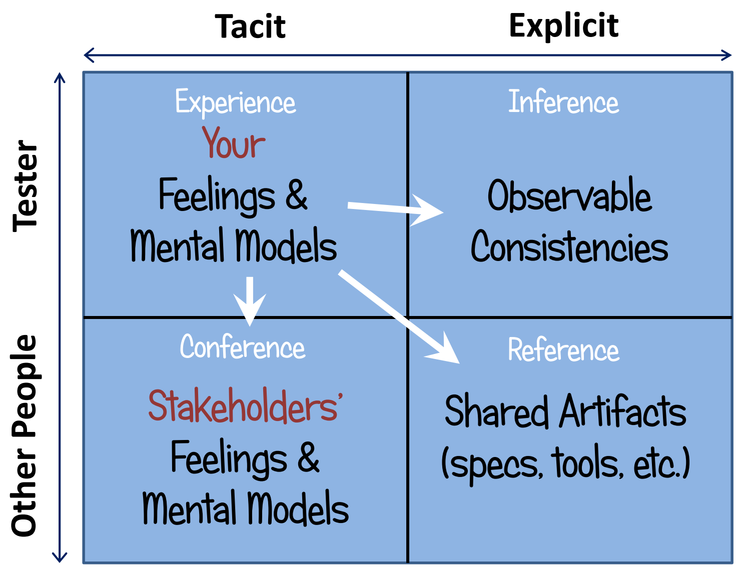
- A process or tool by which the output is checked.
- A reference document with useful information.
- A known good example output.
- A known bad example output.
- A process or tool that helps a tester identify patterns.
- A person whose opinion matters.
- An opinion held by a person who matters.
- A disagreement among people who matter.
- A feeling like confusion or annoyance.

Mechanisms

People

Feelings

Oracles from the Inside Out



Consistency ("this agrees with that")

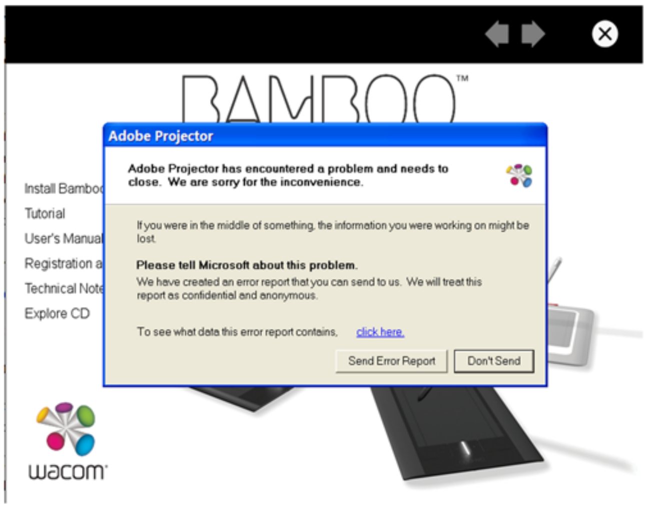
an important theme in oracle principles

- **Familiarity:** The system ***is not consistent*** with the pattern of any familiar problem.
- **Explainability:** The system *is consistent* with our ability to describe it clearly.
- **World:** The system *is consistent* with things that we recognize in the world.
- **History:** The present version of the system *is consistent* with past versions of it.
- **Image:** The system *is consistent* with an image that the organization wants to project.
- **Comparable Products:** The system *is consistent* with comparable systems.
- **Claims:** The system *is consistent* with what important people say it's supposed to be.
- **Users' Expectations:** The system *is consistent* with what users want.
- **Product:** Each element of the system is *consistent* with comparable elements in the same system.
- **Purpose:** The system *is consistent* with its purposes, both explicit and implicit.
- **Standards and Statutes:** The system *is consistent* with applicable laws, or relevant implicit or explicit standards.

Consistency heuristics rely on the quality of your models of the product and its context.

65

Familiar Problems



Adobe Projector

Adobe Projector has encountered a problem and needs to close. We are sorry for the inconvenience.

If you were in the middle of something, the information you were working on might be lost.

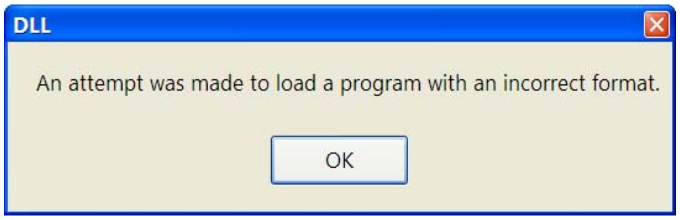
Please tell Microsoft about this problem.

We have created an error report that you can send to us. We will treat this report as confidential and anonymous.

To see what data this error report contains, click here.

If a product is consistent with problems we've seen before, we suspect that there might be a problem.

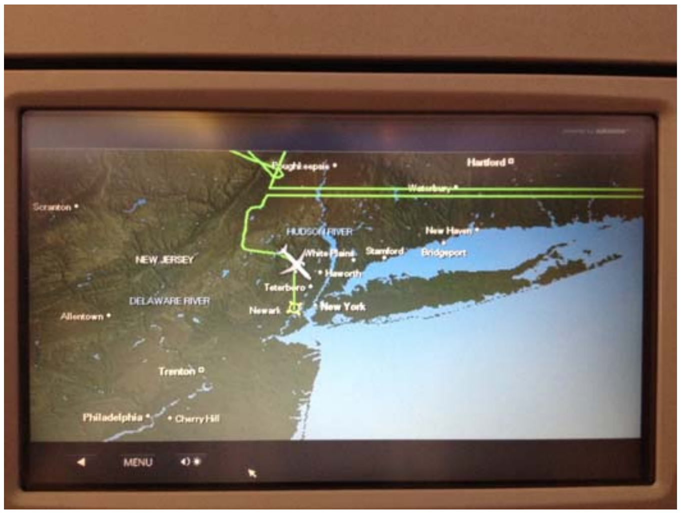
Explainability


DLL

An attempt was made to load a program with an incorrect format.

If a product is inconsistent with our ability to explain it, we suspect that there might be a problem.

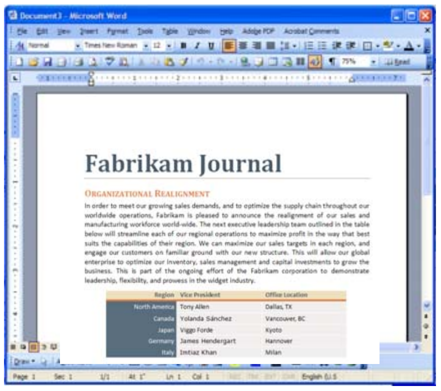
World

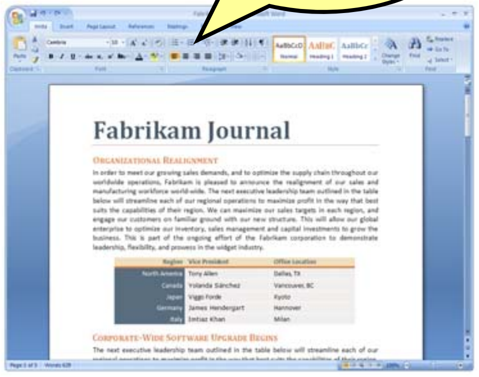


If a product is inconsistent with the way the world works, we suspect that there might be a problem.

History

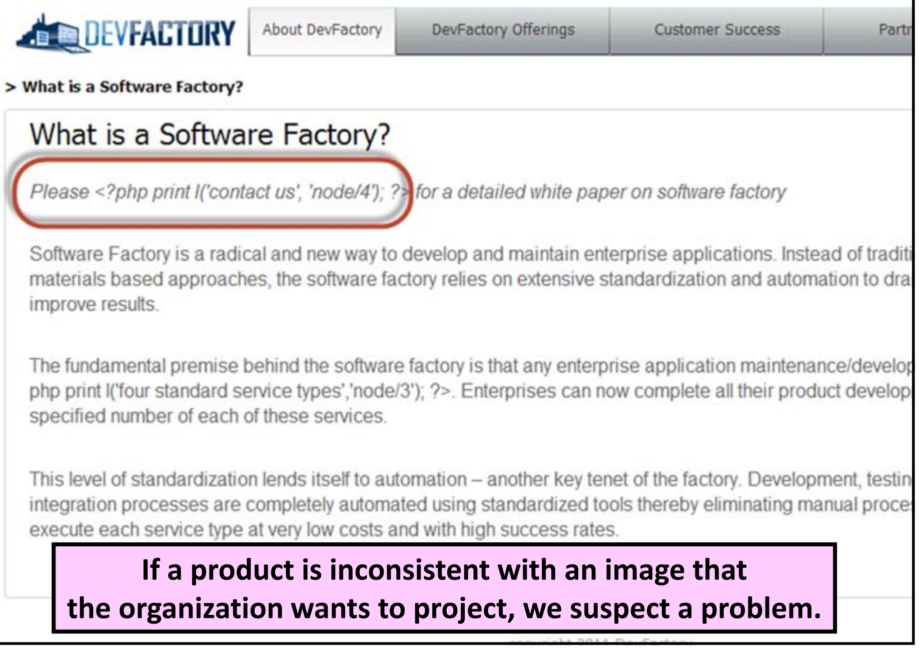
Hey, I liked the menu bar! How the #&@ do I print now?





If a product is inconsistent with previous versions of itself, we suspect that there might be a problem.

Image



> What is a Software Factory?

What is a Software Factory?

Please <?php print l('contact us', 'node/4'); ?> for a detailed white paper on software factory

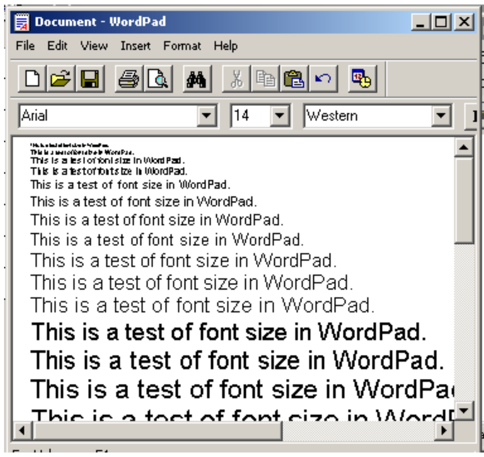
Software Factory is a radical and new way to develop and maintain enterprise applications. Instead of traditi
materials based approaches, the software factory relies on extensive standardization and automation to dra
improve results.

The fundamental premise behind the software factory is that any enterprise application maintenance/develop
php print l('four standard service types','node/3'); ?>. Enterprises can now complete all their product develop
specified number of each of these services.

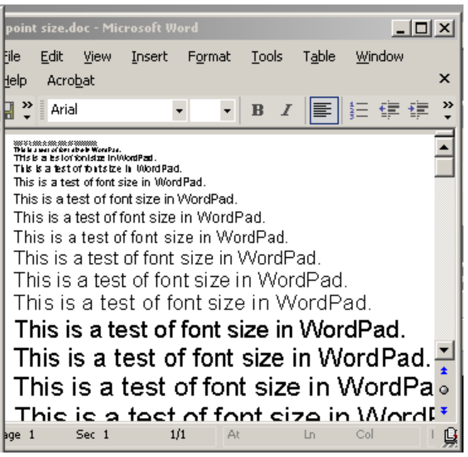
This level of standardization lends itself to automation – another key tenet of the factory. Development, testin
integration processes are completely automated using standardized tools thereby eliminating manual proce
execute each service type at very low costs and with high success rates.

If a product is inconsistent with an image that the organization wants to project, we suspect a problem.

Comparable Products



This is a test of font size in WordPad.
This is a test of font size in WordPad.
This is a test of font size in WordPad.
This is a test of font size in WordPad.
This is a test of font size in WordPad.
This is a test of font size in WordPad.
This is a test of font size in WordPad.
This is a test of font size in WordPad.
This is a test of font size in WordPad.
This is a test of font size in WordPad.
This is a test of font size in WordPad.
This is a test of font size in WordPa


point size.doc - Microsoft Word
File Edit View Insert Format Tools Table Window
Help Acrobat
Arial
Page 1 Sec 1 1/1 At Ln Col

This is a test of font size in WordPad.
This is a test of font size in WordPad.
This is a test of font size in WordPad.
This is a test of font size in WordPad.
This is a test of font size in WordPad.
This is a test of font size in WordPad.
This is a test of font size in WordPad.
This is a test of font size in WordPad.
This is a test of font size in WordPad.
This is a test of font size in WordPad.
This is a test of font size in WordPa

WordPad **Word**

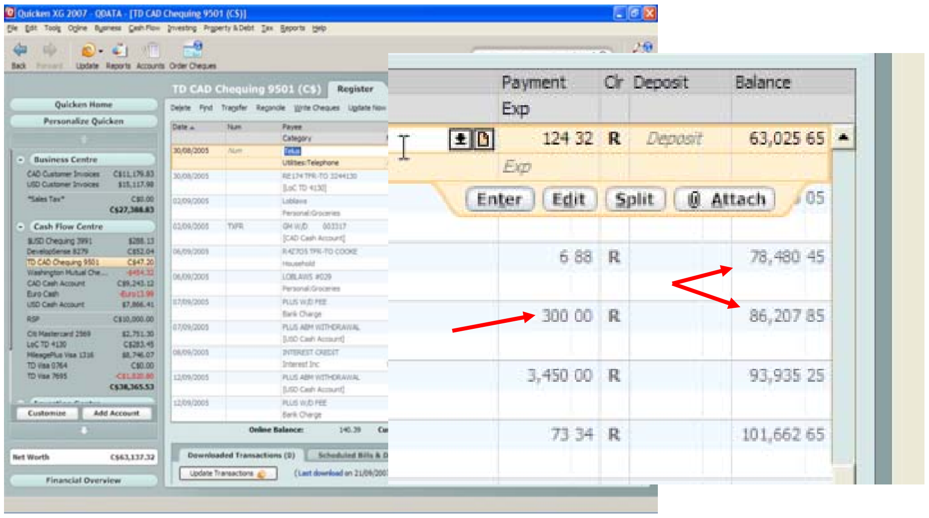
When a product seems inconsistent with a comparable product or algorithm, we suspect that there might be a problem.

Claims

New! Supports Mac OS/X!
URD structure
Part 1 - General Description
Part 2 - Key User Requirements
Part 3 - User Requirements
Part 4 - Context Document
Part 5 - Glossary

When a product is inconsistent with claims that important people make about it, we suspect a problem.

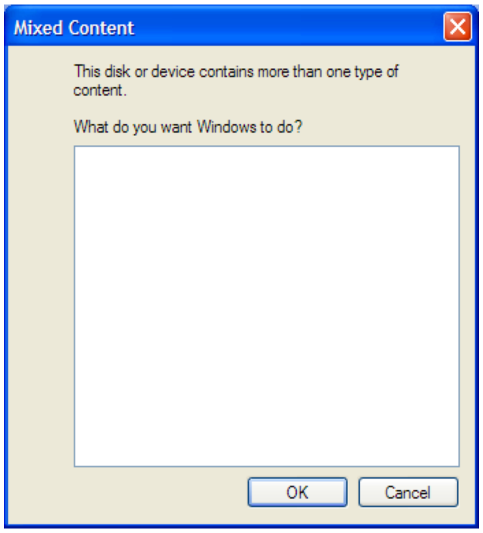
User Expectations


Quicken XG 2007 - QDATA - [TD CAD Chequing 9501 (C$)]
TD CAD Chequing 9501 (C$)
Register
Payment
Exp
Clr
Deposit
Balance
Enter
Edit
Split
Attach

Payment	Clr	Deposit	Balance
124 32	R	Deposit	63,025 65
6 88	R		78,480 45
300 00	R		86,207 85
3,450 00	R		93,935 25
73 34	R		101,662 65

When a product is inconsistent with expectations that a reasonable user might have, we suspect a problem.

Purpose

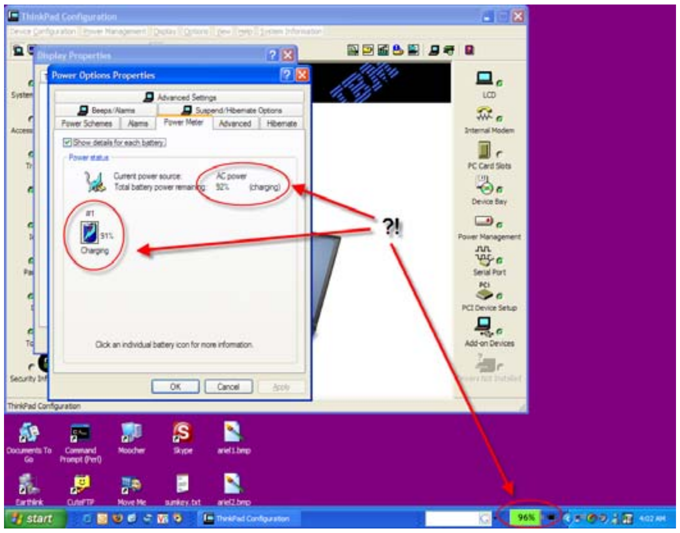


This disk or device contains more than one type of content.

What do you want Windows to do?

When a product is inconsistent with its designers' explicit or implicit purposes, we suspect a problem.

Product



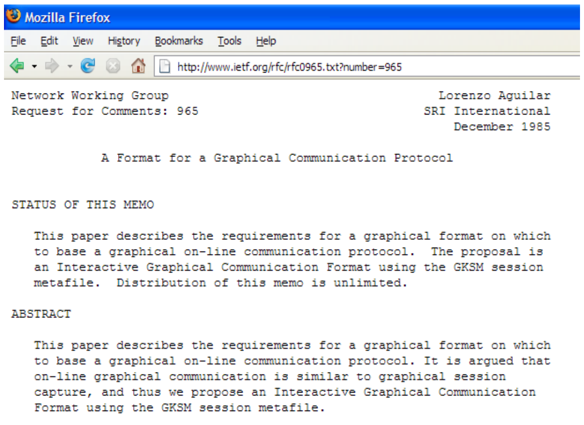
Current power source: AC power
Total battery power remaining: 92% (charging)

#1
91%
Charging

Click an individual battery icon for more information.

When a product is inconsistent internally—as when it contradicts itself—we suspect a problem.

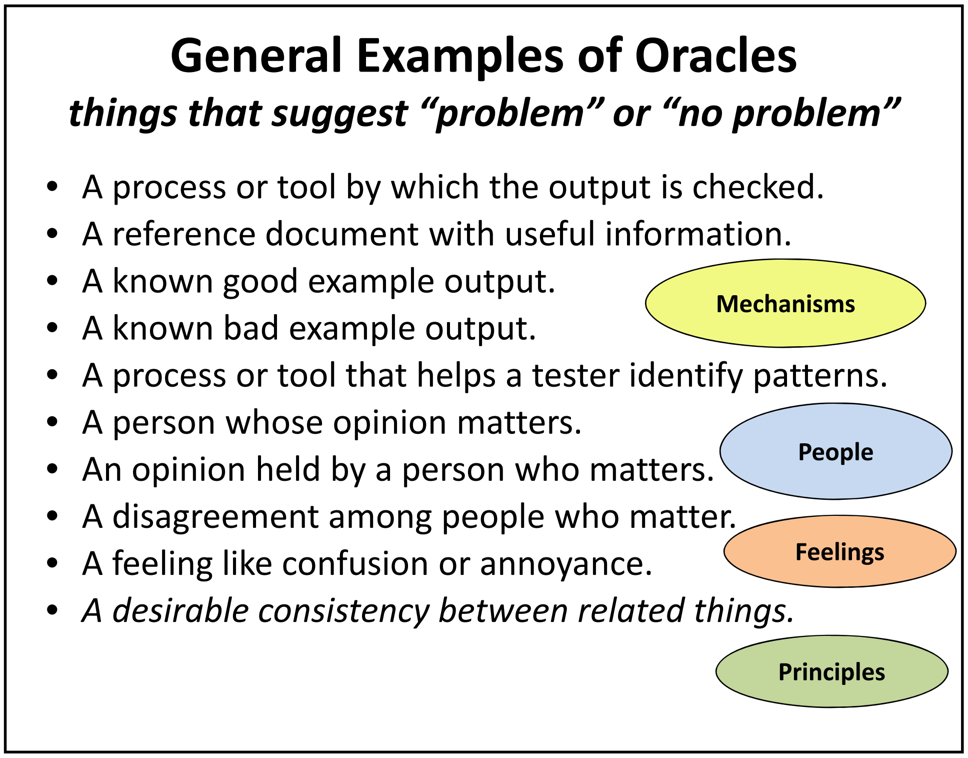
Statutes and Standards



Mozilla Firefox
File Edit View History Bookmarks Tools Help
http://www.ietf.org/rfc/rfc0965.txt?number=965

```
Network Working Group                                   Lorenzo Aguilar
Request for Comments: 965                             SRI International
                                                          December 1985

              A Format for a Graphical Communication Protocol


STATUS OF THIS MEMO

   This paper describes the requirements for a graphical format on which
   to base a graphical on-line communication protocol.  The proposal is
   an Interactive Graphical Communication Format using the GKSM session
   metafile.  Distribution of this memo is unlimited.

ABSTRACT

   This paper describes the requirements for a graphical format on which
   to base a graphical on-line communication protocol. It is argued that
   on-line graphical communication is similar to graphical session
   capture, and thus we propose an Interactive Graphical Communication
   Format using the GKSM session metafile.
```

When a product is inconsistent with laws or widely accepted or relevant standards, we suspect a problem.



General Examples of Oracles

things that suggest "problem" or "no problem"

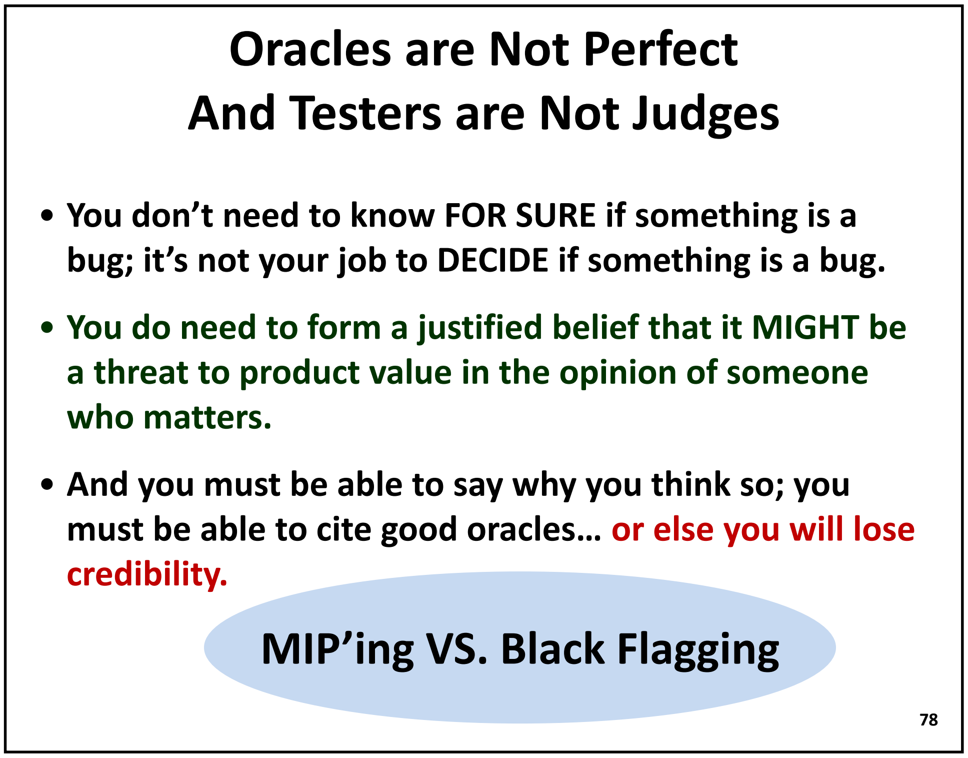
- A process or tool by which the output is checked.
- A reference document with useful information.
- A known good example output.
- A known bad example output.
- A process or tool that helps a tester identify patterns.
- A person whose opinion matters.
- An opinion held by a person who matters.
- A disagreement among people who matter.
- A feeling like confusion or annoyance.
- *A desirable consistency between related things.*

Mechanisms

People

Feelings

Principles



Oracles are Not Perfect And Testers are Not Judges

- **You don't need to know FOR SURE if something is a bug; it's not your job to DECIDE if something is a bug.**
- **You do need to form a justified belief that it MIGHT be a threat to product value in the opinion of someone who matters.**
- **And you must be able to say why you think so; you must be able to cite good oracles… or else you will lose credibility.**

MIP'ing VS. Black Flagging

78

Some Quality Criteria Are Oriented Towards Product Development

Remember Testability!

Supportability
Testability
Maintainability
Portability
Localization

Testability affords us opportunities for observing and controlling the product. Reduced testability gives bugs more time and more opportunities to hide.

79

Testability: Observability & Controllability

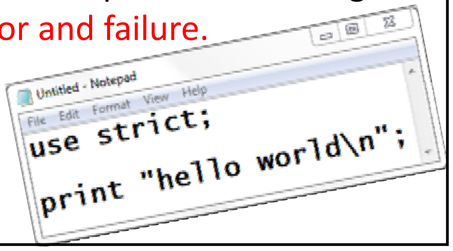
In order to test a product well, I must be able to control the execution to visit each important state that it has, see everything important, and control the variables (in the environment) that might influence it.

Imagine a clock under a glass shield. You are not allowed to come near it or poke or probe it...

Testability: Smallness and Simplicity

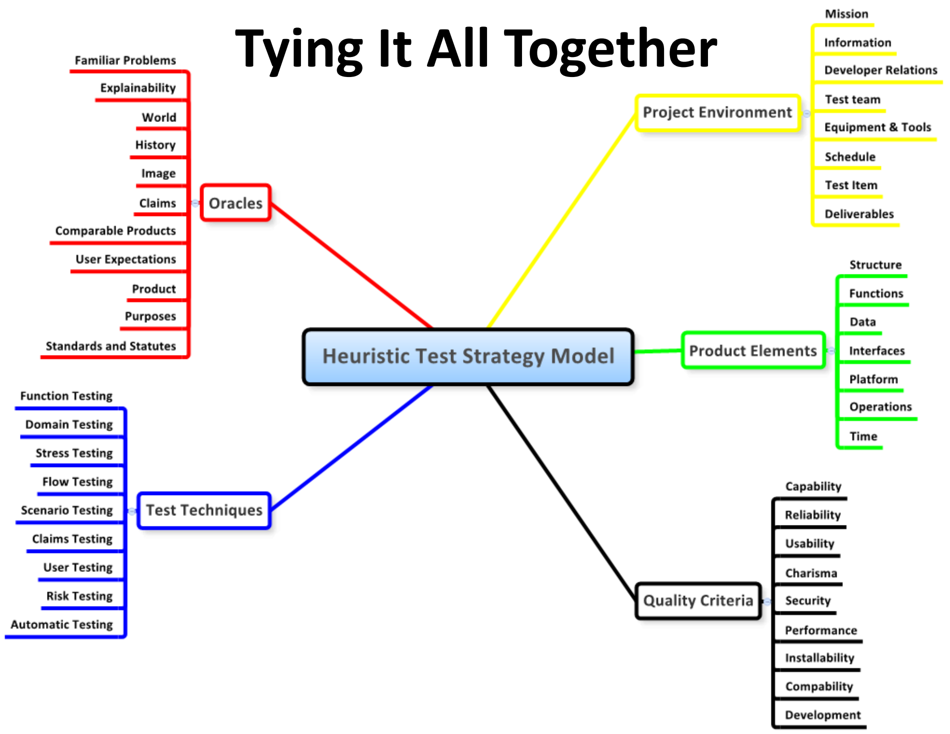
- A product is smaller, in testing terms, when there are fewer interestingly different and risky things that must be looked at.
- It is algorithmically simpler when there are fewer variations of those things to look at, and fewer conditions to control or consider when looking at them.
- This is related to your models of the product and usage as well as your theories of error and failure.

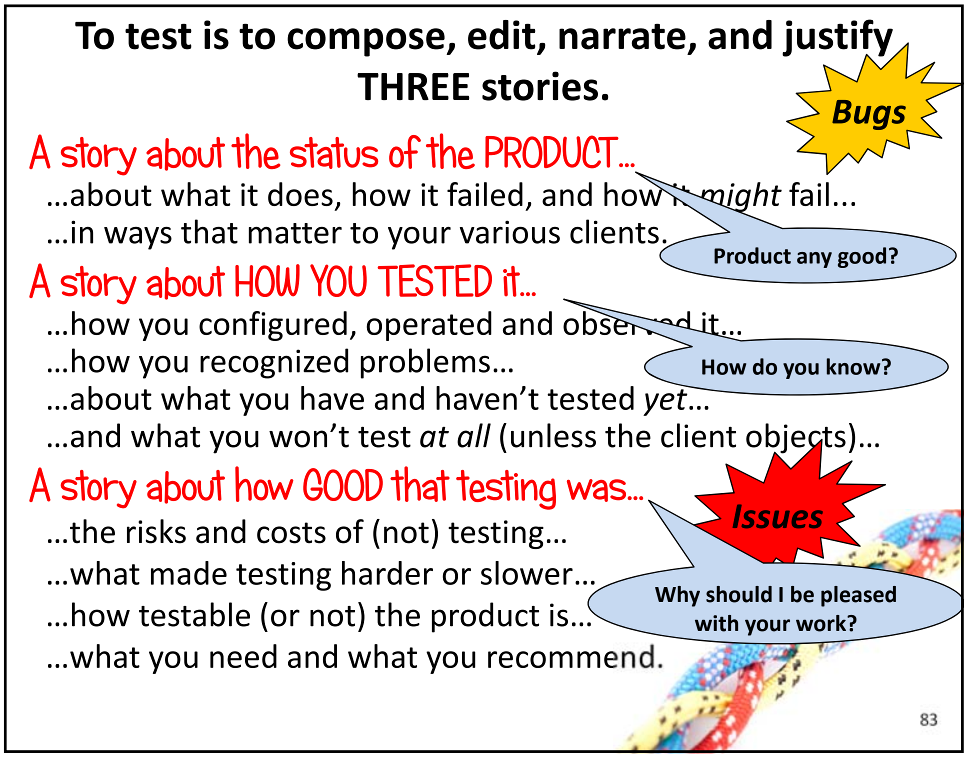
Imagine Notepad vs. Activestate Perl. Both can produce essentially infinite outputs, but one is far simpler and smaller than the other.



```
use strict;
print "hello world\n";
```

Tying It All Together





To test is to compose, edit, narrate, and justify THREE stories.

Bugs

A story about the status of the PRODUCT…

…about what it does, how it failed, and how it *might* fail...
…in ways that matter to your various clients.

Product any good?

A story about HOW YOU TESTED it…

…how you configured, operated and observed it…
…how you recognized problems…
…about what you have and haven't tested *yet*…
…and what you won't test *at all* (unless the client objects)…

How do you know?

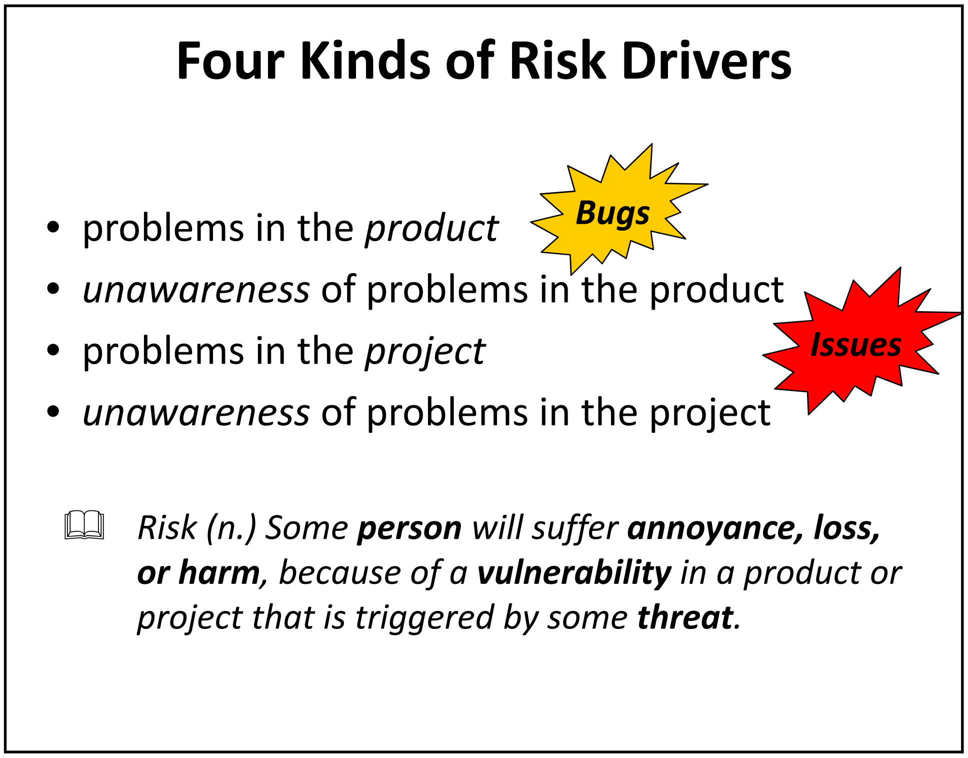
A story about how GOOD that testing was…

…the risks and costs of (not) testing…
…what made testing harder or slower…
…how testable (or not) the product is…
…what you need and what you recommend.

Issues

Why should I be pleased with your work?

83



Four Kinds of Risk Drivers

- problems in the *product*
- *unawareness* of problems in the product
- problems in the *project*
- *unawareness* of problems in the project

Bugs

Issues

📖 *Risk (n.) Some **person** will suffer **annoyance, loss, or harm**, because of a **vulnerability** in a product or project that is triggered by some **threat**.*

Is Regression Your Biggest Risk?

- Before the Agile Manifesto was declared, a group of experienced test managers reported that regression problems ran from 6-15% of discovered problems
- In Agile shops, we now (supposedly) have
  - TDD
  - unit tests
  - pairing
  - configuration management
  - build and version control
  - continuous integration
- Is regression a serious risk?
- If so, can testing (whether automated or not) fix it?
- Is regression really a symptom of problems elsewhere?
- What about all the tests you *haven't* performed yet?

Regression Problems Are *Symptoms*

- If you see a consistent pattern of regression
  - the bugs you're seeing are probably not your biggest problem
  - your biggest problem might be *a favourable environment for regression*
  - is the project simply going too fast in a complex, volatile world?

Not-So-Good Questions for Testers

- Is the product done?
- Are we ready to ship?
- Is it good enough?
- How much time do you need to test?
- How many test cases are passing and failing?
- How many tests cases have you run?
- How many bugs are in the product?

Seek more than data.
Seek information.

Better Questions for Testers

- What is the product story? What can you tell me about important problems in the product?
- What risks I should be aware of?
- What have you done to obtain the product story?
- What important testing remains to be done?
- What problems are slowing testing down or making it harder to find out what we might need to know?
- What do you need to help speed things up?
- What *specific* aspects of testing are taking time?
- How do your tests link to the mission?

What might prevent the on-time, successful completion of the project?

What Interrupts Test Coverage?

- Non-testing work, setup of environments and tools, and bug investigation and reporting *take time away* from test design and execution
- Suppose testing that appropriately covers some aspect of a feature takes two minutes; let's call that a micro-session.
- Suppose also that it takes an extra eight minutes to investigate and report a bug
  - these are highly arbitrary and artificial assumptions—that is, they're *wrong*, but let's run the thought experiment anyway
- In a 90-minute session, we can run 45 micro-sessions *—as long as we don't find any bugs*

How Do We Spend Time?

(assume we're fabulous testers and would find everything our clients deem a bug)

Module	Bug reporting/investigation (time spent on tests that find bugs)	Test design and execution (time spent on tests that find no bugs)	Number of tests
A (good)	0 minutes (no bugs found)	90 minutes (45 tests)	45
B (okay)	10 minutes (1 bug, 1 test)	80 minutes (40 tests)	41
C (bad)	80 minutes (8 bugs, 8 tests)	10 minutes (5 tests)	13

Investigating and reporting bugs means….

SLOWER TESTING **or…**

REDUCED COVERAGE **…or both.**

- For Module A, our *coverage* is great—but if our clients assess us on the number of bugs we're finding, we look bad.
- For Module C, we look good because we're finding and reporting lots of *bugs*—but our *coverage* is suffering severely.
- Note that we haven't included setup time here, either.
- A system that rewards us or increases confidence based on the number of bugs we find might mislead us into believing that our product is well tested.

What Happens The Next Day?

(assume 6 minutes per bug fix verification)

	Fix verifications	Bug reporting and investigation today	Test design and execution today	New tests today	Total over two days
A	0 min	0 min (no new bugs)	90 min (45 tests)	45	90
B	6 min	10 min (1 new bug)	74 min (37 tests)	38	79
C	48 min	40 min (4 new bugs)	2 min (1 test)	5	18

Finding bugs today means….

VERIFYING FIXES LATER

…which means….

EVEN SLOWER TESTING **or…**

EVEN LESS COVERAGE **…or both.**

•…and note the optimistic assumption that all of our fixed verifications worked, and that we found no new bugs while running them. Has this ever happened for you?

91

13 Commitments for Testers to Make to Clients

1. I provide a service. You are an important client of that service. I am not satisfied unless you are satisfied.
2. I am not the gatekeeper of quality. I don't "own" quality. Shipping a good product is a goal shared by all of us.
3. I will test anything as soon as someone delivers it to me. I know that you need my test results quickly (especially for fixes and new features).
4. I will strive to test in a way that allows you to be fully productive. I will not be a bottleneck.
5. I'll make every reasonable effort to test, even if I have only partial information about the product.
6. I will learn the product quickly, and make use of that knowledge to test more cleverly.
7. I will test important things first, and try to find important problems. *(I will also report things you might consider unimportant, just in case they turn out to be important after all, but I will spend less time on those.)*

13 Commitments for Testers to Make to Clients

8. I will strive to test in the interests of everyone whose opinions matter, including you, so that you can make better decisions about the product.
9. I will write clear, concise, thoughtful, and respectful problem reports. *(I may make suggestions about design, but I will never presume to be the designer.)*
10. I will let you know how I'm testing, and invite your comments. And I will confer with you about little things you can do to make the product much easier to test.
11. I invite your special requests, such as if you need me to spot check something for you, help you document something, or run a special kind of test.
12. I will not carelessly waste your time. Or if I do, I will learn from that mistake.
13. I will not FAKE a test project.

The Themes of Rapid Testing

- Put the **tester's mind** at the center of testing.
- Learn to **deal with complexity** and ambiguity.
- Learn to **tell a compelling testing story.**
- Develop **testing skills** through practice, not just talk.
- **Use heuristics** to guide and structure your process.
- **Be a service** to the project community, not an obstacle.
- **Consider cost vs. value** in all your testing activity.
- **Diversify** your team and your tactics.
- Dynamically **manage the focus** of your work.
- Your **context should drive your choices**, both of which evolve over time.

Testers light the way.

This is our role.

We see things for what they are.
We make informed decisions about quality possible,
because we think critically about software.

How to Get What You Really Want from Testing

Michael Bolton
DevelopSense
http://www.developsense.com
@michaelbolton
michael@developsense.com