How It Works

Dice Game
Slack Channel for RST Alumni

http://www.developsense.com/blog/2017/10/rst-slack-channel/
Acknowledgements

- James Bach, Michael Bolton, Huib Schoots, Paul Holland, and Griffin Jones co-create and teach this class.
- Some of this material was developed in collaboration with Cem Kaner.
- Jon Bach has been a long-term collaborator in developing ET management and training methods.
- Lenore Bach (wife of James) maintains the emotional and home environment that motivated James and enabled him to create this.
- Many of the ideas in this presentation were also inspired by or augmented by other colleagues including Doug Hoffman, Bret Pettichord, Brian Marick, Dave Gelperin, Elisabeth Hendrickson, Jerry Weinberg, Noel Nyman, and Mary Alton.
- Some of our exercises were introduced to us by Payson Hall, Ross Collard, James Lyndsay, Dave Smith, Earl Everett, Brian Marick, and Joe McMahon.
- Many ideas were improved by students who took earlier versions of the class going back to 1995.

Although the value of the class is greater because of the people who helped us, the sole responsibility for the content of this class belongs to the authors, James Bach and Michael Bolton. This class is essentially a long editorial opinion about testing. We have no authority to say what is right. There are no authorities in our field. We don’t peddle best practices. We don’t believe in them.

What we will do is share our experiences and challenge your mind. We hope you will challenge us back. You’ll find the class more rewarding, we think, if you assume that our purpose is to make you stronger, smarter and more confident as a tester. Our purpose is NOT to have you listen to what the instructor says and believe it, but to listen and then think for yourself.

All good testers think for themselves. In a way, that’s what testing is about. We look at things differently than everyone else so that we can find problems no one else will find.
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.

Rapid testing is a mind set—a particular way of looking at the world; and a skill set—a particular set of things that we practice and get better at.

We’re advocates of the context-driven approach to software testing.

We’ve applied these approaches in financial institutions, in court cases, in testing of medical devices, in commercial shrink-wrapped software, to games... The methodology is designed to adapt to any kind of testing context.

Rapid testing involves considerations of skill, personal integrity, improved focus on the underlying need for testing tasks, improved appreciation for the stakeholders of testing tasks, and knowledge of the possible techniques and tools that could be brought to bear to improve efficiency.
What is Rapid Testing?

Rapid testing is a methodology focused on how people learn and self-organize under pressure:

1. Learning to cope with complexity and uncertainty.
2. Learning how to think critically about product risk.
3. Learning how to analyze the product and design tests.
4. Learning what the bugs are in the product.
5. Learning how other people can help you test.
6. Learning how to use tools to magnify testing.
7. Learning to construct and share your testing story.

And how to deliver the fruits of that learning both effectively and ethically.

Coping with complexity and uncertainty is a skill that opens the way to everything else. Living productively with uncertainty becomes part of the identity of every great tester or critical thinker.
Primary Goal of this Class

To teach you how to test a product when you have to test it *right now*, under *conditions of uncertainty*, in a way that stands up to *scrutiny*.

Without scrutiny, this would be easy to do. You could test badly, but no one would ever know or care. What makes testing hard is that someone, perhaps your employer, perhaps the users, or perhaps you yourself, will be judging your work. Maybe this scrutiny will be indirect, such as someone unhappy with the product and cursing your company for building it so poorly. Or perhaps it will be the most intimate form of scrutiny— which is your feeling of pride or disappointment in your own work.

If you want your testing to pass muster, you need to work at it.
Secondary Goal of this Class

To help you practice *thinking* like an expert tester.

Do that well, and you’ll be able to handle any testing situation.

You will get some practice in thinking like an expert tester. That will mean some success and some failures. For the successes, we want you to feel good—strong and powerful; that’s a feeling that an expert tester gets. Alas, from time to time, expert testers also get the opportunity to feel embarrassed by failure. There are many traps in the exercises. We want to give you the experience of being trapped, of recognizing the traps, and of working your way out of them. This class is intended to provide you with a safe place to do that, so you can recognize similar situations and respond appropriately when you’re on the job.

We also expect you to challenge our expertise. That’s part of developing your own expertise, but it’s also fundamental to our development. Any “expert” can be wrong, and “expert advice” is always wrong in some context, so we don’t get upset if you argue. On the contrary; we’re delighted if you argue. Challenging our statements shows us that you’re thinking about how things might work—or not—in your own environment. The conversation is often very enlightening.
What is testing?

What is the testing mentality?
Testing lights the way.

This is our role.  
Testers see things for what they are.  
Testing makes informed decisions about quality possible,  
by looking critically at software.
Testing is...

- Acquiring the competence, motivation, and credibility to...
- create the conditions necessary to...
- evaluate a product by learning about it through exploration and experimentation, which includes to some degree: questioning, study, modeling, observation and inference, including...
- verifying propositions about the product...
- ...so that you help your clients to make informed decisions about risk.
- And perhaps help make the product better, too.
Six Mentalities of Responsible Tester

Each of these mentalities is an independent way of thinking to some degree, but the outer ones depend on the foundation provided by the inner ones.
Six Mentalities of Responsible Tester

Agent

"I make choices for myself.
I am accountable for my choices."
Six Mentalities of Responsible Tester

The craftsman mentality is about assuring the quality of your own work; seeking to be ever better at whatever you choose to do with your life. Craftsmanship means taking the time and energy to study your craft, whatever that is. And when your work is not good enough, it means you make it right.

"I take pride in my work.  
I strive to be better in my work."
The analyst mentality includes the skills and attitudes of studying complex things, modeling them, and applying those models. It requires the courage to deal with things that are confusing and poorly explained. It requires curiosity. It involves the design and application of heuristic problem-solving methods.

“IA study the world and figure it out.”
The service provider mentality is about being useful to others. It’s about being able to do a job for money, deliver the goods, and get along with people while you do so. There are different ways to succeed as a provider, including being especially sociable or especially reliable.

"I serve; I support; I provide."
The critic mentality includes the skills and attitudes of thinking critically and scientifically. It’s the art of productively complaining. It’s the habit of questioning safe assumptions. It requires faith that there are problems here even if you haven’t yet found them. You can’t be a good tester if you are a poor critic.

“I will find trouble wherever I look!”
Six Mentalities of Responsible Tester

Tester

The tester mentality adds all the nuts and bolts of designing specific tests for your specific product. It includes how to operate your product, how to use tools, wrangling of test data, configuring test environments, and all that nitty gritty.

“I evaluate the product so my client can make good decisions about it.”
What About Supporting Testers?

A supporting tester is someone who helps test but is ultimately not responsible for the quality of the test strategy or the test process.

“What would you like me to do for you?”

Tester

Critic

Analyst

Craftsman

Agent

Provider
The developer mentality has much in common with that of a tester… but there is one huge difference.
Responsible Developer

Maker
Analyst
Craftsman
Agent

"I make things that end your troubles!"
Critical Distance and Social Distance

* A Creative Tension

Social distance is generally bad.

Critical distance is generally good.
RST Agile Testing Contexts v1.2
James Bach and Michael Bolton

Realization
Deep testing can help us discover subtle, rare, or unanticipated problems that frustrate the intent of the product.

Intention
Design-focused testing can help us discover what the customer really wants.

Testing is...
Evaluating a product by learning about it through experimentation and exploration.

Preparation
Part of testing is advocating to make the product reasonably testable and organizing the project and systems to facilitate testing work.

Discipline
Shallow testing can help us discover if what we just did is reasonably close to what we intended to do.

High Value of Product

Low Cost of Development

Studying It

Building It
Why do some people think a role is a prison or a fortress?

This happens when role is defined as the *only things you do* and what *no one else does.*

**HIGH SOCIAL DISTANCE**
We like to think of roles this way.

A role is like a villa. It is a semi-private space. Someone dwells in it. Someone is the responsible tester. But visitors may offer support.

- Devs help testers.
- Testers help devs.
- But responsible testers are ACCOUNTABLE for test process.

VARIABLE SOCIAL DISTANCE
The Testing Role Heuristic

The testing specialist role is a heuristic for developing, operating, and expressing good testing mentality across the organization.

If testing is nobody’s focus, it will not be anybody’s competence.
Section 2

Where are the bugs?

To find bugs, you must cover the product. For that, you need good models of it.
Testing lights the way.

- **Quality** is value to some person (who matters).
- A **bug** is anything about the product that threatens its value.

- These definitions are designed to be inclusive.
- Inclusive definitions minimize the chance that you will inadvertently overlook an important problem.

The first question about quality is always “whose opinions matter?” If someone who doesn’t matter thinks your product is terrible, you don’t necessarily change anything. So, an important bug is something about a product that really bugs someone important.

One implication of this is that, if we see something that we think is a bug and we’re overruled, we don’t matter. That’s a fact: testers don’t matter. That is to say, our standards are not the standards to which the product must adhere; we don’t get to make decisions about quality. This can be hard for some testers to accept, but it’s the truth. We provide information to managers; they get to make the hard decisions. They get the big bucks, and we get to sleep at night.
The important parts of testing don’t take place in the computer or on your desk.
Models Link Observation and Inference

- **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 represents another idea, activity, or object.**
  such as something complex that you need to work with or study.

- **...A GOOD model is one that helps you understand or manipulate the thing that it represents.**
  - A map helps navigate across a terrain.
  - “2 + 2 = 4” is a model for adding two apples to a basket that already has two apples in it.
  - Atmospheric models help predict where hurricanes will go.
  - A fashion model helps people to understand how clothing would look on actual humans.
  - Your beliefs about what you test are a model of what you test.
Testers must distinguish observation from inference!
- Our mental models form the link between them
- Defocusing is lateral thinking.
- Focusing is logical (or “vertical”) thinking.
A Heuristic Test Strategy Model

Project Environment

Tests

Quality Criteria

Product Elements

Testing and Quality Story
A Heuristic Test Strategy Model

Risk!

Procedures

Diversification

Cost vs. Value

Testability

Tools

Testing and Quality Story

Risk! Oracles

Risk! Coverage

Risk!
Elements of Testing

1. Model the product space and risks.
2. Consider what to cover.
3. Consider oracles.
4. Determine test procedures that cover product and apply oracles.
5. Configure the test system.
6. Operate the test system.
7. Observe the test system.
8. Evaluate the test results.

Most testing is driven by questions about risk...

...So, it helps to relate test results back to risk.

By “model”, we mean that we must build a representation in our minds of what the product must do. If our model is incorrect or limited, then we automatically will not test what should be tested.

The first four items on the list are usually called “test design”. The latter five items are usually called “test execution”.

Coverage

____ coverage is how thoroughly you have examined the product with respect to some model of ____.

- **Interesting kinds of coverage**
  - Product coverage: *What aspects of the product did you look at?*
  - Risk coverage: *What risks have you tested for?*
  - Requirements coverage: *What requirements have you tested for?*
In RST methodology, a test factor is something that could be examined in a test, or that could influence the outcome of a test.

All of these are in your testing...
Which of them are you talking about?
Typical industry practice...

All of these are in your testing...
Which of them are you talking about?
Our suggested improvement...

All of these are in your testing...
Which of them are you talking about?
Want to cover the product?

**SFDIPOT**

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

*Remember: “San Francisco Depot”*
Product Elements: *Structural Coverage*

Test what it's made of.

- Print testing example
  - Files associated with printing
  - Code modules that implement printing
  - Code statements inside the modules
  - Code branches inside the modules
Product Elements: *Functional Coverage*

- Print testing example
  - Print, page setup and print preview
  - Print range, print copies, zoom
  - Print all, current page, or specific range
Product Elements: 
*Data Coverage*

- Test what it does it to.

### Print testing example
- Types of documents
- Items in documents, size and structure of documents
- Data about how to print (e.g. zoom factor, no. of copies)
Product Elements: *Interface Coverage*

- Test the ways we can interact with it.

**Print testing example**
- User interface elements; menus, windows, dialogs, controls
- Programming interfaces for launching and controlling a print job
- Printing to or from files
- USB printer interface protocol
- Physical connections between printer and computers and/or network
Product Elements: 
*Platform Coverage*

- **Test what it depends upon.**

- **Print testing example**
  - Printers, spoolers, network behavior
  - Computers
  - Operating systems
  - Printer drivers
Product Elements: Operations Coverage

- Test how it's used.

- Print testing example
  - Use defaults
  - Use realistic environments
  - Use realistic scenarios
  - Use complex flows
Product Elements: *Time Coverage*

- **Print testing example**
  - Try different network or port speeds
  - Print one document right after another, or after long intervals
  - Try time-related constraints—spooling, buffering, or timeouts
  - Try printing hourly, daily, month-end, and year-end reports
  - Try printing from two workstations at the same time
  - Try printing again, later.
SECTION 3

How do you think like a tester?

Well, think like a SCIENTIST!
Learn to love solving mysteries.
Do not confuse testing with fact checking.

A mechanistic process for verifying propositions about a product... means

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

Evaluate: Apply algorithmic decision rules to those observations.

Report: Report any rule violation as a FAIL, otherwise call it PASS.
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)

Even though a check itself is skill-free, good checking is surrounded by activities that require many skills, including testing skill, programming skill, and project management skill.

Before the check happens,

• someone must recognize and identify a risk. That takes testing skill.
• someone must translate that risk into a test idea. That takes testing skill.
• someone must express a test idea as a one-bit, yes-or-no question. That takes testing and programming skill.
• someone must turn the question into code. That takes programming skill.
• someone must determine the trigger—an event or action that launches the check. That takes testing skill.
• someone must encode the trigger so that it can be acted on automatically. That takes programming skill.

After the check happens,

• someone must read the bit. That takes programming skill.
• someone must aggregate bits. That takes programming skill.
• someone must design a report. That takes design and testing skill.
• someone must encode the report. That takes programming skill.
• someone must observe the report. That takes testing skill.
• someone must determine meaning. That takes testing and product management skill.
• someone must determine significance. That takes testing and project management skill.
• someone must respond. That takes at least one of several of the skills above.

Variability is not necessarily a liability of human testing. Requisite variety is essential to finding problems that can appear in diverse ways in diverse places.
Healthy Perspective About Checks

1. We like checks. We use checks. They help us discover if the product can work.
3. And demonstrating that a product can work is far from showing that it will work.
4. The performance of a check can be automated. But designing and programming and interpreting and improving checks is testing.
Managing Tacit Knowledge
Explicit knowledge can be like a trellis...

It can be a matrix onto which tacit knowledge grows...
Tacit and Explicit Knowledge

**EXPLICIT** means it is 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.
- **Collective Tacit Knowledge** is tacit in your community.

(see Collins, *Tacit and Explicit Knowledge*)
Reflex is IMPORTANT
But Critical Thinking is About Reflection

System 2

Slower
Surer

Faster
Looser

logical thinking
de-bias
consider alternatives
solicit review
get more data
suspended judgment
formal models

REFLECTION

statistical thinking

cognitive heuristics
priming

REFLEX

norms
habits
emotions
memorable examples
inattentional blindness

See Thinking Fast and Slow, by Daniel Kahneman
Scientific Skills

- posing useful questions
- observing what’s going on
- describing what you perceive
- thinking critically about what you know
- recognizing and managing bias
- designing hypotheses and experiments
- thinking despite already knowing
- analyzing someone else’s thinking
- reasoning about cause and effect
- treating “facts” as merely what we believe we know as of this moment
Testing is about questions...
*posing them and answering them*

**Huh?**
- You may not understand. (errors in interpreting and modeling a situation, communication errors)

**Really?**
- What you understand may not be true. (missing information, observations not made, tests not run)

**And?**
- You may not know the whole story. (perhaps what you see is not all there is)

**So?**
- The truth may not matter, or may matter much more than you think. (poor understanding of risk)
Testing is about questions...

you can apply them to anything.

- Words and Pictures
- Causation
- The Product
  - Design
  - Behavior
- The Project
  - Schedule
  - Infrastructure
- The Test Strategy
  - Coverage
  - Oracles
  - Procedures

Sometimes stating an assumption is a useful proxy for asking a question.
Thinking Like A Tester:
Seeing the Words that Aren’t There

- Among other things, testers question premises.
- A suppressed premise is an unstated premise that an argument needs in order to be logical.
- A suppressed premise is something that should be there, but isn’t...
- (...or is there, but it’s invisible or implicit.)
- Among other things, testers bring suppressed premises to light and then question them.
- A diverse set of models can help us to see the things that “aren’t there.”
Thinking Like A Tester:
Spot the missing words!

▪ “I performed the tests. All my tests passed. Therefore, the product works.”
▪ “The programmer said he fixed the bug. I can’t reproduce it anymore. Therefore it must be fixed.”
▪ “Microsoft Word frequently crashes while I am using it. Therefore it’s a bad product.”
▪ “Step 1. Reboot the test system.”
▪ “Step 2. Start the application.”
Think Critically About Requirements

**How do you test this?**

“The system shall operate at an input voltage range of nominal 100 - 250 VAC.”

**Poor answer:**

“Try it with an input voltage in the range of 100-250.”
## General Requirements for Observation

<table>
<thead>
<tr>
<th>Observer</th>
<th>Object</th>
<th>Environment</th>
</tr>
</thead>
<tbody>
<tr>
<td>Motivated</td>
<td>Motivating</td>
<td>Consequential</td>
</tr>
<tr>
<td>Attentive</td>
<td>Noticeable</td>
<td>Non-Distracting</td>
</tr>
<tr>
<td>Literate</td>
<td></td>
<td>Legible</td>
</tr>
<tr>
<td>Discerning</td>
<td></td>
<td>Distinguishable</td>
</tr>
<tr>
<td>Sensitive</td>
<td>Visible</td>
<td>Transparent</td>
</tr>
</tbody>
</table>
General Requirements for Observation

<table>
<thead>
<tr>
<th>Motivated</th>
<th>Motivating</th>
</tr>
</thead>
<tbody>
<tr>
<td>Consequential</td>
<td>Attentive</td>
</tr>
<tr>
<td>Noticeable</td>
<td></td>
</tr>
<tr>
<td>Literate</td>
<td></td>
</tr>
<tr>
<td>Discerning</td>
<td></td>
</tr>
<tr>
<td>Visible</td>
<td></td>
</tr>
</tbody>
</table>

In testing, *MOST* of the product is invisible unless made visible by **specific operations**. "Alternation blink oracles" can greatly help our ability to see small differences. Computer-assisted oracles, such as conditional highlighting, also help.

Being a subject matter expert makes it far easier to spot problems, because you know what important things are supposed to mean.

When a product is designed to use standardized interface conventions it becomes much easier to make sense of it.
General Requirements for Observation

<table>
<thead>
<tr>
<th></th>
<th>Object</th>
<th>Environment</th>
</tr>
</thead>
<tbody>
<tr>
<td>Motivated</td>
<td>Motivating</td>
<td>Consequential</td>
</tr>
<tr>
<td>Attentive</td>
<td>Noticeable</td>
<td>Non-Distracting</td>
</tr>
<tr>
<td>Literate</td>
<td>Legible</td>
<td></td>
</tr>
<tr>
<td>Discerning</td>
<td>Distinguishable</td>
<td></td>
</tr>
</tbody>
</table>

Sometimes important bugs such as creeping data corruption are quite hard to notice unless you have tool support or special preparation.

People who follow scripts that they didn’t write and don’t understand tend to be less attentive and motivated.

When what you observe doesn’t matter then you are far less likely to observe it well.

“Inattentional blindness” may cause us to miss even large “obvious” problems.
Sometimes important bugs such as creeping data corruption are quite hard to notice unless you have tool support or special preparation.

When management doesn’t seem to care about problems, or makes it too hard to argue about them, we tend to lose motivation.

"Inattentional blindness" may cause us to miss even large “obvious” problems.

People who follow scripts that they didn’t write and don’t understand tend to be less attentive and motivated.

Testability!
To describe what you see, you must **decide** what you are seeing.

<table>
<thead>
<tr>
<th>Observer</th>
<th>Object</th>
<th>Environment</th>
</tr>
</thead>
<tbody>
<tr>
<td>Motivated</td>
<td>Motivating</td>
<td>Consequential</td>
</tr>
<tr>
<td>Attentive</td>
<td>Noticeable</td>
<td>Non-Distracting</td>
</tr>
<tr>
<td>Literate</td>
<td>Legible</td>
<td></td>
</tr>
<tr>
<td>Discerning</td>
<td>Distinguishable</td>
<td></td>
</tr>
<tr>
<td>Sensitive</td>
<td>Visible</td>
<td>Transparent</td>
</tr>
</tbody>
</table>
Consider the Spellchecker in FocusWriter

No spell checking dictionary is chosen by default
Magic Tricks and Blindness

- **Our thinking is limited**
  - We misunderstand probabilities
  - We use the wrong heuristics
  - We lack specialized knowledge
  - We forget details
  - We don’t pay attention to the right things

- **The world is hidden**
  - states
  - sequences
  - processes
  - attributes
  - variables
  - identities

Magic tricks work for the same reasons that bugs exist.

Studying magic can help you develop the imagination to find better bugs.

Testing magic is indistinguishable from testing sufficiently advanced technology.

---

*Our colleague, Jeremy Kominar, is an expert amateur magician. We asked him to provide us with some remarks on the links between software testing and magic. Here’s an edited version of what he had to say—we can’t quote him exactly, because he asked us not to give away his secrets!*

- Learning to perform magic causes the magician to be aware of not only his own perspective, but also (and more importantly) others’ perception of the trick. This becomes invaluable for testers because it helps us to recognize that the interests and perspectives of many stakeholders must be considered when testing software. The trick may “look good” or “work” from your angle, but there are always other angles to cover.

- When learning to perform magic tricks, you generally want to simplify things—minimize the number of moves and figure out how you can reach the same end result without as many steps to get there. In a testing context this idea can be used for creating different usage scenarios—taking multiple paths and variables to attain the same goal. (We might want to minimize the number of steps, but we might also want to vary the paths to shake out more bugs.)

- Learning to perform magic and learning to figure out magic tricks require us to develop observational skills, and to recognize things that deceive us. Most non-magicians fail to consider possibilities that come easily to magicians; devices, coins, or cards can be gimmicked. Non-magicians only think about normal cards. You need to think outside of the box to be a magician or to figure out how the magician does his work. This kind of reasoning and deduction are key assets to being a tester. There isn’t *really* such a thing as magic but there clearly is such a thing as deception. As the magician becomes more experienced and more wise to the practice, he should be able to reverse-engineer tricks by using patterns that he already knows about magic. Once you’ve seen one kind of trick, similar tricks aren’t so mysterious because you have heuristics that you can use to recognize how it’s done.*
KEY IDEA

**When a bug happens, how do you know?**

*Know your oracles.*
Jerry Weinberg has suggested that “it works” may mean “We haven't tried very hard to make it fail, and we haven't been running it very long or under very diverse conditions, but so far we haven't seen any failures, though we haven't been looking too closely, either.” In this pessimistic view, you have to be on guard for people who say it works without checking even once to see if it could work.
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

Mipping is a way of bug reporting when you aren’t confident in the value of a bug. It’s an acronym that stands for “mention in passing.” To MIP a bug is to report it in an informal and inexpensive way, such as through email or chat. That way, you don’t get in trouble for not reporting it, and you don’t get in trouble for over-reporting it.

Black flagging is the opposite of that. When you black flag a bug you put a lot of energy into the report. You may write a white paper about it, and call a meeting to discuss it. This is necessary in cases where a simple bug report won’t capture the extent of the risk, or when you feel that fixing the bug will not adequately deal with the risks related to that bug (perhaps because you feel that the product may need a comprehensive re-design. Usually we black flag a bug when it seems to represent a cluster of potentially related problems that collectively pose a large risk.
General Examples of Oracles
Some are more authoritative than others

- A person whose opinion matters.
- An opinion held by a person who matters.
- A disagreement among people who matter.
- A known good example output.
- A known bad example output.
- A reference document with useful information.
- A process or tool by which the output is checked.
- A process or tool that helps a tester identify patterns.
- A feeling you have, such as confusion or annoyance.
- A desirable consistency between related things.
Consistency ("this agrees with that")

an important theme in oracles

- **Acceptability**: The product is consistent with how good it could reasonably be; not just good, but also good enough.
- **Familiarity**: The product is not consistent with the pattern of any familiar problem.
- **Explainability**: The product is consistent with our ability to explain it.
- **World**: The product is consistent with objects and states, in the world, that it represents.
- **History**: The product is consistent with itself over time.
- **Image**: The product is consistent with an image that the organization wants to project.
- **Comparable Products**: The product is consistent with other products or parts of products that work in a comparable way.
- **Claims**: The product is consistent with what important people say it’s supposed to be.
- **Users’ Desires**: The product is consistent with what users want.
- **Product**: Each part of the product is consistent with comparable parts of that product.
- **Purpose**: The product is consistent with its purposes, both explicit and implicit.
- **Statutes & Standards**: The product is consistent with laws and relevant standards.
In exploratory testing, the first oracle to trigger is usually a personal feeling. Then you move from that to something more explicit and defensible. We need to move in that direction because testing is social. We probably can’t just go with our gut and leave it at that, because we have to convince other people that it’s really a bug.
All Oracles Are Heuristic

- We often do not have oracles that establish a definite correct or incorrect result, in advance. **That’s why we use oracles heuristically.**
- No single oracle can tell us whether a program (or a feature) is working correctly at all times and in all circumstances. **That’s why we use a variety of oracles.**
- Any program that looks like it’s working, to you, may in fact be failing in some way that happens to fool all of your oracles. **That’s why we proceed with humility and critical thinking.**
- You (the tester) can’t know the deep truth about any result. **That’s why we report whatever seems likely to be a bug.**
Coping With Difficult Oracle Problems

- **Ignore the Problem**
  - Ask “so what?” Maybe the value of the bug doesn’t justify the cost of detecting it.

- **Simplify the Problem**
  - Ask for testability. It usually doesn’t happen by accident.
  - Automated oracle. Internal error detection or external output analyzer.
  - Lower the standards. You may be using an unreasonable standard of correctness.

- **Shift the Problem**
  - Parallel testing. Compare with another instance of a comparable algorithm.
  - Live oracle. Find an expert who can tell if the output is correct.
  - Reverse the function. (e.g. 2 x 2 = 4, then 4/2 = 2)

- **Divide and Conquer the Problem**
  - Spot check. Perform a detailed inspection on one instance out of a set of outputs.
  - Blink test. Compare or review overwhelming batches of data for patterns that stand out.
  - Easy input. Use input for which the output is easy to analyze.
  - Easy output. Some output may be obviously wrong, regardless of input.
  - Unit test first. Gain confidence in the pieces that make the whole.
  - Test incrementally. Gain confidence by testing over a period of time.

Types of “Easy Input” oracles:

*Fixed Markers.* Use distinctive fixed input patterns that are easy to spot in the output.

*Statistical Markers.* Use populations of data that have distinguishable statistical properties.

*Self-Referential Data.* Use data that embeds metadata about itself. (e.g. counterstrings)

*Easy Input Regions.* For specific inputs, the correct output may be easy to calculate.

*Outrageous Values.* For some inputs, we expect error handling.

*Idempotent Input.* Try a case where the output will be the same as the input.

*Match.* Do the “same thing” twice and look for a match.

*Progressive Mismatch.* Do progressively differing things over time and account for each difference. (code-breaking technique)
Quality Criteria are Diverse

**CRUCSS CPID**

- Capability
- Reliability
- Usability
- Charisma
- Security
- Scalability
- Compatibility
- Performance
- Installability
- Development

Many test approaches focus on capability (functionality) and underemphasize the other criteria

**Development. How well can we create, test, and modify it?**

- **Supportability:** How economical will it be to provide support to users of the product?
- **Testability:** How effectively can the product be tested?
- **Maintainability:** How economical is it to build, fix or enhance the product?
- **Portability:** How economical will it be to port or reuse the technology elsewhere?
- **Localizability:** How economical will it be to adapt the product for other places?
KEY IDEA

How do you invent the right tests, at the right time?

Treat testing as an exploratory process. Resist premature formalization.
Deep Testing Defined

Deep testing is testing which seeks to **maximize** the chance of finding all the important bugs.

Deep testing **requires** substantial skill, effort, preparation, time, or tooling, etc., to fulfill its charter.
Not Everything Needs Deep Testing

- Only build easy things.
- Stop caring if it’s broken.
- Build in a testable way.

Testing is hard!

Testing is easy!
How do you get deep?

- **Time**: All this takes time to develop, maintain, and perform.
- **Determination**: It’s hard to blunder into deep testing. You have to want it.
- **Skills**: You need to know how to model products, identify testable elements, and design experiments to evaluate them.
- **Learning**: You need to develop a rich and detailed model of the product and its risks (may be a mental model, formal model, or both).
- **Requisite Variety of Test Activities**: You need to work out a pattern of test activities that will find the obscure, yet important bugs, based on a good theory of risk.
- **Tooling**: You may need tools to help you cover large areas or to reach otherwise inaccessible areas of the product.
- **Environments for Testing**: You need a requisite variety of test platforms configured and available for tester use.
- **Data for Testing**: You need a requisite variety of test data so you can trigger the important bugs.
- **Team Support**: You may need lots of eyes and minds poring over it. Developers can help immensely by exposing the code.
- **Testability**: You may need special features in the product that help you observe and control it.
The Testing Formality Continuum
*Mixing Scripting and Exploration*

**INFORMAL**
Not done in any specific way, nor to verify specific facts.

- Survey
- Exploratory
- Play

**FORMAL**
Done in a specific way, or to verify specific facts.

- Analytical
- Matrix/Outline of Test Factors
- Vague/Generic Test Scripts
- "Human Transceiver"
- Specific Test Data
- Human Checking
- Machine Checking

When I say “exploratory testing” and don’t qualify it, I mean anything on the informal side of this continuum.
Testing

Evaluate by learning through experimentation and exploration including questioning, modeling, observation, inference, etc.

Human Checking  Both  Machine Checking
Exploratory Process is Structured

- Testing, as I teach it, is a structured process conducted by a skilled tester, or by lesser skilled testers or users working under supervision.

- The structure of testing comes from many sources:
  - Test design heuristics
  - Chartering
  - Time boxing
  - Perceived product risks
  - The nature of specific tests
  - The structure of the product being tested
  - The process of learning the product
  - Development activities
  - Constraints and resources afforded by the project
  - The skills, talents, and interests of the tester
  - The overall mission of testing

In other words, it’s not “random”, but systematic.
Heuristics bring useful structure to problem-solving skill.

- **adjective:**
  “serving to discover.”

- **noun:**
  “a fallible means of solving a problem.”

“The engineering method is the use of heuristics to cause the best change in a poorly understood situation within the available resources.”
-- Billy Vaughan Koen, *Discussion of The Method*

Key ideas about heuristics:

- A heuristic is not an *edict*. Heuristics require guidance and control of skilled practitioner.
- Heuristics are context-dependent.
- Heuristics may be useful even when they contradict each other—especially when they do!
- Heuristics can substitute for complete and rigorous analysis.

Types of heuristics:

- Guideword Heuristics: Words or labels that help you access the full spectrum of your knowledge and experience as you analyze something.
- Trigger Heuristics: Ideas associated with an event or condition that help you recognize when it may be time to take an action or think a particular way. Like an alarm clock for your mind.
- Subtitle Heuristics: Help you reframe an idea so you can see alternatives and bring out assumptions during a conversation.
- Heuristic Model: A representation of an idea, object, or system that helps you explore, understand, or control it.
- Heuristic Procedure or Rule: A plan of action that may help solve a class of problems.
- Cognitive Heuristics: Heuristics built into your mind by virtue of how human brains work. These show themselves as systematic biases in our senses (optical illusions) and decision-making.
Heuristics

They are **applied**, not followed.

This...

The skilled tester remains in control of the process.

...not this.

1. Do this
2. Then do this
3. Then do this
4. Then do this
5. And then this...

Scripted procedures give the illusion of control over unskilled testers.

Everyday examples of heuristics:

- It’s dangerous to drink and drive.
- A bird in hand is worth two in the bush.
- Nothing ventured, nothing gained.
- Sometimes people stash their passwords near their computers. Try looking there.
- Stores are open later during the Holidays.
- If your computer is behaving strangely, try rebooting. If it’s very strange, reinstall Windows.
- If it’s a genuinely important task, your boss will follow-up, otherwise, you can ignore it.
Testing with a Formal Procedure

Follow wisely
Follow ignorantly
Deviate ignorantly
Deviate wisely

Powerful Test (probably)

Improves the test
Finds a great bug
Comprises a sanity check
Comprises a reasonable test of some kind

Fresh Powerful Test (probably)

Spoils the test
Misses a great bug

Competent Supervision Required
ET is a Structured Process

In excellent exploratory testing, one structure tends to dominate all the others:

The Testing Story

Exploratory testers construct a compelling story of their testing. It is this story that gives ET a backbone.
To test is to construct three stories
*(plus one more)*

**Level 1:** A story about the status of the PRODUCT...
...about how it failed, and how it *might* fail...
...in ways that matter to your various clients.

**Level 2:** A story about HOW YOU TESTED it...
...how you configured, operated and observed it...
...about what you haven’t tested, yet...
...and won’t test, at all...

**Level 3:** A story about the VALUE of the testing...
...what the risks and costs of testing are...
...how testable (or not) the product is...
...things that make testing harder or slower...
...what you need and what you recommend...

*(Level 3+ : A story about the VALUE of the stories.)*
...do you know what happened? can you report? does report serve its purpose?
If You are Frustrated

De-Focus!

1. Look over your recent tests and find a pattern there.
2. With your next few tests, *violate* the old pattern.
3. Prefer *MFAT* (multiple factors at a time).
4. Broaden and vary your observations.
If You are Confused

Focus!

1. Simplify your tests.
2. Conserve states.
3. Frequently repeat your actions.
4. Frequently return to a known state.
5. Prefer OFAT heuristic (one factor at a time).
6. Make precise observations.

“State Conservation” means that you minimize disturbances to the product state, maintaining it as much as possible while using the product.
Test Design and Execution

Achieve excellent test design by exploring different test designs while actually testing.

Guide testers with personal supervision and concise documentation of test ideas. Meanwhile, train them so that they can guide themselves and be accountable for increasingly challenging work.
Allow some *disposable time*
*Self-management is good!*

- How often do you account for your progress?
  - Every minute?
  - Every hour?
  - Every day?

- If you have *any autonomy at all*, you can risk investing *some time* in
  - learning
  - thinking
  - refining approaches
  - better tests
Allow some *disposable time*

😊 If it turns out that you’ve made a good investment, you might have
- learned something about the product
- invented a more powerful test
- found a bug
- done a better job
- avoided going down a dead end for too long
- surprised and impressed your manager

😁 If it turns out that you’ve made a bad investment... *oh well*
“Plunge in and Quit” Heuristic

Whenever you are called upon to test something very complex or frightening, plunge in! After a little while, if you are very confused or find yourself stuck, quit!

- This benefits from *disposable time*—that part of your work not scrutinized in detail.
- Plunge in and quit means you can start something without *committing* to finish it successfully, and therefore *you don’t need a plan*.
- Several cycles of this is a great way to *discover* a plan.
Exploratory Branching:

*Distractions are good!*

New test ideas occur continually during an ET session.
“Long Leash” Heuristic

Let yourself be distracted...

‘cause you never know what you’ll find...

but periodically take stock of your status against your mission
KEY IDEA

How do you perform a test?

Know your procedures.
A **test activity** is a line of investigation that fulfills some part of the test strategy. It can encompass many test cases.

A **test case** is one particular instance or variation of a test or test idea.

A **test procedure** is a way of performing a test.

- What role do you play in it?
- What role do tools play?
- Who controls your procedures?
- Should they be documented? How?

Some people prefer a more restrictive definition of procedure. Cem Kaner considers a sequential set of steps to be an important aspect of procedure. He would prefer that we use the word “test implementation” instead of procedure, in this section.

However, in the Rapid Testing methodology, we usually opt for the most inclusive definitions that still express the key ideas, since that makes the conversation easier. In our view, the word procedure as it is commonly used is not limited only to sequential instruction sets.

As always, it’s up to you what words you use.
Test Procedure

*has four elements*

- **Configure**
  - *(if necessary)* Obtain product for testing
  - *(if necessary)* Install on a platform.
  - *(if necessary)* Prepare test data and tools for test execution.
  - Assure that the product is in a “clean enough” starting state.

- **Operate**
  - Control the product and platform with inputs to exercise the product.
  - Cause the product to exercise the right functions/states in the right sequence with the right data.

- **Observe**
  - Collect information about how the product behaves (collect both direct and indirect output) so that it can be evaluated.

- **Evaluate**
  - Apply oracles to detect bugs.
To maximize test integrity...

Focus!

1. Start the test from a known (clean) state.
2. Prefer simple, deterministic actions.
3. Trace test steps to a specified model.
4. Follow established and consistent lab procedures.
5. Make specific predictions, observations and records.
6. Make it easy to reproduce (automation helps).
To find *unexpected problems*, *elusive problems* that occur in sustained field use, or more problems *quickly* in a complex product...

**De-Focus!**

1. Start from *different states* (not necessarily clean).
2. Prefer *complex, challenging* actions.
3. Generate tests from a *variety* of models.
4. *Question* your lab procedures and tools.
5. Try to *see everything* with open expectations.
6. Make the test *tougher*, instead of more reproducible.
KEY IDEA

How do you test a whole product?

Use a diversified risk-based strategy.
The Test Project Context

Where are you and what are you doing there?

Mission
- Find Important Problems
- Assess Quality
- Certify to Standard
- Fulfill Process Mandates
- Satisfy Stakeholders
- Assess Accountability
- Advise about QA
- Advise about Testing
- Advise about Quality
- Maximize Efficiency
- Minimize Cost
- Minimize Time

Development
- Product
- Project Lifecycle
- Project Management
- Configuration Management
- Defect Prevention
- Development Team

Requirements
- Product Mission
- Stakeholders
- Quality Criteria
- Reference Material

Test Team
- Expertise
- Loading
- Cohesion
- Motivation
- Leadership
- Project Integration

Test Lab
- Test Platforms
- Test Tools
- Test Library
- Problem Tracking System
- Office Facilities

Test Process
- Strategy
- Logistics
- Work-products
Test Strategy

- **Strategy:** “The set of ideas that guide your test design.”
- **Logistics:** “The set of ideas that guide your application of resources to fulfilling the test strategy.”
- **Plan:** “The set of ideas that guide your test project.”

A good test strategy is:
- Product-Specific
- Risk-focused
- Diversified
- Practical

\[ \text{plan} = \text{strategy} + \text{logistics} \]
One way to make a strategy...

1. Learn the product.
2. Think of important potential problems.
3. Think of how to search the product for \textit{those problems}.
4. Think of how to search the product, \textit{in general}.

Think of ways that:
\begin{itemize}
  \item will take advantage of the \textit{resources} you have.
  \item comprise a \textit{mix} of different techniques.
  \item comprise something that you really \textit{can} actually do.
  \item serve the \textit{specific} mission you are expected to fulfill.
\end{itemize}
What does “taking advantage of resources” mean?

**MIDTESTD**

- **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...”?

The General Test Techniques

- Function testing
- Domain testing
- Stress testing
- Flow testing
- Scenario testing
- Claims testing
- User testing
- Risk testing
- Automatic checking

We’ll come back to these later. For now, note that we’re more likely to catch different bugs with different nets; a diverse set of test techniques will find a diverse set of bugs.
How much is enough?

*Diverse Half-Measures*

- There is no single technique that finds all bugs.
- We can’t do any technique perfectly.
- We can’t do all conceivable techniques.

Use “*diverse half-measures*” -- lots of different points of view, approaches, techniques, even if no one strategy is performed completely.
The Product Risk Knowledge Gap:
*How much do we already know?*

Our knowledge of the status of the product.
Value (or Risk) as a Simplifying Factor

*Find problems that matter*

- In general it can vastly simplify testing if we focus on whether the product has a problem that matters, rather than whether the product merely satisfies all relevant standards.
- Effective testing requires that we understand standards as they relate to how our clients value the product.

Instead of thinking pass vs. fail, consider thinking problem vs. no problem.
Risk-Based Test Project Cycle: 
*Testing itself is risk analysis.*

New Project → Analyze Potential Risks → Perform Appropriate Testing → Experience Problems & Potentialities → Analyze Actual Risks → Experience Problems In the Field → New Project

Risk analysis involves identifying potential and actual risks, analyzing them, and then performing appropriate testing. The cycle is continuous, with short and long loops.

Testing is itself a form of risk analysis: risk-based testing begins with a rumor of risk and transforms that into information.

Even if testing reveals a problem that cannot be reproduced, you have learned something about risk.

*Exploratory* testing is a popular and powerful approach to analyzing risk. *Scripted* testing is better suited for monitoring known risk.
What is testing?
Serving Your Client

If you don’t have an understanding and an agreement on what is the mission of your testing, then doing it “rapidly” would be pointless.

Know your mission.

Begin sympathetically...

...Then chase the risk!
Cost as a Simplifying Factor

Try quick tests as well as careful tests

A quick test is a cheap test that has some value but requires little preparation, knowledge, or time to perform.

- Happy Path
- Tour the Product
  - Sample Data
  - Variables
  - Files
  - Complexity
  - Menus & Windows
  - Keyboard & Mouse
- Interruptions
- Undermining
- Adjustments
- Dog Piling
- Continuous Use
- Feature Interactions
- Click on Help

Happy Path: Use the product in the most simple, expected, straightforward way, just as the most optimistic programmer might imagine users to behave. Perform a task, from start to finish, that an end-user might be expected to do. Look for anything that might confuse, delay, or irritate a reasonable person.

Documentation Tour: Look in the online help or user manual and find some instructions about how to perform some interesting activity. Do those actions. Improvise from them. If your product has a tutorial, follow it. You may expose a problem in the product or in the documentation; either way, you've found something useful. Even if you don't expose a problem, you'll still be learning about the product.

Sample Data Tour: Employ any sample data you can, and all that you can—the more complex or extreme the better. Use zeroes where large numbers are expected; use negative numbers where positive numbers are expected; use huge numbers where modestly-sized ones are expected; and use letters in every place that's supposed to handle numbers. Change the units or formats in which data can be entered. Challenge the assumption that the programmers have thought to reject inappropriate data.

Variables Tour: Tour a product looking for anything that is variable and vary it. Vary it as far as possible, in every dimension possible. Identifying and exploring variations is part of the basic structure of my testing when I first encounter a product.

Complexity Tour: Tour a product looking for the most complex features and using challenging data sets. Look for nooks and crowds where bugs can hide.

File Tour: Have a look at the folder where the program's .EXE file is found. Check out the directory structure, including subs. Look for READMEs, help files, log files, installation scripts, .cfg, .ini, .rc files. Look at the names of .DLLs, and extrapolate on the functions that they might contain or the ways in which their absence might undermine the application.

Menus and Windows Tour: Tour a product looking for all the menus (main and context menus), menu items, windows, toolbars, icons, and other controls.

Keyboard and Mouse Tour: Tour a product looking for all the things you can do with a keyboard and mouse. Run through all of the keys on the keyboard. Hit all the F-keys. Hit Enter, Tab, Escape, Backspace. Run through the alphabet in order. Combine each key with Shift, Ctrl, and Alt. Also, click on everything.

Interruptions: Start activities and stop them in the middle. Stop them at awkward times. Perform stoppages using cancel buttons, O/S level interrupts (ctrl-alt-delete or task manager), arrange for other programs to interrupt (such as screensavers or virus checkers). Also try suspending an activity and returning later.

Undermining: Start using a function when the system is in an appropriate state, then change the state part way through (for instance, delete a file while it is being edited, eject a disk, pull net cables or power cords) to an inappropriate state. This is similar to interruption, except you are expecting the function to interrupt itself by detecting that it no longer can proceed safely.

Adjustments: Set some parameter to a certain value, then, at any later time, reset that value to something else without resetting or recreating the containing document or data structure.

Dog Piling: Get more processes going at once; more states existing concurrently. Nested dialog boxes and non-modal dialogs provide opportunities to do this.

Continuous Use: While testing, do not reset the system. Leave windows and files open. Let disk and memory usage mount. You're hoping that the system ties itself in knots over time.

Feature Interactions: Discover where individual functions interact or share data. Look for any interdependencies. Tour them. Stress them. I once crashed an app by loading up all the fields in a form to their maximums and then traversing to the report generator. Look for places where the program repeats itself or allows you to do the same thing in different places.

Click for Help: At some point, some users are going to try to bring up the context-sensitive help feature during some operation or activity. Does
the product’s help file explain things in a useful way, or does it offend the user’s intelligence by simply restating what’s already on the screen? Is help even available at all?
Cost as a Simplifying Factor

Try quick tests as well as careful tests

A quick test is a cheap test that has some value but requires little preparation, knowledge, or time to perform.

- Input Constraint Attack
- Click Frenzy
- Shoe Test
- Blink Test
- Error Message Hangover
- Resource Starvation
- Multiple Instances
- Crazy Configs
- Cheap Tools

Input Constraint Attack: Discover sources of input and attempt to violate constraints on that input. For instance, use a geometrically expanding string in a field. Keep doubling its length until the product crashes. Use special characters. Inject noise of any kind into a system and see what happens. Use Satisfice’s PerlClip utility to create strings of arbitrary length and content; use PerlClip’s counterstring feature to create a string that tells you its own length so that you can see where an application cuts off input.

Click Frenzy: Ever notice how a cat or a kid can crash a system with ease? Testing is more than “banging on the keyboard”, but that phrase wasn’t coined for nothing. Try banging on the keyboard. Try clicking everywhere. I broke into a touchscreen system once by poking every square centimeter of every screen until I found a secret button.

Shoe Test: This is any test consistent with placing a shoe on the keyboard. Basically, it means using auto-repeat on the keyboard for a very cheap stress test. Look for dialog boxes so constructed that pressing a key leads to, say, another dialog box (perhaps an error message) that also has a button connected to the same key that returns to the first dialog box. That way you can place a shoe (or Coke can, as I often do, but sweeping off a cowboy boot has a certain drama to it) on the keyboard and walk away. Let the test run for an hour. If there’s a resource or memory leak, this kind of test will expose it.

Blink Test: Find some aspect of the product that produces huge amounts of data or does some operation very quickly. For instance, look a long log file or browse database records very quickly. Let the data go by too quickly to see in detail, but notice trends in length or look or shape of the data. Some bugs are easy to see this way that are hard to see with detailed analysis. Use Excel’s conditional formatting feature to highlight interesting distinctions between cells of data.

Error Message Hangover: Make error messages happen and test hard after they are dismissed. Often developers handle errors poorly.

Resource Starvation: Progressively lower memory, disk space, display resolution, and other resources until the product collapses, or gracefully (we hope) degrades.

Multiple Instances: Run a lot of instances of the app at the same time. Open the same files. Manipulate them from different windows.

Crazy Configs: Modify the operating system’s configuration in non-standard or non-default ways either before or after installing the product. Turn on “high contrast” accessibility mode, or change the localization defaults. Change the letter of the system hard drive. Consider that the product has configuration options, too—change them or corrupt them in a way that should trigger an error message or an appropriate default behavior.

Cheap Tools: Learn how to use InCtrl5, Filemon, Regmon, AppVerifier, Perfmon, and Process Explorer, and Task Manager (all of which are free). Have these tools on a thumb drive and carry it around. Also, carry a digital camera. I now carry a tiny 3 megapixel camera and a tiny video camera. Both fit into my coat pockets. I use them to record screen shots and product behaviors. While it’s not cheap, you can usually find Excel on most Windows systems; use it to create test matrices, tables of test data, charts that display performance results, and so on. Use the World-Wide Web Consortium’s HTML Validator at http://validator.w3c.org. Pay special attention to tools that hackers use; these tools can be used for good as well as for evil. Netcat, Burp Proxy, wget, and fuzzer are but a few examples.
Can tests be repeated?

Only partly...

- You can’t be certain that you control all the factors that might matter to fulfill the purpose of a test.
- So, to “repeat a test” means that you believe you are repeating some part of a test that matters, while other parts of that test may not be repeated.
- Even if you repeat “just 1% of a test”, it may be fair to say that you have repeated that test in every way that matters.

Exact repetition is not an option.
Should tests be repeated?

*Think twice before repeating yourself...*

- Consider cost vs. value.
- Consider what could be gained from some of the many tests you have not yet *ever* performed.
- To test is to question:
  - Why ask the same question again?
  - Why ask the same question again?
  - Why ask the same question again?
  - Why ask the same question again?
- You should know why, and why *not*, to repeat tests.
Should Tests Be Repeated?

*Exploiting Variation To Find More Bugs*

- **Micro-behaviors:** Unreliable and distractible humans make each test a little bit new each time through.
- **Randomness:** Can protect you from unconscious bias.
- **Data Substitution:** The same actions may have dramatically different results when tried on a different database, or with different input.
- **Platform Substitution:** Supposedly equivalent platforms may not be.
- **Timing/Concurrency Variations:** The same actions may have different results depending on the time frame in which they occur and other concurrent events.
- **Scenario Variation:** The same functions may operate differently when employed in a different flow or context.
- **State Pollution:** Hidden variables of all kinds frequently exert influence in a complex system. By varying the order, magnitude, and types of actions, we may accelerate state pollution, and discover otherwise rare bugs.
- **Sensitivities and Expectations:** Different testers may be sensitive to different factors, or make different observations. The same tester may see different things at different times or when intentionally shifting focus to different things.
KEY IDEA

How do you record your work?

Use concise, modular documents that help tell the testing story.
Common Claims About Test Documentation

- It’s needed for new testers.
- It’s needed to deflect legal liability.
- It’s needed to share testing with other testers.
- It’s needed for “repeatability.”
- It’s needed for accountability.
- It forces you to think about test strategy.

Something is wrong with these claims. Can you see the problem?
The First Law of Documentation

“That should be documented.”

*really means...*

“That should be documented if and when and how it serves our purposes.”

Who will read it? Will they understand it? Is there a better way to communicate that information? What does documentation cost you?
Common Problems with Test Documentation

- Distracts, and even prevents, testers from doing the best testing they know how to do.
- Authors don’t understand testing.
- Authors don’t own format.
- Templates help hide poor thinking.
- Full of fluff.
- Fluff discourages readers and increases cost of maintenance.
- No one knows documentation requirements.
- Too much formatting increases the cost of maintenance.
- Information for different audiences is mixed together.
- Catering to rare audiences instead of probable user.
- Disrupts the social life of information.
- Long term and short term goals conflict.
- Most people don’t read.
What Does Rapid Testing Look Like?

**Concise Documentation Minimizes Waste**

### General
- Testing Heuristics
- Risk Catalog

### Project-Specific
- Product Coverage Outline
- Product Risks
- Test Strategy
- Schedule
- Issues
- Bugs
- Status Dashboard

Detailed procedural documentation is expensive and largely unnecessary.

Tutorial documentation is also usually unnecessary, but if you do it, then keep it separate from the working documents.
Consider Automatic Logging

- Exploratory testing works better when the product produces an automatic log of everything that was covered in each test.
- You can also use external logging tools such as Spector (www.spectorsoft.com).
- Automatic logging means that you get something like a retrospective script of what was tested.
Taking Notes

- Coverage
- Oracles
- Procedures (key elements only)
- Test Ideas
- Bugs/Risks

- Issues, Questions & Anomalies
  - It would be easier to test if you changed/added...
  - How does ... work?
  - Is this important to test? How should I test it?
  - I saw something strange...
High Accountability Exploratory Testing

1) Charter
2) Time Box
3) Reviewable Result
4) Debriefing

VS.
Charter:
*A clear mission for the session*

- A charter may suggest what should be tested, how it should be tested, and what problems to look for.
- A charter is not meant to be a detailed plan.
- General charters may be necessary at first:
  - “Analyze the Insert Picture function”
- Specific charters provide better focus, but take more effort to design:
  - “Test clip art insertion. Focus on stress and flow techniques, and make sure to insert into a variety of documents. We’re concerned about resource leaks or anything else that might degrade performance over time.”
Time Box:

*Focused test effort of fixed duration*

- Brief enough for accurate reporting.
- Brief enough to allow flexible scheduling.
- Brief enough to allow course correction.
- Long enough to get solid testing done.
- Long enough for efficient debriefings.
- Beware of overly precise timing.

Short: 60 minutes (+-15)

**Normal: 90 minutes (+-15)**

Long: 120 minutes (+-15)
Debriefing:

*Measurement begins with observation*

- The manager reviews *session sheet* to assure that he understands it and that it follows the protocol.
- The tester answers any questions.
- Session metrics are checked.
- Charter may be adjusted.
- Session may be extended.
- New sessions may be chartered.
- Coaching happens.
Reviewable Result:  
*A scannable session sheet*

- Charter  
  - #AREAS
- Start Time  
- Tester Name(s)  
- Breakdown  
  - #DURATION  
  - #TEST DESIGN AND EXECUTION  
  - #BUG INVESTIGATION AND REPORTING  
  - #SESSION SETUP  
  - #CHARTER/OPPORTUNITY  
- Data Files

- Test Notes
- Bugs  
  - #BUG
- Issues  
  - #ISSUE
The Breakdown Metrics

*Testing is like looking for worms*

- Test Design and Execution
- Session Setup
- Bug Investigation and Reporting
Reporting the TBS Breakdown  
*A guess is okay, but follow the protocol*

- Test, Bug, and Setup are orthogonal categories.
- Estimate the percentage of charter work that fell into each category.
- Nearest 5% or 10% is good enough.
- If activities are done simultaneously, report the highest precedence activity.
- Precedence goes in order: T, B, then S.
- All we really want is to track interruptions to testing.
- Don’t include Opportunity Testing in the estimate.
Activity Hierarchy

All test work fits here, somewhere

all work

non-
session
inferred

session

opportunity

on charter

test

bug

setup
Work Breakdown: 
*Diagnosing the productivity*

- Do these proportions make sense?
- How do they change over time?
- Is the reporting protocol being followed?
Coverage:

*Specifying coverage areas*

- These are text labels listed in the Charter section of the session sheet. (e.g. “insert picture”)
- Coverage areas can include anything
  - areas of the product
  - test configuration
  - test strategies
  - system configuration parameters
- Use the debriefings to check the validity of the specified coverage areas.
Coverage:

Are we testing the right stuff?

- Is this a risk-based test strategy?

  or

- Is it a lop-sided set of coverage areas?
- Is it distorted reporting?
Using the Data to Estimate a Test Cycle

1. How many perfect sessions (100% on-charter testing) does it take to do a cycle? *(let’s say 40)*

2. How many sessions can the team (of 4 testers) do per day? *(let’s say 3 per day, per tester = 12)*

3. How productive are the sessions? *(let’s say 66% is on-charter test design and execution)*

4. Estimate: $40 / (12 \times .66) = 5 \text{ days}$

5. We base the estimate on the data we’ve collected. When any conditions or assumptions behind this estimate change, we will update the estimate.
KEY IDEA

How do you effectively report your work?

Learn to tell a compelling story that provokes the right questions.
Reporting Considerations

- **Reporter safety**: What will they think if I made no progress?
- **Client**: Who am I reporting to and how do I relate to them?
- **Rules**: What rules and traditions are there for reporting, here?
- **Significance of report**: How will my report influence events?
- **Subject of report**: On what am I reporting?
- **Other agents reporting**: How do other reports affect mine?
- **Medium**: How will my report be seen, heard, and touched?
- **Precision and confidence levels**: What distinctions make a difference?

*Take responsibility for the communication.*
The Dashboard Concept

Large dedicated whiteboard “Do Not Erase”

Project conference room

Project status meeting
Test Cycle Report

Product Areas

vs.

Test Effort

Test Coverage

Quality Assessment

vs.

Time
<table>
<thead>
<tr>
<th>Area</th>
<th>Effort</th>
<th>C.</th>
<th>Q. Comments</th>
</tr>
</thead>
<tbody>
<tr>
<td>file/edit</td>
<td>high</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>view</td>
<td>low</td>
<td>1+</td>
<td>1345, 1363, 1401</td>
</tr>
<tr>
<td>insert</td>
<td>low</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>format</td>
<td>low</td>
<td>2+</td>
<td>automation broken</td>
</tr>
<tr>
<td>tools</td>
<td>blocked</td>
<td>1</td>
<td>crashes: 1406, 1407</td>
</tr>
<tr>
<td>slideshow</td>
<td>low</td>
<td>2</td>
<td>animation memory leak</td>
</tr>
<tr>
<td>online help</td>
<td>blocked</td>
<td>0</td>
<td>new files not delivered</td>
</tr>
<tr>
<td>clipart</td>
<td>none</td>
<td>0</td>
<td>need help to test...</td>
</tr>
<tr>
<td>converters</td>
<td>none</td>
<td></td>
<td>need help to test...</td>
</tr>
<tr>
<td>install</td>
<td>start 3/17</td>
<td>0</td>
<td>lab time is scheduled</td>
</tr>
<tr>
<td>compatibility</td>
<td>start 3/17</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>general GUI</td>
<td>low</td>
<td>3</td>
<td></td>
</tr>
</tbody>
</table>
Product Area

- 15-30 areas (keep it simple)
- Avoid sub-areas: they’re confusing.
- Areas should have roughly equal value.
- Areas together should be inclusive of everything reasonably testable.
- “Product areas” can include tasks or risks - but put them at the end.
- Minimize overlap between areas.
- Areas must "make sense" to your clients, or they won’t use the board.

<table>
<thead>
<tr>
<th>Area</th>
</tr>
</thead>
<tbody>
<tr>
<td>file/edit</td>
</tr>
<tr>
<td>view</td>
</tr>
<tr>
<td>insert</td>
</tr>
<tr>
<td>format</td>
</tr>
<tr>
<td>tools</td>
</tr>
<tr>
<td>slideshow</td>
</tr>
<tr>
<td>online help</td>
</tr>
<tr>
<td>clipart</td>
</tr>
<tr>
<td>converters</td>
</tr>
<tr>
<td>install</td>
</tr>
<tr>
<td>compatibility</td>
</tr>
<tr>
<td>general GUI</td>
</tr>
</tbody>
</table>
## Test Effort

<table>
<thead>
<tr>
<th>Level</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>None</td>
<td>Not testing; not planning to test.</td>
</tr>
<tr>
<td>Start</td>
<td>No testing yet, but expect to start soon.</td>
</tr>
<tr>
<td>Low</td>
<td>Regression or spot testing only; maintaining coverage.</td>
</tr>
<tr>
<td>High</td>
<td>Focused testing effort; increasing coverage.</td>
</tr>
<tr>
<td>Pause</td>
<td>Temporarily ceased testing, though area is testable.</td>
</tr>
<tr>
<td>Blocked</td>
<td>Can’t effectively test, due to blocking problem.</td>
</tr>
<tr>
<td>Ship</td>
<td>Going through final tests and signoff procedure.</td>
</tr>
</tbody>
</table>
Test Effort

- Use red to denote significant problems or stoppages, as in blocked, none, or pause.
- Color ship green once the final tests are complete and everything else on that row is green.
- Use neutral color (such as black or blue, but pick only one) for others, as in start, low, or high.
## Test Coverage

<table>
<thead>
<tr>
<th>Rating</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>We don’t have good information about this area.</td>
</tr>
<tr>
<td>1</td>
<td>Sanity Check: major functions &amp; simple data.</td>
</tr>
<tr>
<td>1+</td>
<td>More than sanity, but many functions not tested.</td>
</tr>
<tr>
<td>2</td>
<td>Common &amp; Critical: all functions touched; common &amp; critical tests executed.</td>
</tr>
<tr>
<td>2+</td>
<td>Some data, state, or error coverage beyond level 2.</td>
</tr>
<tr>
<td>3</td>
<td>Complex Cases: strong data, state, error, or stress testing.</td>
</tr>
</tbody>
</table>
Test Coverage

- Color green if coverage level is acceptable for ship, otherwise color black.
- Level 1 and 2 focus on functional requirements and capabilities: can this product work at all?
- Level 2 finds any “obvious” bug.
- Level 2+ and 3 focus on information to judge performance, reliability, compatibility, and other “ilities”: will this product work under realistic usage?
- Level 3 or 3+ implies “if there were a bad bug in this area, we would probably know about it.”
<table>
<thead>
<tr>
<th>Quality Assessment</th>
</tr>
</thead>
<tbody>
<tr>
<td>“We know of no problems in this area that threaten to stop ship or interrupt testing, nor do we have any definite suspicions about any.”</td>
</tr>
<tr>
<td>“We know of problems that are possible showstoppers, or we suspect that there are important problems not yet discovered.”</td>
</tr>
<tr>
<td>“We know of problems in this area that definitely stop ship or interrupt testing.”</td>
</tr>
</tbody>
</table>
Comments

Use the comment field to explain anything colored red, or any non-green quality indicator.

- Problem ID numbers.
- Reasons for pausing, or delayed start.
- Nature of blocking problems.
- Why area is unstaffed.
Using the Dashboard

- **Updates:** 2-5/week, or at each build, or prior to each project meeting.
- **Progress:** Set expectation about the duration of the “Testing Clock” and how new builds reset it.
- **Justification:** Be ready to justify the contents of any cell in the dashboard. The authority of the board depends upon meaningful, actionable content.
- **Going High Tech:** Sure, you can put this on the web, but will anyone actually look at it???
KEY IDEA

Rapid testing is a personal discipline. It requires practice, but no one's permission.
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.
- Understand your users, don’t just read specs.
- Dynamically manage the focus of your work.
- Your context should drive your choices, both of which evolve over time.
Getting started...

- **Take advantage of resources**
  - Email me
  - See my book Lessons Learned in Software Testing
  - Browse the appendices of the class
  - Join forum: software-testing@yahoogroups.com

- **Do some focused exploration**
  - Try a three-hour chartered bug hunt with a group of testers. *See what happens.*
  - Watch yourself as you learn a new product. Notice how your thoughts evolve.

- **Practice solving puzzles**
  - Solve jigsaw puzzles, logic puzzles, Sudoku, cryptograms, or lateral thinking puzzles.
Getting started...

- **Try using guideword heuristics**
  - Use my Heuristic Test Strategy Model
  - Or modify my model
  - Or make your own model

- **Defrost your procedures**
  - Pick some procedures and simplify them.
  - Generalize some of your procedures.
  - Include design information to help test executors contribute.

- **Start a lunch-time test coaching meeting**
  - Pick a part of the product each week and talk about how to test it.
    Get the whole team to brainstorm better ways.
  - Do a risk brainstorm.
  - Review bug reports and talk about how to write better ones.
Getting started...

- **Practice your systems thinking**
  - Read “The Art of Systems Thinking” or “An Introduction to General Systems Thinking”
  - Try the exercises in those books.
  - Take any part of your product and model it in terms of its most important dimensions.

- **Put a coverage/risk outline inside your test documentation**
  - You don’t have to throw out your old test documentation, but how about putting a new section in that consists of a coverage outline, a risk outline, or a combination of the two?

- **Question value vs. cost for all your testing**
  - Is any of your testing not giving you much value relative to its cost? Why are you doing it, then?