A Very Rapid Introduction to Rapid Software Testing

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

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

1

Updates

- This presentation is ALWAYS under construction
- Updated slides at http://www.developsense.com/past.html
- All material comes with lifetime free technical support

Flight Itinerary

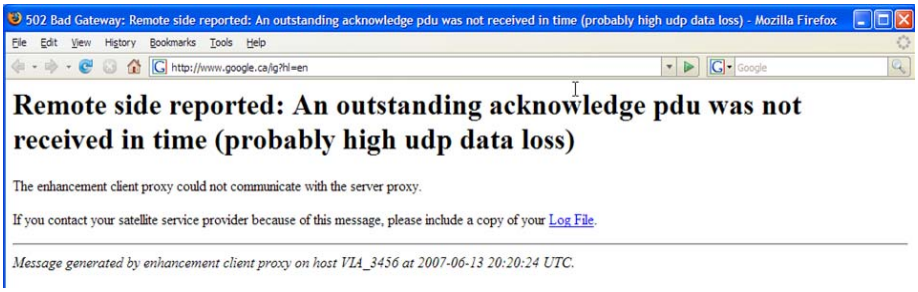
Flight	From	To	Stops	Aircraft	Fa
AC597	**Toronto, Pearson Int'l (YYZ)** Sat 16-Oct 2010 13:10 - Terminal 1	**Las Vegas, Mccarran Int'l (LAS)** Sat 16-Oct 2010 14:54 - Terminal 2	1	319	Ta
	AC597: This flight includes a stop in null.				
AC5233*	**Las Vegas, Mccarran Int'l (LAS)** Tue 19-Oct 2010 19:15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

aircanada.com

ght Itinerary

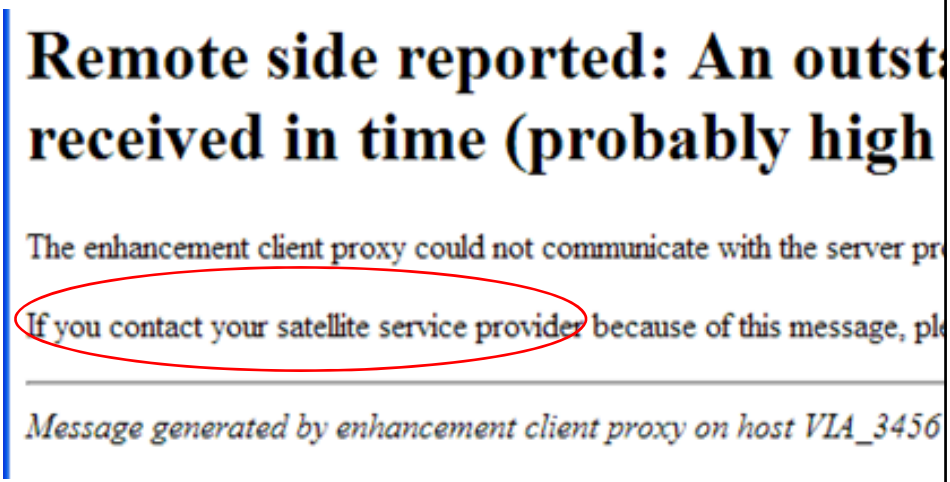
ght	From	To
597	**Toronto, Pearson Int'l (YYZ)** Sat 16-Oct 2010 13:10 - Terminal 1	**Las Vegas, Mccarran I** Sat 16-Oct 2010 14:54 - Terminal 2
	AC597: This flight includes a stop in null.	
5233*	**Las Vegas, Mccarran Int'l (LAS)** Tue 19-Oct 2010 19:15 - Terminal 1	**San Francisco, San Fr** **Int'l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

aircanada.com



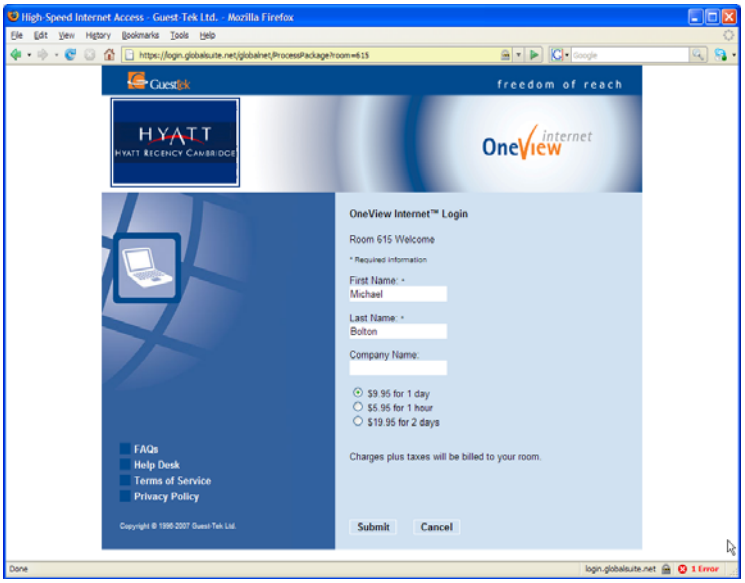
Error Message
Via Rail, between Montreal and Toronto, 2007

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



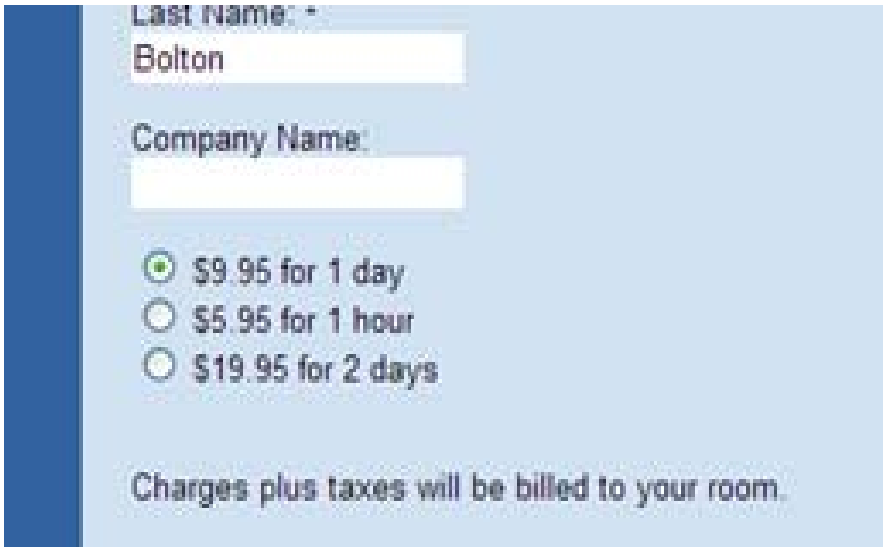
Error Message
Via Rail, between Montreal and Toronto, 2007

With rooms at $195 a night, they CHARGE you?



Hyatt Regency, Cambridge, USA, 2008

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



Hyatt Regency, Cambridge, USA, 2008

On the other hand, why not charge it to someone else's room?

Hyatt Regency, Cambridge, USA, 2008

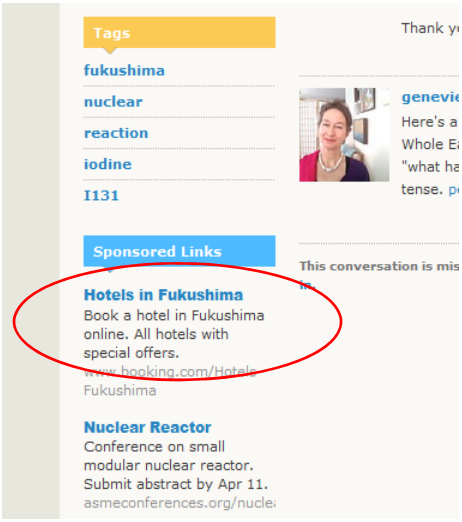
Why you shouldn't let an unsupervised algorithm choose your sponsored links (1).

Vimeo's Web Page
Spring 2010

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

Vimeo's Web Page
Spring 2010

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

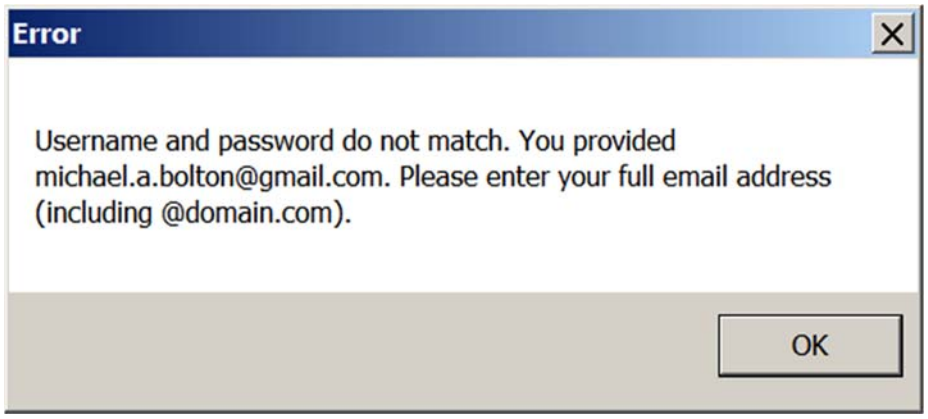


Vimeo's Web Page
Spring 2010

At least it wasn't 1 or 2.

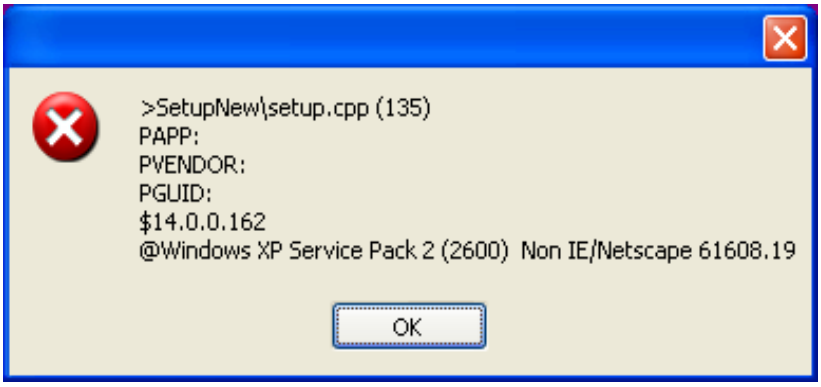
Google Chrome

Should that be michael.a.bolton@gmail.com@domain.com?

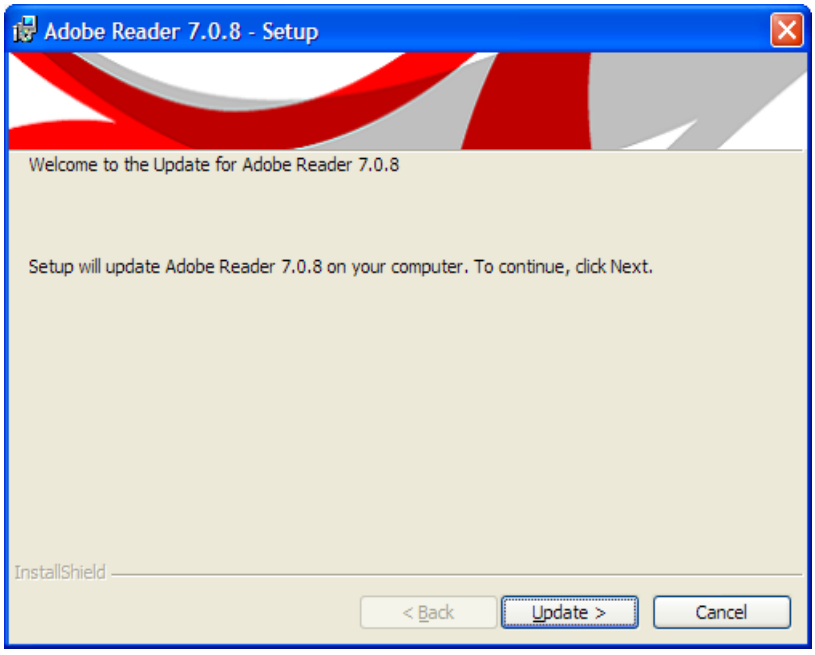


Google Calendar Update

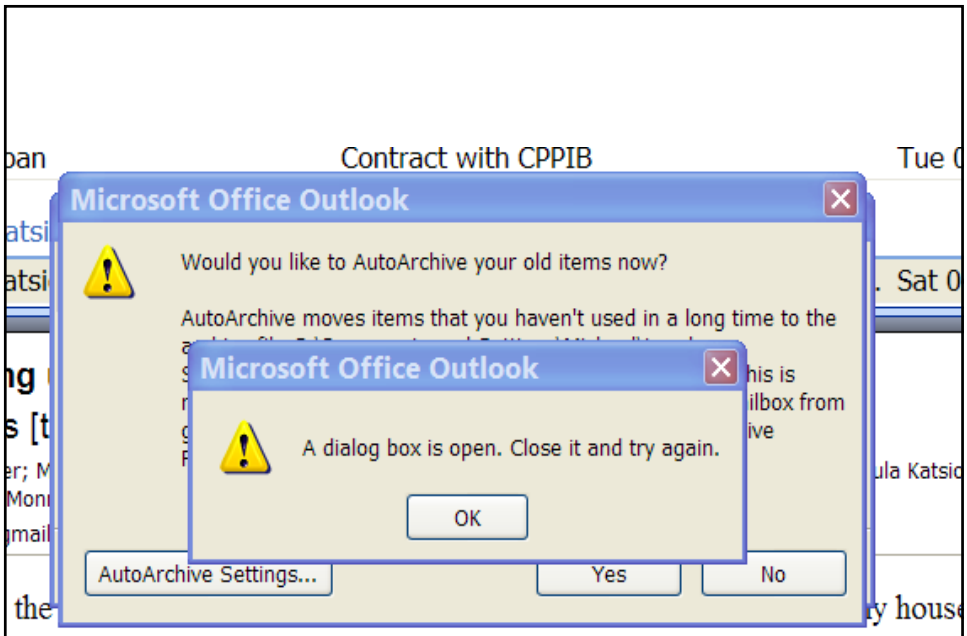
OK, fine.



Don't Know Who This Was

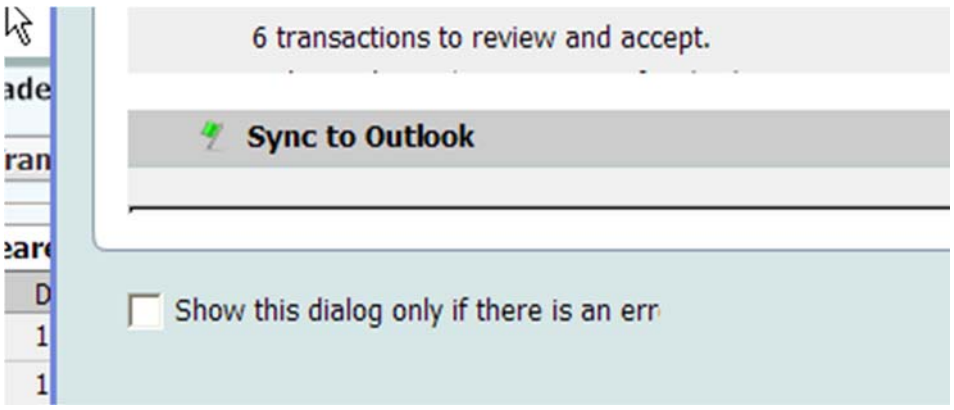


Adobe Acrobat



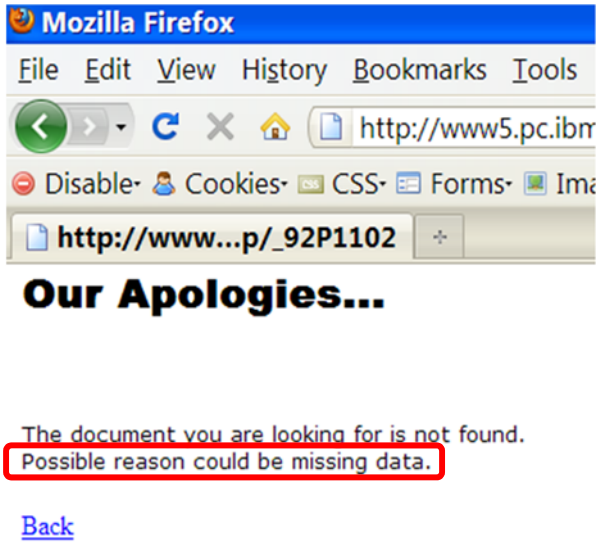
Microsoft Outlook

I think there IS an err.

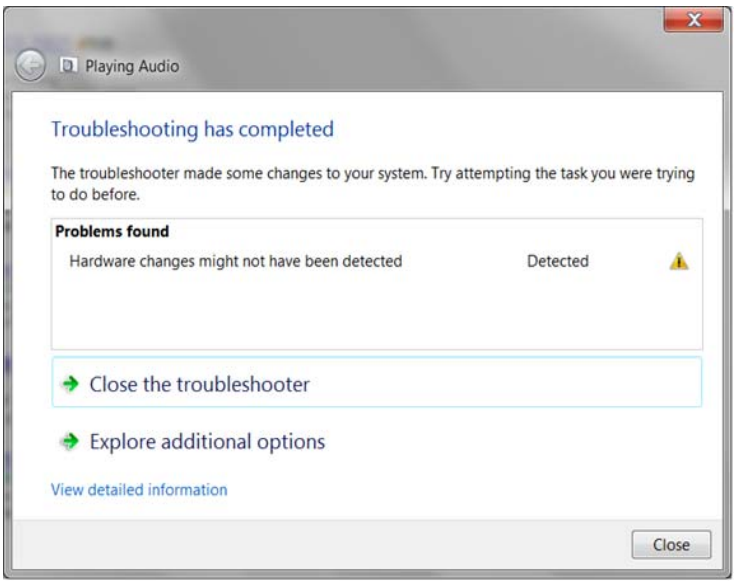


Intuit Quicken

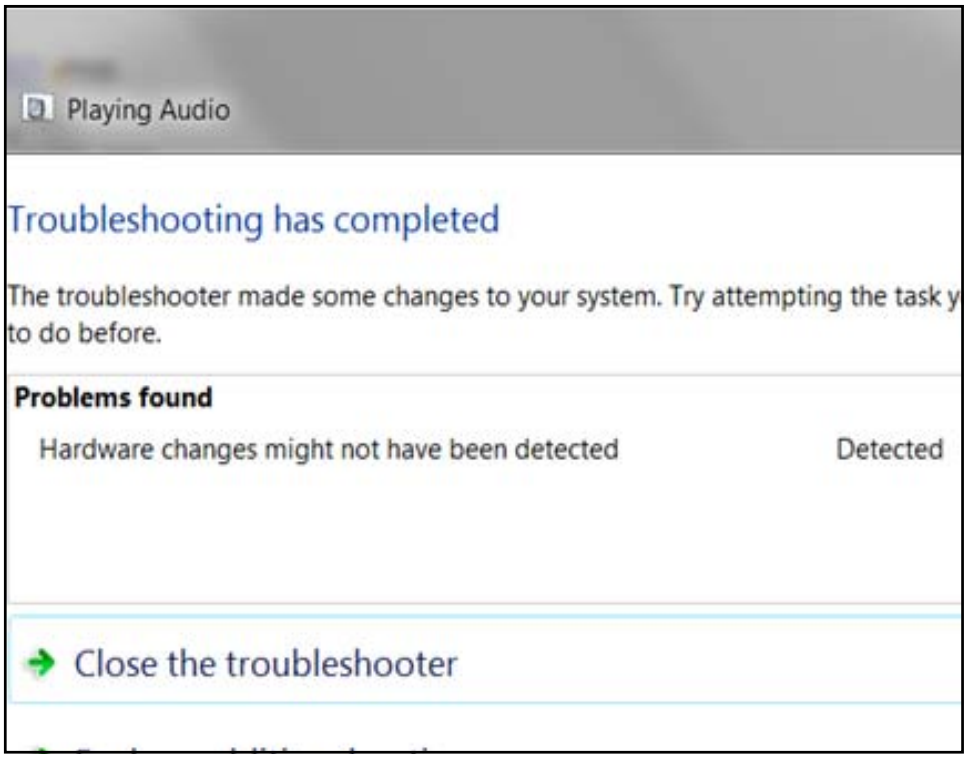
What was your first clue?



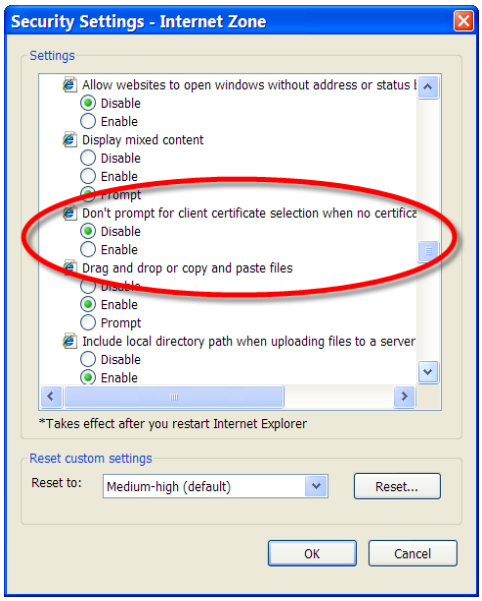
ibm.com



Windows

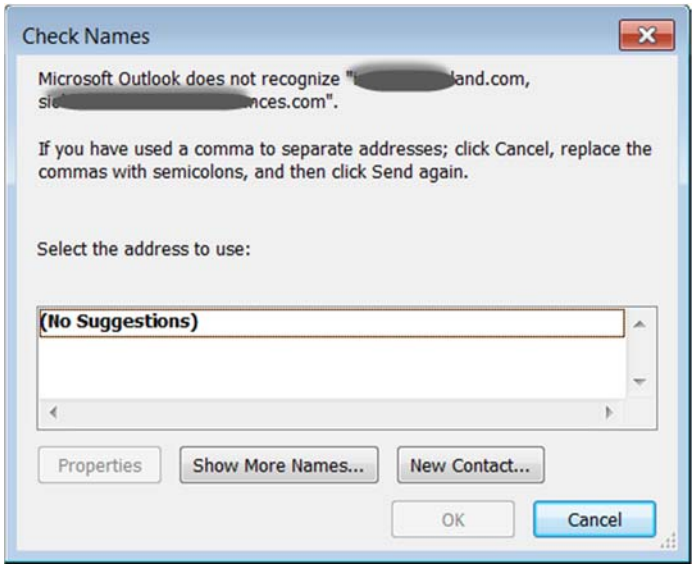


Go ahead; make sense of this.

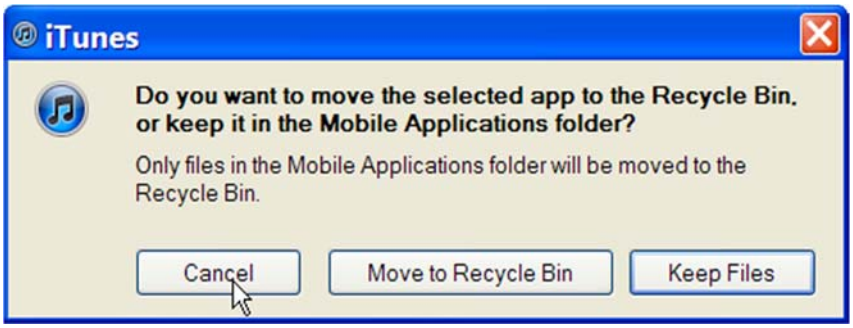


Windows

You're the computer. You do it.



23

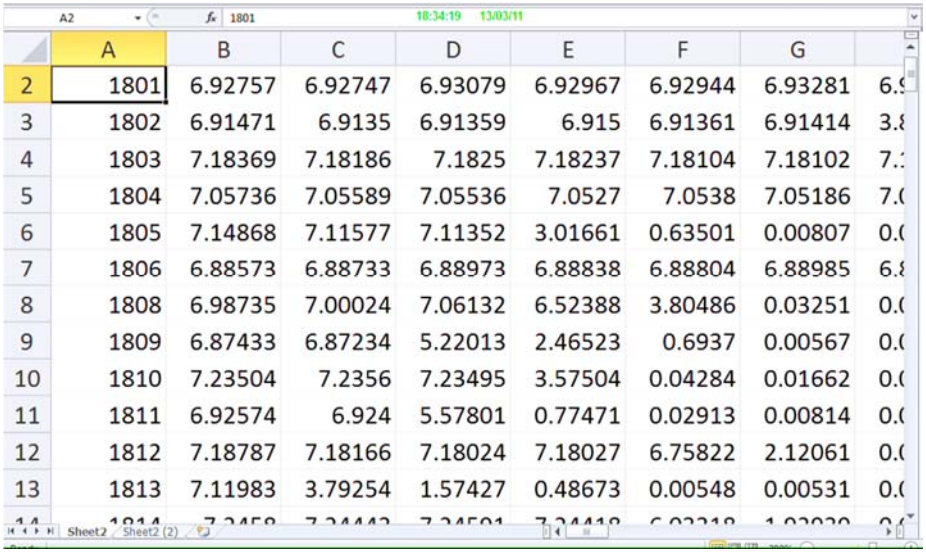


iTunes

This One's From Just This Morning

25

Deeper Rapid Testing



	A	B	C	D	E	F	G	
2	1801	6.92757	6.92747	6.93079	6.92967	6.92944	6.93281	6.9
3	1802	6.91471	6.9135	6.91359	6.915	6.91361	6.91414	3.8
4	1803	7.18369	7.18186	7.1825	7.18237	7.18104	7.18102	7.1
5	1804	7.05736	7.05589	7.05536	7.0527	7.0538	7.05186	7.0
6	1805	7.14868	7.11577	7.11352	3.01661	0.63501	0.00807	0.0
7	1806	6.88573	6.88733	6.88973	6.88838	6.88804	6.88985	6.8
8	1808	6.98735	7.00024	7.06132	6.52388	3.80486	0.03251	0.0
9	1809	6.87433	6.87234	5.22013	2.46523	0.6937	0.00567	0.0
10	1810	7.23504	7.2356	7.23495	3.57504	0.04284	0.01662	0.0
11	1811	6.92574	6.924	5.57801	0.77471	0.02913	0.00814	0.0
12	1812	7.18787	7.18166	7.18024	7.18027	6.75822	2.12061	0.0
13	1813	7.11983	3.79254	1.57427	0.48673	0.00548	0.00531	0.0

26

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.

Call this "Checking" not Testing

operating a product to check specific facts about it…

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

Evaluate — Apply algorithmic decision rules to those observations.

Report — 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

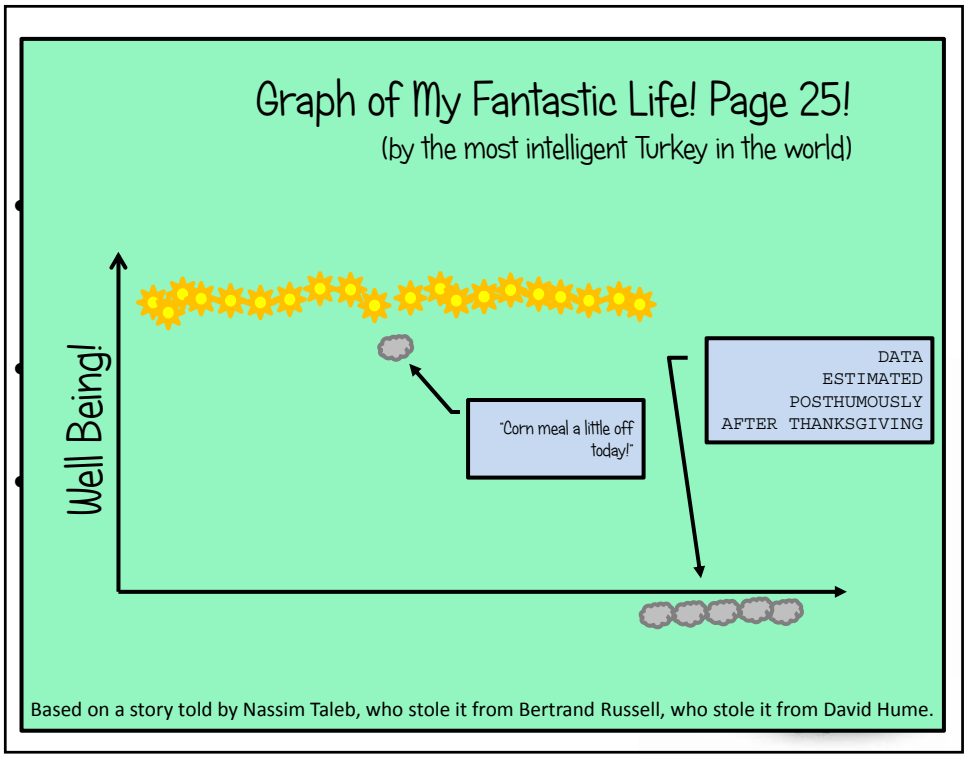
by a machine
that *can't* think
(but that is quick and precise)

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

Testing Is More Than *Checking*

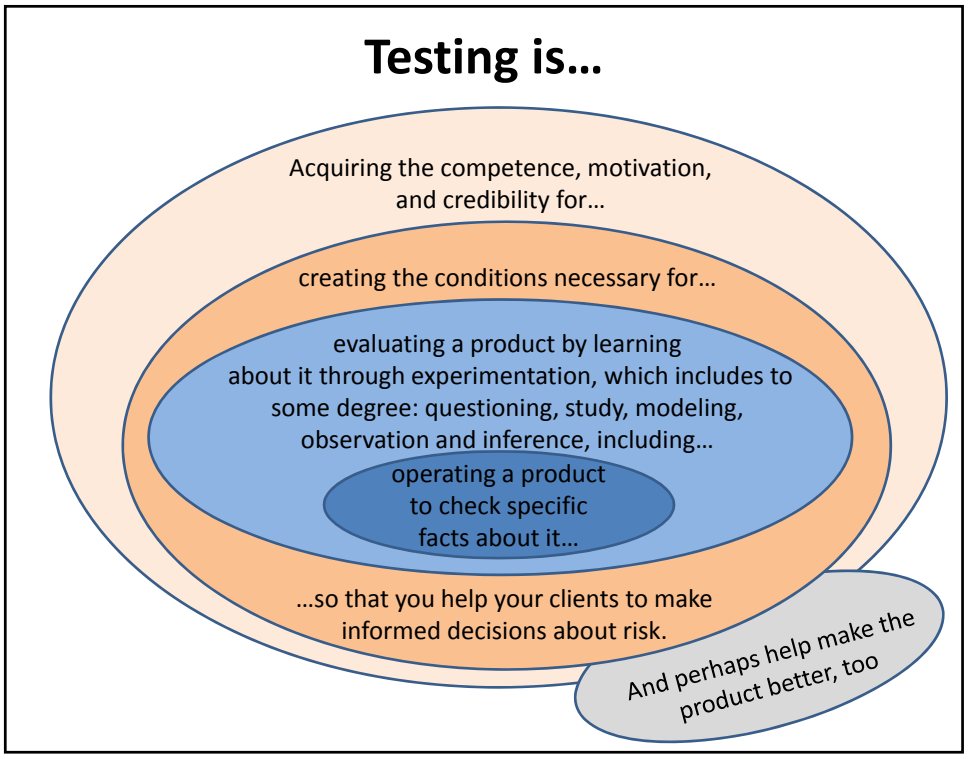
- *Checking* is a process of confirming and verifying existing beliefs
  - Checking can (and we argue, largely should) be done mechanically
  - It is a *non-sapient* process

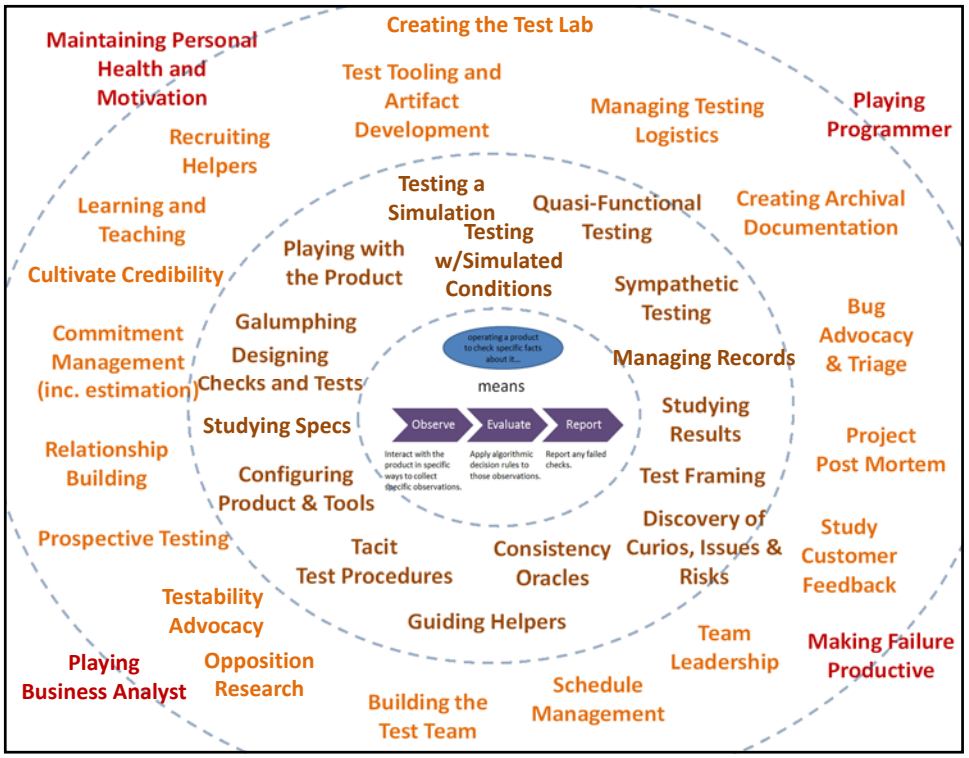
See http://www.developsense.com/2009/08/testing-vs-checking.html



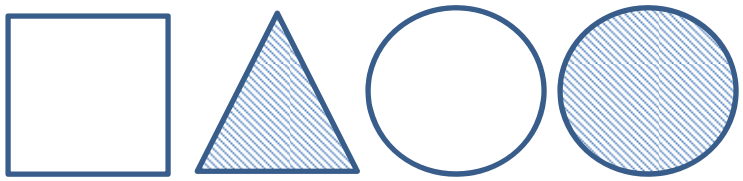
Don't Be A Turkey

- 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" (CHECKS) CANNOT PROVE THAT SOFTWARE IS GOOD.**





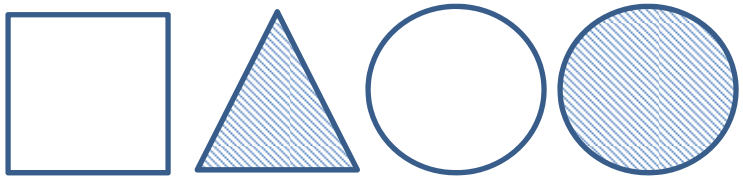
Wait, let's try something simple…



Can we agree? Can we share common ground?

"There are four geometric figures on this slide."
"There is one **square** among those figures."
"The square is shaded in **blue**."

Wait, let's try something simple…



This is Shallow Agreement!

What is the Point?

Any communication among humans…

(especially technical communication about complex systems)

…involves making reasonable, socially-situated assumptions about things not spoken...

(or else people will see you as a child, a robot, or a very unpleasant person)

…and that is a risk to be managed.

What is a "reasonable" assumption?

Tacit and Explicit Knowledge

EXPLICIT means it can be represented completely in the form of a string of bits: words, pictures, even actions can be explicit. (software is explicit)

TACIT means it is not manifested in a form that can be equated to a string of bits: it is unspoken, unwritten, unpictured.

- **Relational Tacit Knowledge** is tacit by convenience.
- **Somatic Tacit Knowledge** is tacit in your body.
- **Social Tacit Knowledge** is tacit in a community.

(see Collins, *Tacit and Explicit Knowledge*)

Examples of the Need for Tacit Knowledge

- You consider how users will **interact with and adjust** to a product.
- You consider what a specification and product was **intended to say or do**, not just what they literally say and do.
- You notice what the specification and product **strangely omits**.
- You focus on business risks even when **no one tells you** what they are.
- You notice changes in the product's behavior over time—none of which are failures in and of themselves—**you form conjectures** of why that happens, and you connect those conjectures with an evolving understanding of **plausible error and failure patterns**.
- You use tools, you notice when those tools **misbehave**, and you make **adjustments or workarounds** to get the job done.
- You report your test results based partly on **how you think your clients will react**. You **anticipate** their questions based on your insight about **how they will understand you.**

Exercise: Calculator Test

"You are carrying a calculator.

You **drop** it!

Perhaps it is **damaged**!

What might you do to test it?"

When did I drop it? Was I in the middle of a calculation? If so part of my testing might be to visually inspect the status of the display to determine whether the calculator appears to still be in the state it was at the time I dropped it. If so, I might continue the calculation from that point, unless I believe that the drop probably damaged the calculator.

Did I drop it on a hard surface with a force that makes me suspect internal damage? If so then I would expect possible hairline fractures. I imagine that would lead to intermittent or persistent short circuits or broken circuits. I also suspect damage to moving parts, battery or solar cell connections, or screen.

Did I drop it into a destructive chemical environment? If so, I might worry more about the progressive decay of the components.

Did I drop it into a dangerous biological or radiological environment? If so, the functions of the calculator maybe less concern than contaminants. I may have to test it with a Geiger counter.

Was the calculator connected to anything else whereby the connection (data cable or AC/cable or duct tape that fastened it to a Faberge egg) could have been damaged, or could have damaged the thing it was connected to?

Did I detect anything while it was dropping that leads me to suspect any damage in particular (e.g. an electrical flash, or maybe a loud popping sound)?

Am I aware of a history of "drop" related problems with this calculator? Have I ever dropped it before?

Is the calculator ruggedized? Is it designed to be dropped in this way?

What is my relationship to this calculator? Is it mine or someone else's? Maybe I'm just borrowing it.

What is the value of this calculator. I assume that this is not a precious artifact from a museum. The exercise as presented appears to be about a calculator as calculating machine, rather than as a precious Minoan urn that happens to have calculator functions built into it.

What am I using the calculator for? If it's a component of a space craft, it may be irresponsible to use it unless I comprehensively re-qualify it. Or maybe it's a throwaway item used only for simple calculations

What is the Point?

You have MANY expectations about how a calculator should work or might work…

(you acquire them automatically)

…but you aren't aware of many of them...

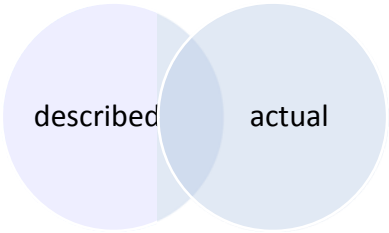
(they are encoded or generated in your mind, but not in words or pictures)

…and that means no explicit test procedure can ever duplicate the value of tacit knowledge of a skilled human tester.

We need a process that respects tacit knowledge.

This is what people think testers do

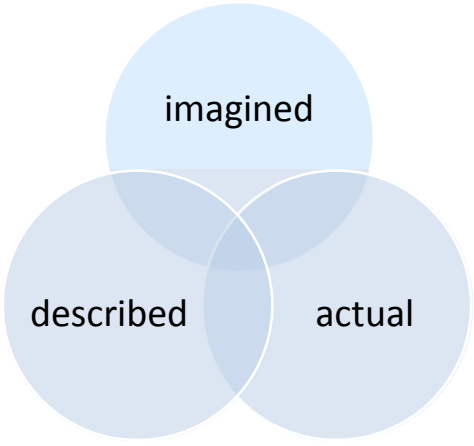
"Compare the product to its specification"



This is more like what testers really do

"Compare the idea of the product to a description of it"

"Compare the idea of the product to the actual product"



"Compare the actual product to a description of it"

This is what testers find…

The designer **INTENDS** the product to be Firefox compatible, but never says so, and it actually is not.

The designer **INTENDS** the product to be Firefox compatible, **SAYS SO IN THE SPEC**, but it actually is not.

The designer **INTENDS** the product to be Firefox compatible, **MAKES IT FIREFOX COMPATIBLE**, but forgets to say so in the spec.

The designer **INTENDS** the product to be Firefox compatible, **SAYS SO**, and **IT IS**.

The designer assumes the product is not Firefox compatible, and it actually is not, but the **ONLINE HELP SAYS IT IS**.

The designer assumes the product is not Firefox compatible, but it **ACTUALLY IS**, and the **ONLINE HELP SAYS IT IS**.

The designer assumes the product is not Firefox compatible, and no one claims that it is, but it **ACTUALLY IS**.

We Are Sensory Instruments For Our Clients

Premises of Rapid Testing

1. Software projects and products are relationships between people.
2. Each project occurs under conditions of uncertainty and time pressure.
3. Despite our best hopes and intentions, some degree of inexperience, carelessness, and incompetence is normal.
4. A test is an activity; it is performance, not artifacts.

Premises of Rapid Testing

5. Testing's purpose is to discover the status of the product and any threats to its value, so that our clients can make informed decisions about it.
6. We commit to performing credible, cost-effective testing, and we will inform our clients of anything that threatens that commitment.
7. We will not knowingly or negligently mislead our clients and colleagues or ourselves.
8. Testers accept responsibility for the quality of their work, although they cannot control the quality of the product.

One Big Problem in Testing

Formality Bloat

- Much of the time, your testing doesn't need to be very formal*
- Even when your testing *does* need to be formal, you'll need to do substantial amounts of informal testing in order figure out how to do *excellent* formal testing.
  - Who says? The FDA. See http://www.satisfice.com/blog/archives/602
- Even in a highly regulated environment, you do *formal* testing primarily for the auditors. You do informal testing to make sure you don't lose money, blow things up, or kill people.

* Formal testing means testing that must be done *to verify a specific fact*, or that must be done *in a specific way*.

What It Means To Test Rapidly

- Since testing is about finding a potentially infinite number of problems in an infinite space in a finite amount of time, testers must *quickly and expertly*
  - understand our mission and obstacles to fulfilling it
  - know how to recognize problems quickly
  - produce diversified models of the product and the test space to know where to look for problems
  - prefer inexpensive, lightweight, effective tools
  - reduce dependence on expensive, time-consuming artifacts, while getting value from the ones we've got
  - do nothing that wastes time or effort
  - tell a credible story about all that

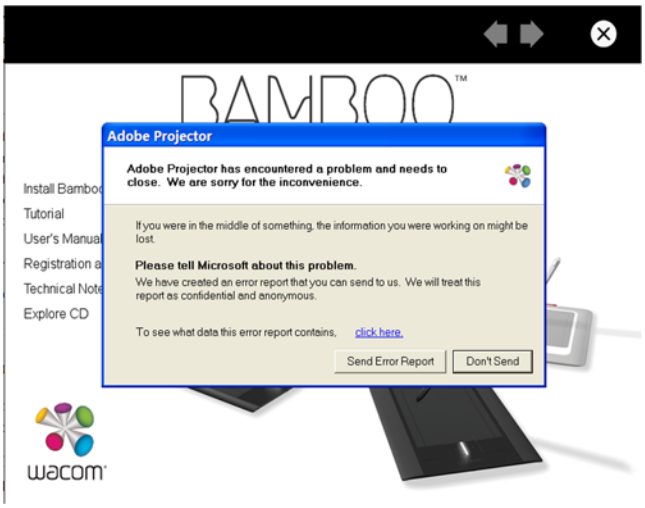
54

How Do We Recognize Problems?

An oracle is...

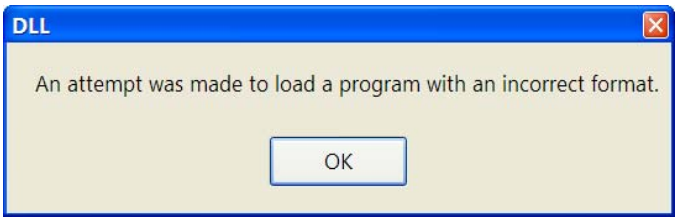
a way to recognize a problem.

Familiar Problems



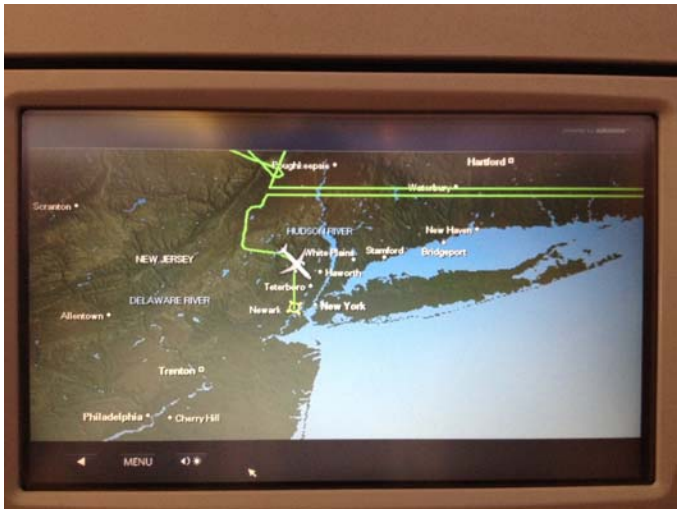
If a product is consistent with problems we've seen before, we suspect that there might be a problem.

Explainability



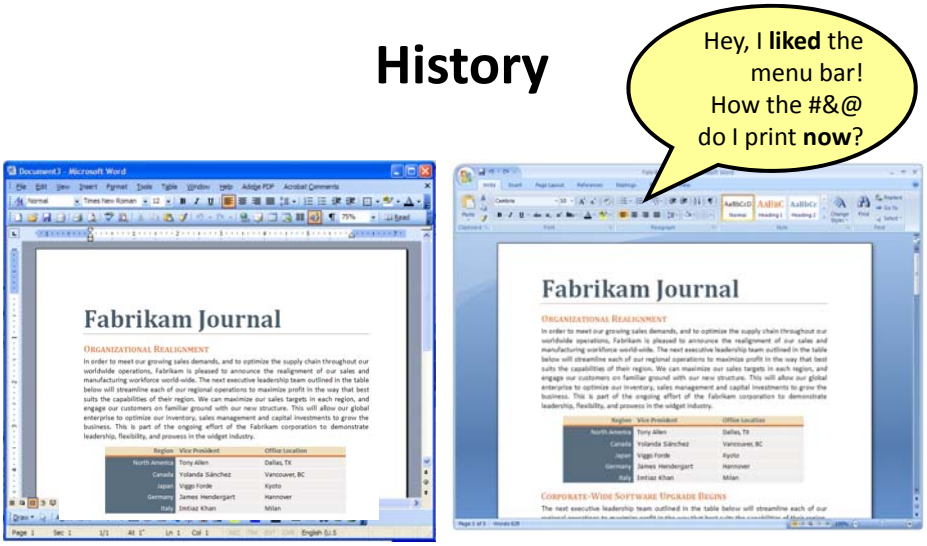
If a product is inconsistent with our ability to explain it, we suspect that there might be a problem.

World



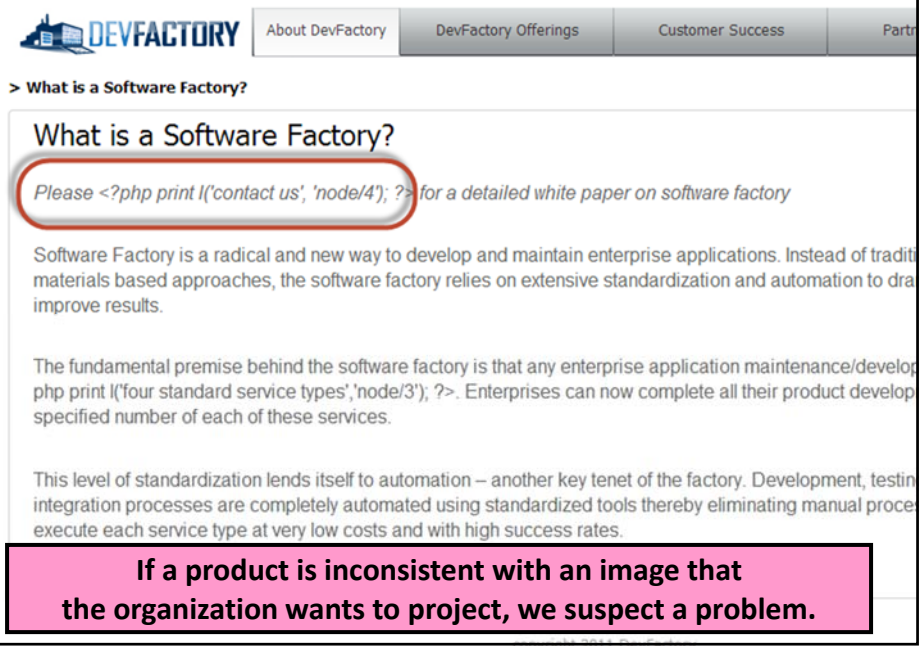
If a product is inconsistent with the way the world works, we suspect that there might be a problem.

History



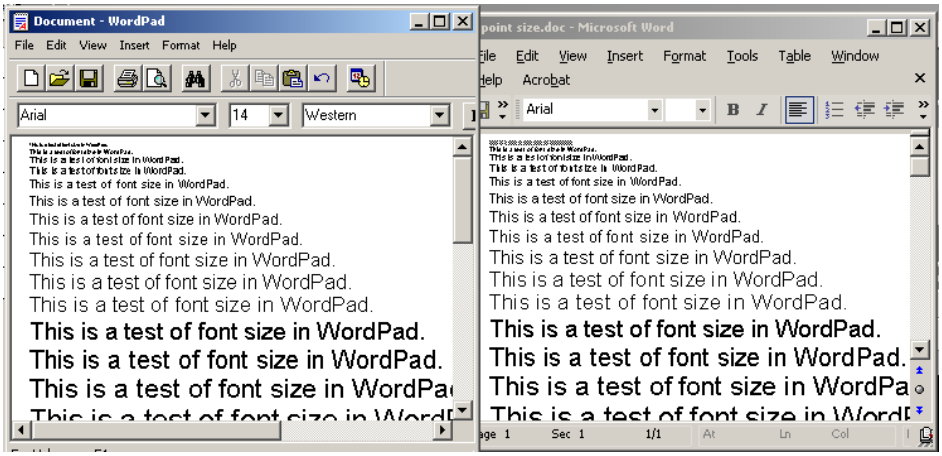
If a product is inconsistent with previous versions of itself, we suspect that there might be a problem.

Image



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

Comparable Products



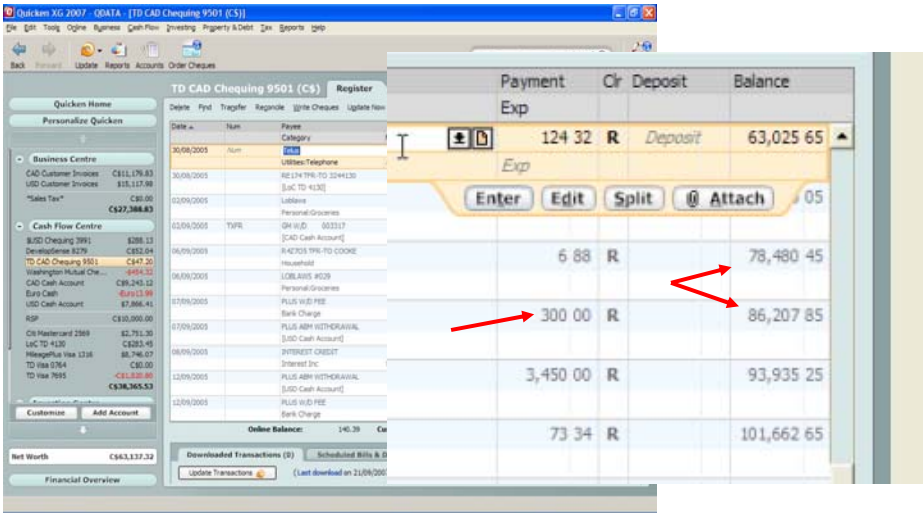
WordPad Word

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

Claims

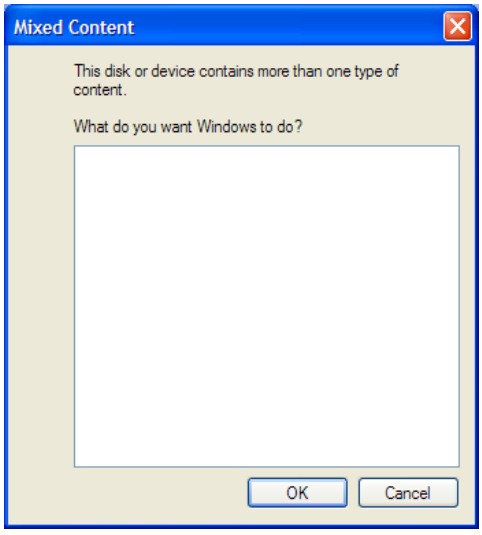
When a product is inconsistent with claims that important people make about it, we suspect a problem.

User Expectations



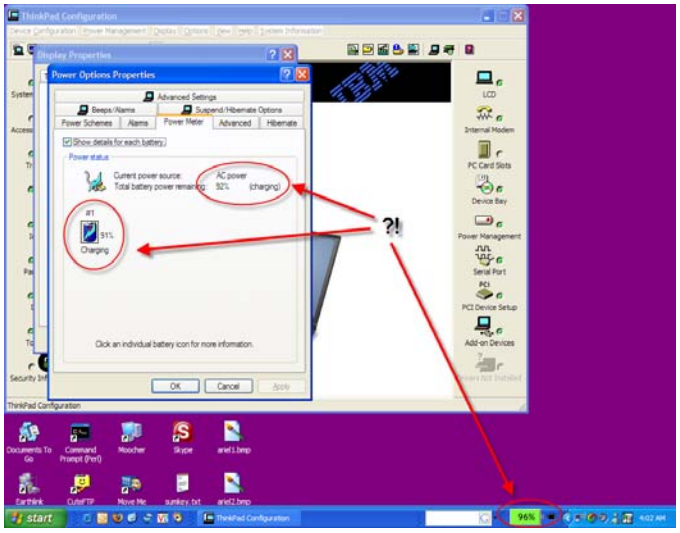
When a product is inconsistent with expectations that a reasonable user might have, we suspect a problem.

Purpose



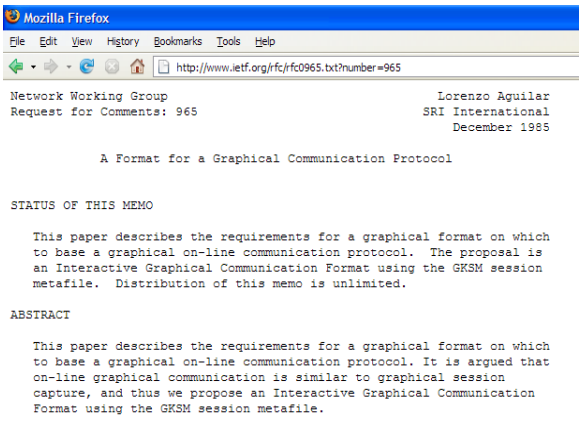
When a product is inconsistent with its designers' explicit or implicit purposes, we suspect a problem.

Product



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

Statutes and Standards



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

Where Do We Look For Problems?

Coverage is...

how much of the
product has been tested.

Where Do We Look For Problems?

____ coverage is...

how much of the product
has been tested
with respect to a model of _____.

Models

- **A model is an idea, activity, or object…**

  such as an *idea in your mind*, a *diagram*, a *list of words*, a *spreadsheet*, a *person*, a *toy*, an *equation*, a *demonstration*, or a *program*

- **…that heuristically represents (literally, re-presents) another idea, activity, or object…**

  such as something complex that you need to work with or study

- **…whereby understanding something about the model may help you to understand or manipulate the thing that it represents.**

  - A *map* is a model that helps to navigate across a terrain.
  - *2+2=4* is a model for adding two apples to a basket that already has two apples.
  - *Atmospheric models* help predict where hurricanes will go.
  - A *fashion model* helps understand how clothing would look on actual humans.
  - *Your beliefs about what you test are a model of what you test.*

There are as many kinds of test coverage as there are ways to model the system.

Structure
Business Risk
Time
Functions
Platform
Technical Risk
Operations
Interfaces
Data

And each could be covered…

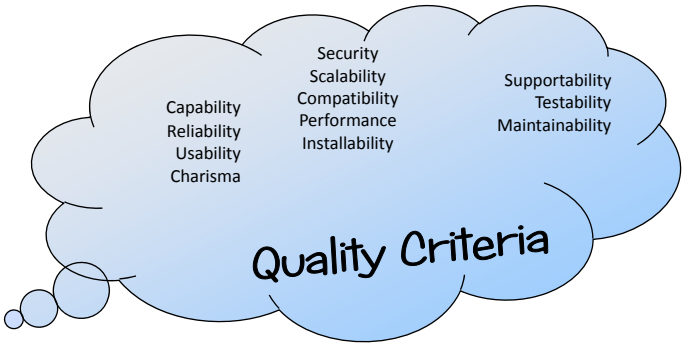
Intentionally OR Incidentally

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

SFDIPOT -- San Francisco Depot

Product Elements

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



Contrasts with Traditional Approaches

Traditional

- Correct
- Precise
- Formal
- Explicit
- Confirmation
- Demonstration
- Experiment (in the lab)
- Test cases (nouns)
- Artifacts
- Code

Rapid Software Testing

- Good enough
- Accuracy
- Informal
- Tacit
- Discovery
- Experiment
- Experience (in the world)
- Testing (verbs)
- Activities
- Social Systems

Contrasts with Traditional Approaches

Traditional	Rapid Software Testing
• Explicit requirements	• Value and threats to value
• Mediated skill	• Direct skill
• Fungible tester	• Responsible tester
• Attributes	• Relationships
• Rational	• Emotional
• Numbers	• Stories
• Certification	• Personal mastery
• Planning	• Preparation
• Standardization	• Diversity
• Following procedures	• Applying heuristics

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.

74