



An Exploratory Tester's Notebook

Michael Bolton
DevelopSense
QUEST Chicago
April 2009

Who I Am



Michael Bolton

(not the singer, not the guy in Office Space)

DevelopSense, Toronto, Canada

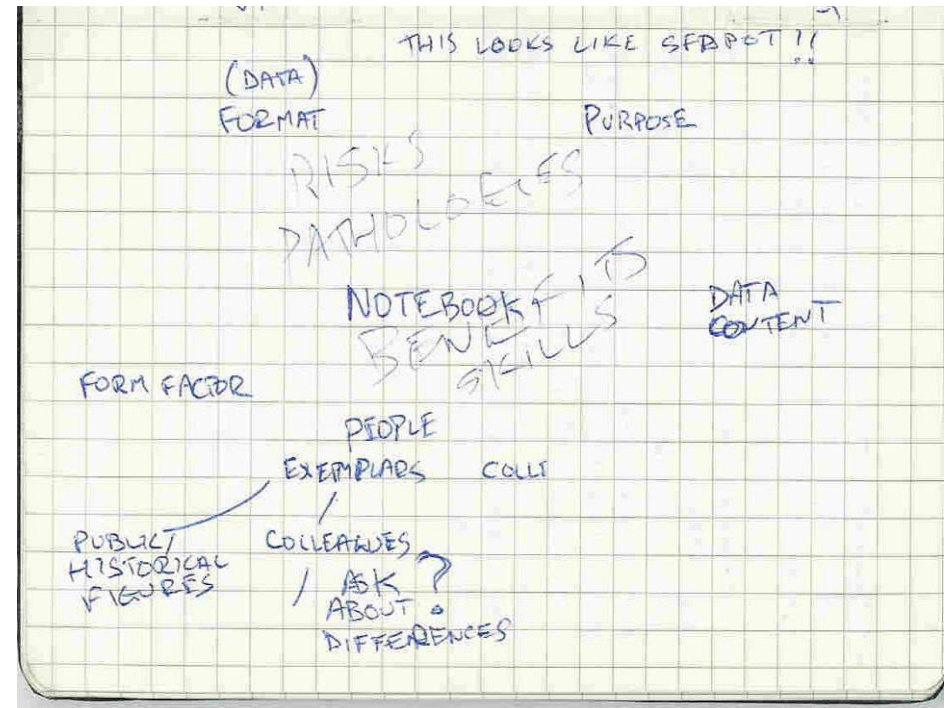
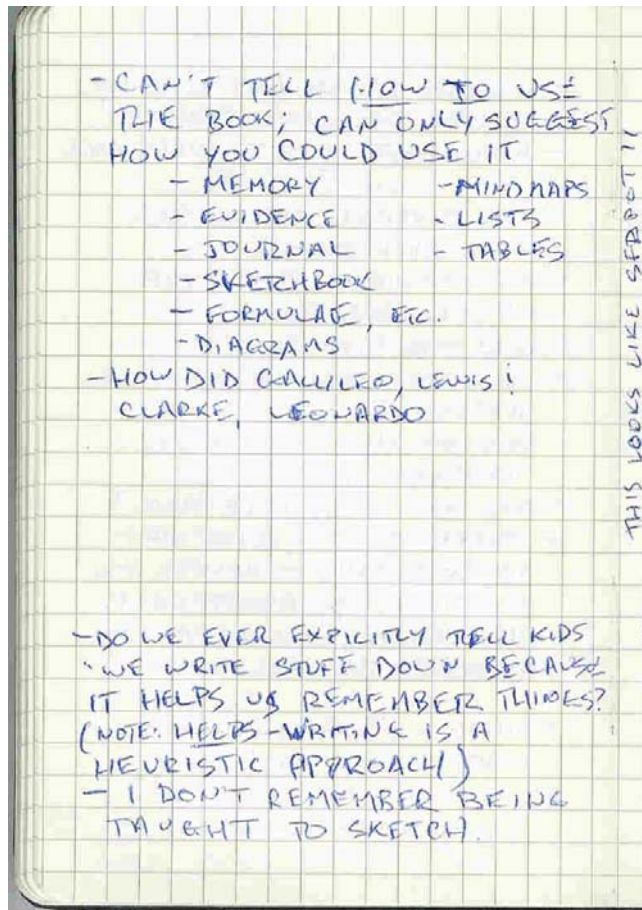
mb@developsense.com

+1 (416) 992-8378

<http://www.developsense.com>

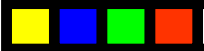
I solve testing problems that other people can't solve.

This Presentation Is Under Continuous Development



- For updated notes AND a more formal paper on notebooks: eurostar@developsense.com

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?**

Documentation: Product or Tool?



Paradigm:



Audience:



Purpose:



Notebooks: A Personal View



- Over the last I've been keeping a set of notebooks
- This is an experience report on how *one* exploratory tester and consultant (me) has used them
- This is a context-driven talk; this is not a best-practices talk


My Introduction to the Moleskine



- I've kept documents (mostly for school or work) all my life
 - scribblers
 - legal pads
 - ASCII text files
 - Word documents
- In January 2004, I noticed Jon Bach's Moleskine notebook
- In January 2005, James Bach suggested I get one. I did.
- It turns out there's a something of a cult...
 - <http://www.moleskinerie.com/>
 - <http://www.moleskinacity.com>
 - <http://en.wikipedia.org/wiki/Moleskine>

So What's the Big Deal?

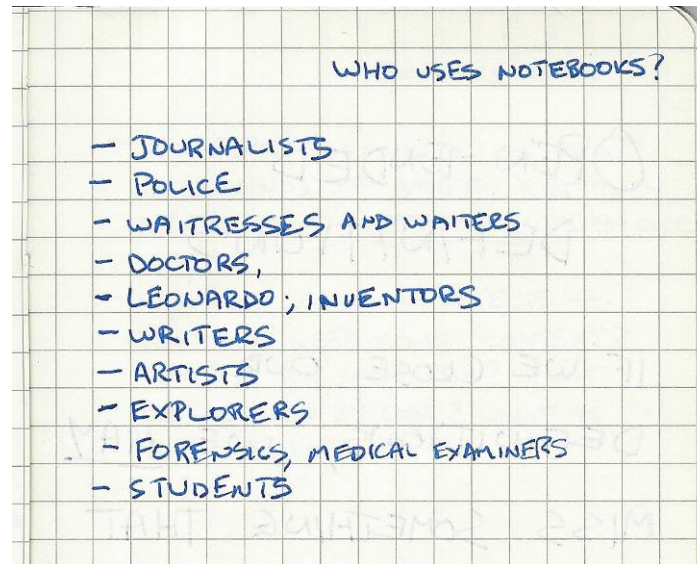


- Several form factors
 - larger notebook
 - smaller notebook (pocket size)
 - reporter style
 - memo pockets
 - Three line styles
 - plain
 - ruled
 - squared
 - Page marker
 - Elastic closure
 - Back pocket
 - Sewn binding, lies flat
 - Geek-chic-mystique-boutique appeal
- 
- Well-constructed
 - Durable
 - Somewhat expensive

Who Uses Notebooks?



People Who Need
to Remember Things



Investigators

Creators

Exploratory testers are all three, and more.

ALL testers are, at some point, explorers.

Maybe we should learn explorers' tools.

Paradigmatic Examples

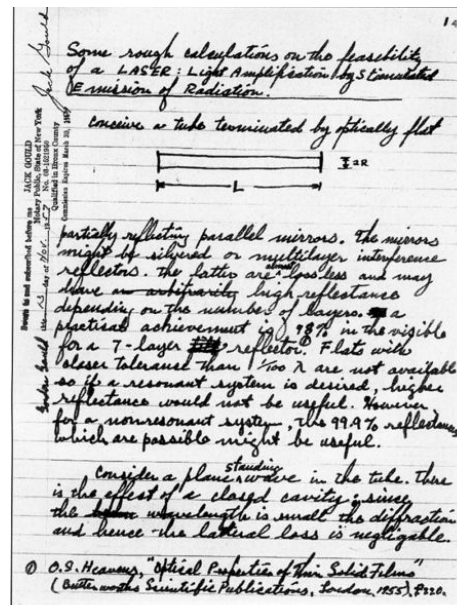
- Leonardo: inventor's notebook

- Codex Arundel, Codex Leicester, Codex Foster, etc.
- contents: sketches, inventions, architecture, elements of mechanics, painting ideas, human anatomy, grocery lists and even people that owed him money (Wikipedia)



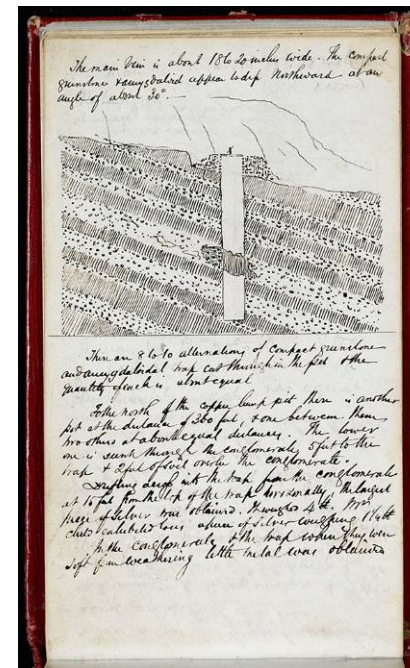
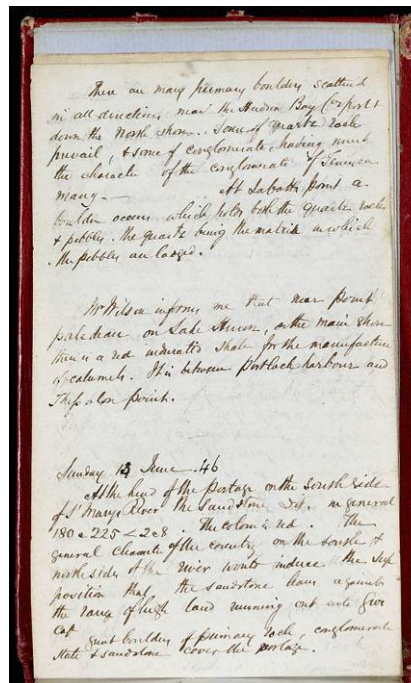
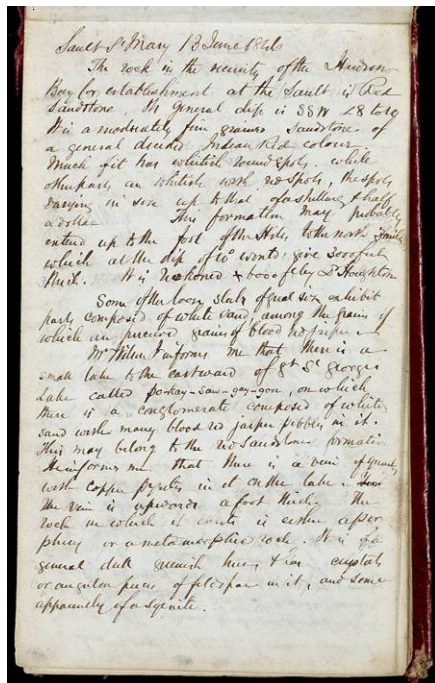
Paradigmatic Examples

- Gordon Gould: inventor's notebook
 - One of the people involved in the invention of the laser.
 - Notes created after meeting with Charles Hard Townes, November 1957 include the acronym "LASER" and several aspects of its design.



Paradigmatic Examples

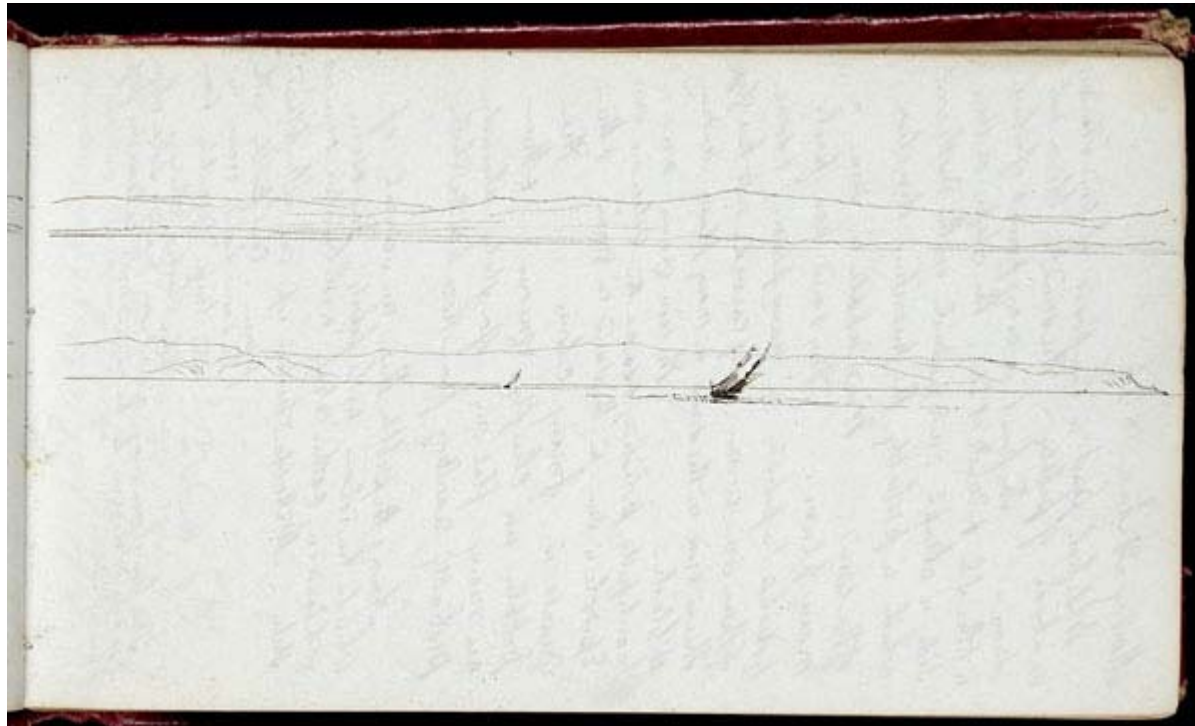
- William Logan: explorer/geologist's notebook
 - Written in Stone—geological explorations of Canada
 - <http://www.collectionscanada.ca/logan/021014-1000-e.html>



1846 Lake Superior

Paradigmatic Examples

- William Logan – Explorer's notebook
 - Written in Stone—geological explorations of Canada
 - <http://www.collectionscanada.ca/logan/021014-1000-e.html>



Why Notebooks Now?

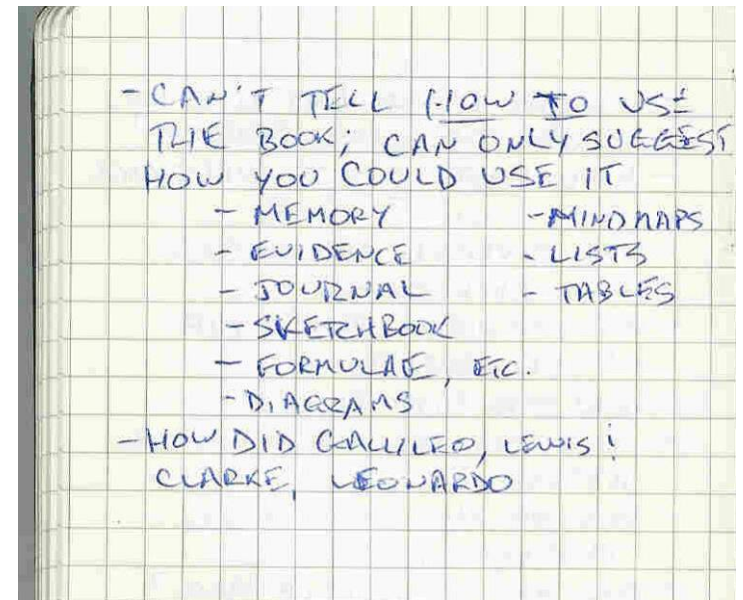


- In the age of the Palm (I have one) and the smartphone (I have one) and the laptop (I have one), why use notebooks?
- They're portable
- They never crash
- They never forget to save
- Battery doesn't wear out
- They're free-form
- They're available
- They're personal

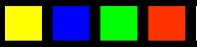
What I Use Notebooks For



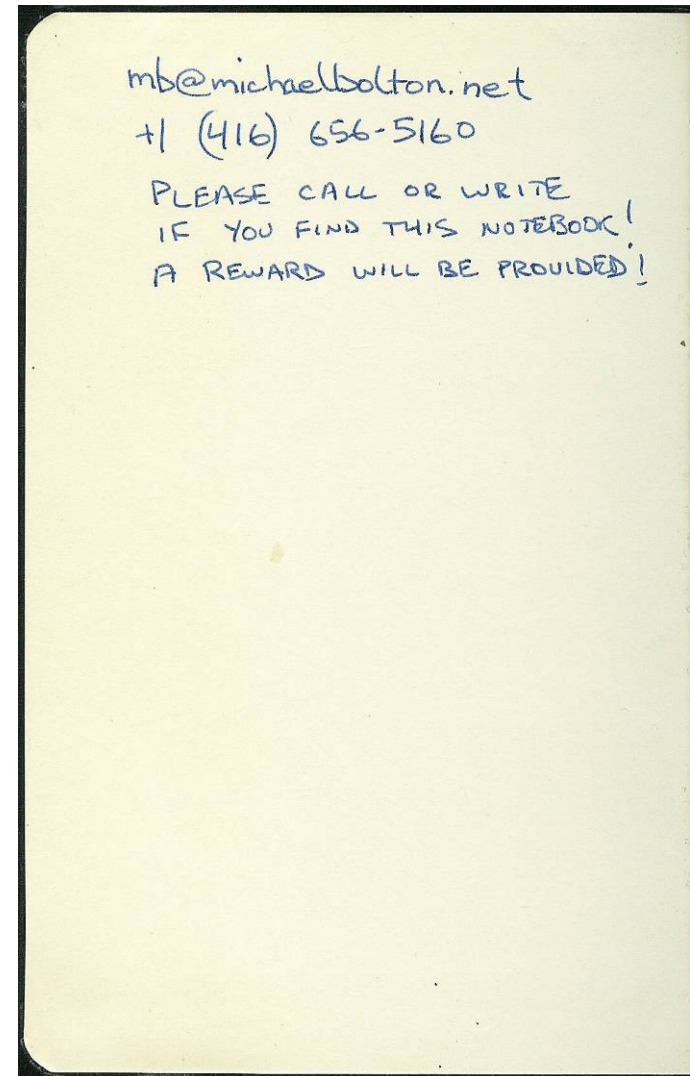
- Brainstorming
- Sketches
- Catalogs of heuristics
- Mind maps
- Diagrams
- Action items and reminders
- “Fieldstones” and blog entries
- Conference or workshop sessions
- Test notes, and practice taking them



My Notebooks



- I thought I lost my notebook once. Now...



A Busy Couple of Days, With Rant



"PARTIAL ANSWERS THAT MIGHT
BE USEFUL."

OBSERVER BIAS APPLIES
TO TESTING AND TESTS -
YOU'LL GET THE EXPECTED
ANSWER.

—
LOOKING @ OTHERS - RECORD

CANON DC 40 X

PANA VDR 300 ✓

SONY HC 96 ✓

SONY DVD 40S/50S X

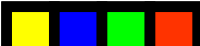
*SONY SR-100 ✓

CANON OPTURA 60 ✓ 50x

USE THE DAMN THING

- YOU CAN TAKE YOUR TIME
DOING AUTOMATION, OR
WRITING SCRIPTS, OR
DRAWING PATH DIAGRAM,
BUT FOR MOST OF THE TESTING
THAT MOST OF US DO, WOULDN'T
IT BE A GOOD IDEA TO USE
THE DAMNED THING FOR
SOMETHING LIKE ITS INTENDED
PURPOSE? WE'RE TEACHING
RST TO ALLOW YOU TO DO
ALL THE OTHER STUFF
QUICKLY, STRAIGHT FROM
YOUR HEAD WHEREVER
POSSIBLE. BUT DON'T
THINK THAT A USER
OF YOUR PRODUCT IS LIKELY

An Exploratory Testing Session



ENTERTAINMENT CONTROL SEAT 37A, KLM 872
HARDWARE


MARCH 18 2006

- TOP OF CASE (GAME/CONTROL PAD) SEPARATING FROM BOTTOM (CELL PHONE) OF CASE @ TOP END OF CONTROL PAD
- NOTICED A KIND OF LATCH OR KEYHOLE @ TOP OF UNIT - BROKEN
- COMPARED UNIT TO SEATMATE'S UNIT - IT WAS BROKEN IN THE SAME WAY, BUT NOT AS BADLY (37B)
- LOOKED FOR MANUFACTURER'S ID - DIDN'T FIND IT - ON-SCREEN? IN MAGAZINE?
- FOUND BIG BREAK NEAR CATCH
- # ISSUE - WOULD HELP TO HAVE CAMERA TO SHORTEN DESCRIPTION
- IDEA THAT CRACKED CASE MAY CAUSE UP BUTTON MALFUNCTION, BUT REMEMBERED OSD DIDN'T SHOW UP AS AN OPTION.
- N.B. THERE'S A CELL PHONE KEYPAD ON THE BOTTOM - WON'T TEST THAT
- THERE'S A CREDIT CARD SWIPING SLOT ON L. SIDE - WON'T TEST THAT
- WEIRD SCREEN TRAVERSAL - HARD TO TELL IF H/W OR S/W
- WHAT'S ON SCREENS
- MOVIE PREVIEWS MUCH MORE BLURRY THAN REST OF OSD -
- DURING MOVIE PREVIEWS, HOME BUTTON OVERLOADED w/ FF >> BUTTON, SO CAN'T GO HOME DURING PREVIEWS
- AT LEAST ONCE ~TOP 1/3 OF SCREEN COVERED BY RED BLOCK

An Exploratory Testing Session

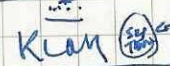


8 HOLES ACROSS

R.G.  17 DOWN

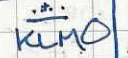
RED BLOCK CONTAINED A GRID OF SINGLE-LINE TRANSPARENT HOLES - CORRECT PICTURE VISIBLE THROUGH THE HOLES

- WAITED TIL END OF PROMO - RED BLOCK DISAPPEARED WHEN MAIN (ENGLISH) MENU REDISPLAYED

| | |
|-----------------------------------------------------------------------------------------------------------------------|--------------------------------------------------------------------------------------|
|  SKY TEAM LOGO | |
| WHAT'S ON MOVIES TV MUSIC GAMES KIDS FLIGHT TRACKING SMS/EMAIL FEEDBACK INFO SCREEN OFF | PICTURE OPTIONS CASE - SENSITIVE NOTE R. SUPPORT FOR AT LEAST 3 ACCENTED CHARS |

0 HOME
0 HELP

- PRESSING HELP FROM MAIN SCREEN DOESN'T PROVIDE HELP; TAKES ME TO LANGUAGE SELECTION SO DOES

| | |
|-------------------------------------------------------------------------------------|--|
|  | |
| ENGLISH NEDERLANDS FRANÇAIS PORTUGUÊS ESPAÑOL DEUTSCH 日本語 中文 | |

3:40 - GETTING SLEEPY
3:43 VERY SLEEPY (STOP)

An Exploratory Testing Session

7:40 AM

- RESTARTED. DID A VERY TENTATIVE AND LOW-IMPACT TEST OF RETRACTING CORD. THE BIG BREAK IS AT THE POINT WHERE THE UNIT HITS A VERY HARD PART OF THE RECEPTACLE — RUBBERIZE THEM?

- TURNED ON THE UNIT. NAVIGATED TO FLIGHT TRACKING SCREEN. SCREEN GOES COMPLETELY OFF, IT APPEARS, WHEN FETCHING THE DATA OR SWITCHING TO THAT MODE

- PRESSED MODE BUTTON — PROMPTS TO RESUME FLIGHT TRACKING OR RETURN TO MENU. CENTRE BUTTON DOESN'T SELECT THE HIGHLIGHTED OPTION — YOU HAVE TO PICK ► ON THE NAV BUTTON

MOVIES

KLM LATEST ~~KLM~~
(15 MOVIES)

~~DOUBT~~ AT ▼ SHOWN AS ONLY OPTION (IMPLICIT) WHEN ▲ ALSO WORKS ◀ TOO

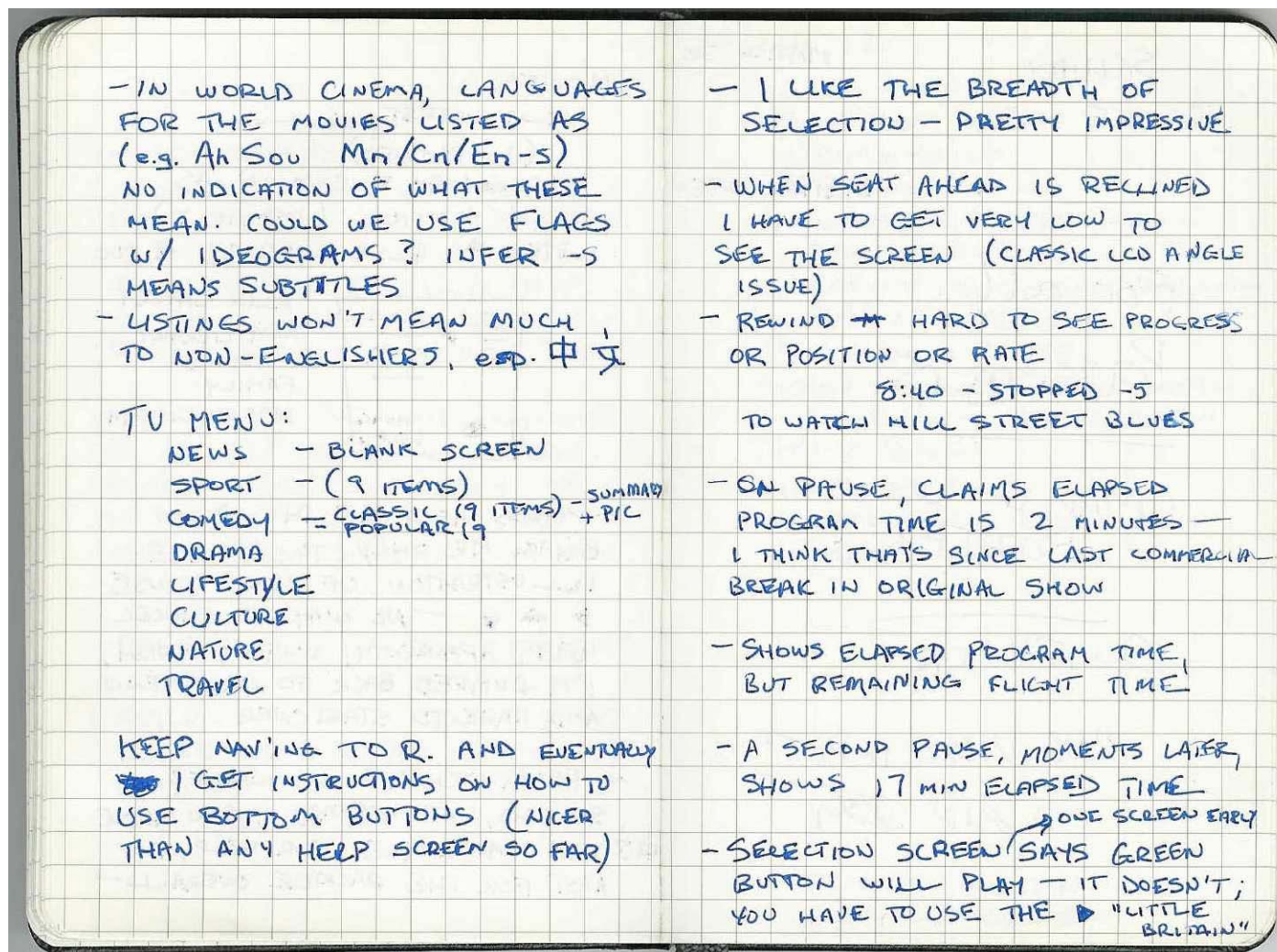


KLM LATEST
KLM LIBRARY
CLASSICS
FAMILY
WORLD CINEMA

- PRESS HELP AT THIS POINT BRING ME ONLY TO A LAME ILLUSTRATION OF HOW TO USE ► ◀ ▼ — NO WAY TO CANCEL HELP, APPARENTLY; WHEN I FINISH, I'M DUMPED BACK TO MAIN MENU AND HAVE TO START OVER

→ FROM "CHOOSE LANGUAGE" SCREEN, IT APPEARS I CAN CHOOSE THE LANGUAGE FOR HELP, BUT NOT FOR THE PACKAGE OVERALL —

An Exploratory Testing Session



Observation on Building Skill

BLOG FIELDSTONE
BARCELONA

2006 07 02

HOW PICASSO GOT TO BE A MASTER
LOBS OF NATURAL TALENT BUT...

- ABILITY TO CAPTURE THE IMPORTANT STUFF i.e. MODELING
- HUGE VOLUME OF WORK
- TRIED LOTS OF DIFFERENT STYLES
- WENT BACK TO BASICS
- LOTS OF PRACTICE
- LOTS OF STUDIES
- HOW MUCH WORK DID WE NOT SEE?
- LOTS OF THROWAWAY EXERCISES
- A PAINTING IS A MODEL

ALSO: NOW THAT WE VIEW HIM
AS A GREAT ARTIST, HIS
MINOR WORKS HAVE GREAT
SIGNIFICANCE ASCRIBED TO THEM.
MAYBE YOUR "INSIGNIFICANT"
WORKS HAVE THE SAME IMPACT.

→ LOTS OF EARLY STUFF IS QUITE
IMITATIVE OF OTHER STYLES -
ATTEMPTS TO DEVELOP SKILLS AND
TO FIGURE OUT WHAT'S IMPORTANT

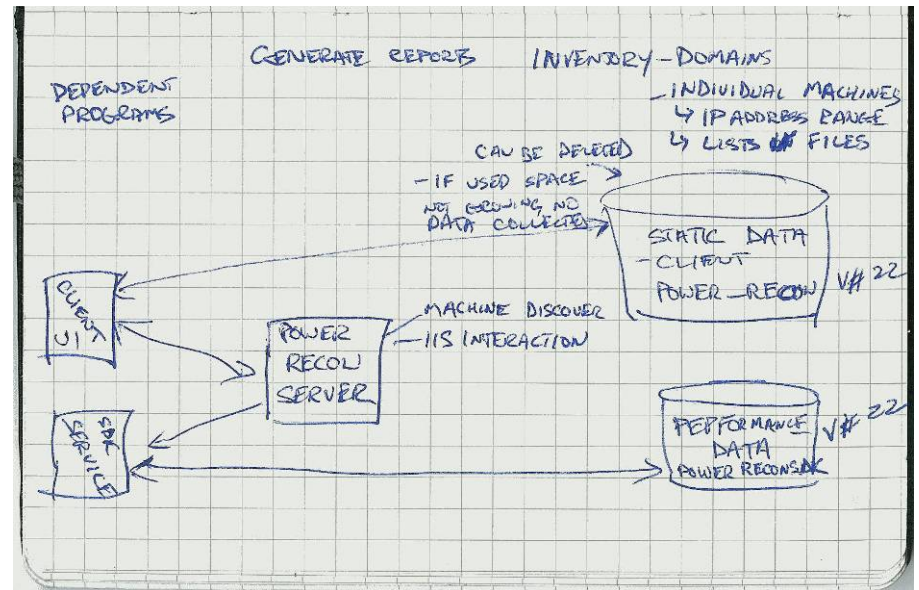
- THERE'S A LINK TO THAT WYNTON
MARRSALLS / CHICK COREA ARTICLE
IN MUSICIAN MAGAZINE, WHERE
ONE HELD THAT LOUIS ARMSTRONG
WAS GREAT BECAUSE OF SKILL, AND
THE OTHER BECAUSE OF HEART AND
SOUL. BUT L.A. HAD BOTH; PEOPLE
DON'T CONNECT AS STRONGLY WITH
ROBERT FRIPP, WHO HAS BAGS OF
SKILL; AND LOTS OF PEOPLE
NEVER GET PUBLISHED BECAUSE,
ALTHOUGH THEY HAVE LOTS TO
SAY, THEY DON'T HAVE THE
SKILL TO SAY IT WELL ENOUGH.

Diagramming

- “The diagram is nothing; the diagramming is everything.”
 - Jerry Weinberg

James Bach on White-Box Risk-Based Analysis, with Diagrams

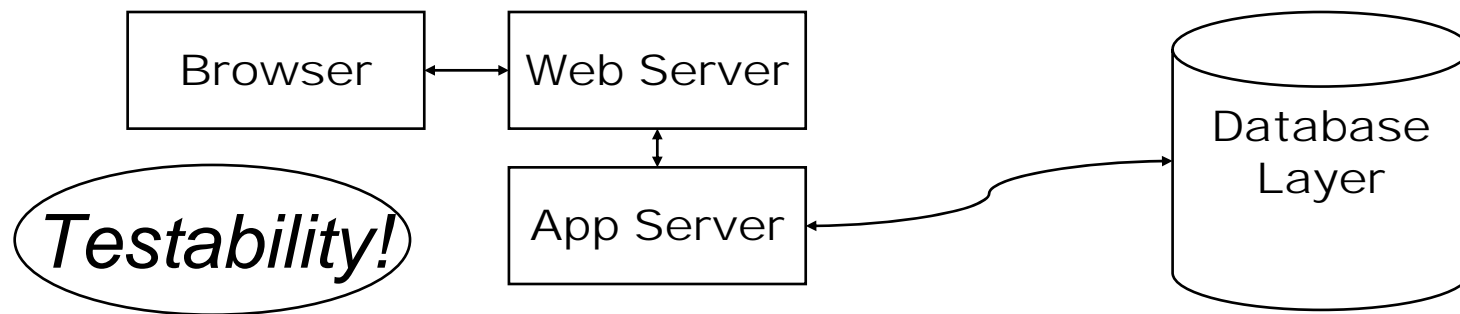
- [pointing at a box] *What if the function in this box fails?*
- *Can this function ever be invoked at the wrong time?*
- [pointing at any part of the diagram] *What error checking do you do here?*
- [pointing at an arrow] *What exactly does this arrow mean? What would happen if it was broken?*



Guideword Heuristics for Diagram Analysis



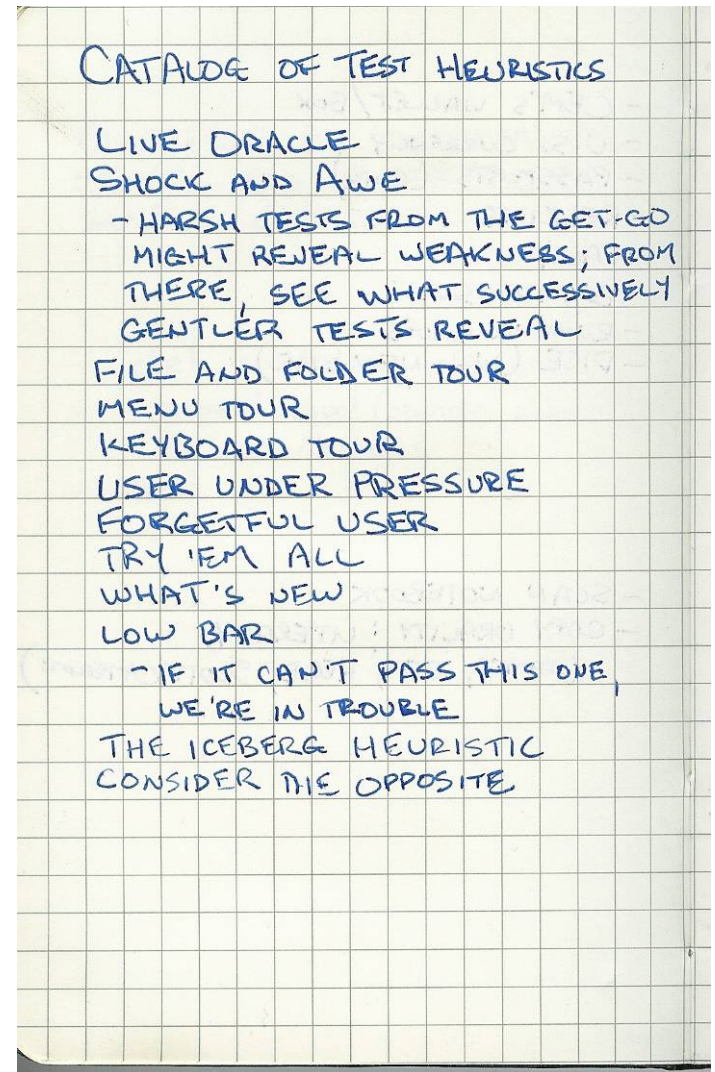
- Boxes
 - *Missing/Drop-out*
 - *Extra/Interfering*
 - *Incorrect*
 - *Timing/Sequencing*
 - *Contents/Algorithms*
 - *Conditional behavior*
 - *Limitations*
 - *Error Handling*
- Lines
 - *Missing/Drop-out*
 - *Extra/Forking*
 - *Incorrect*
 - *Timing/Sequencing*
 - *Status Communication*
 - *Data Structures*
- Paths
 - *Simplest*
 - *Popular*
 - *Critical*
 - *Complex*
 - *Pathological*
 - *Challenging*
 - *Error Handling*
 - *Periodic*



- James Bach, from our Rapid Software Testing course.

Incremental Catalogs

- As ideas occur to me, I might reserve a single page or two to consolidate them.



An Ongoing Bug Catalog

BUGS IN MY CELL PHONE NOV 12

- WHEN MP3 PLAYER IS RUNNING, NO OTHER FUNCTIONS AVAILABLE - I HAVE TO QUIT THE PLAYER TO DO ANYTHING
 - THERE'S NO FAST-FORWARD; ONLY SKIP TRACK (WHY DOES THIS MATTER? SOME PEOPLE LISTEN TO PODCASTS THAT ARE LONG)
 - VOICE RECOGNITION SAYS "EASE SAY THE NAME"
 - WHEN SETTING ALARM, IF THE CURRENT TIME IS 01:34, PRESSING "02" CHANGES THE ALARM-SET TIME TO 10:; THEN SKIPS OUT OF THE FIELD. IF THE TIME SHOWING IS 02:34, PRESSING 02 ACCEPTS BOTH DIGITS, THEN FOCUS MOVES. INCONSISTENT
 - ALARM TONE APPEARS EDITABLE - I.E. IS NOT GRAYED OUT - BUT IS NOT EDITABLE HERE.
LG 8100
- SW VERSION CX81TE17
FW VERSION 4.1
PRL VERSION PRL 11055
BROWSER VERSION 6.2.3.2
- SHOULD DO A TEST COVERAGE OUTLINE & RISK LIST ONE OF THESE DAYS.
 - WHEN TAKING FILES OFF THE MEMORY CARD, THE TIME/DATE STAMP IS , NO MATTER WHAT THE ACTUAL DATE/TIME WAS

Portable Presentations!

- Easier than booting the laptop!

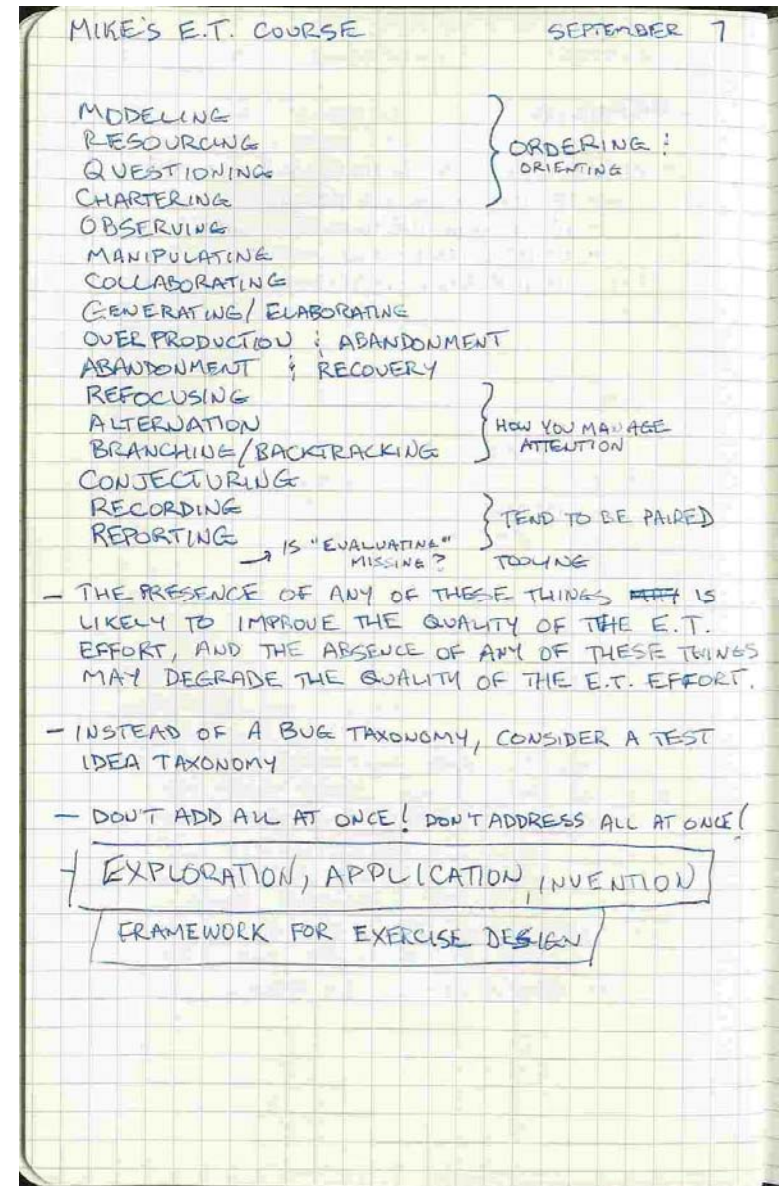
SUPPOSE EACH TEST ACTIVITY TAKES 2 MINUTES
SUPPOSE EACH BUG COSTS 10 MINUTES - 2 FOR THE TEST,
PLUS 8 MINUTES FOR INVESTIGATION & REPORTING

| BUGS FOUND | TEST DESIGN & EXECUTION | BUG INVESTIGATION & REPORTING | TEST ACTIVITIES |
|---------------|----------------------------|----------------------------------|--------------------|
| 0 | 90 mins (45 TESTS) | 0 mins | 45 |
| 1 | 80 mins (40 TESTS) | 10 min (1 TEST) | 41 |
| 8 | 10 mins (5 TESTS) | 80 min (8 TESTS) | 13 |

WHEN MORE BUGS ARE FOUND, TESTING SLOWS
OR COVERAGE WORSENS OR BOTH

E.T. Skills and Tactics

- Mike Kelly elaborated on this list of exploratory skills and tactics, which was originally written by James and Jon Bach.
- In writing down the list, I reckoned that tooling (distinct from resourcing) and evaluating were (for me) missing.



KEY IDEA



How do you record your work?

**Use concise, modular documents
that help tell the testing story.**

E.T. Notetaking Online: Session-Based Test Management



- Charter
 - A clear, concise mission for a session of testing
- Time Box
 - 90-minute (+/- 30), long enough for setup and focused work; short enough to make sure things don't get off track
- Reviewable Result
 - next slide!
- Debriefing
 - conversation between tester and manager
 - problems, bugs and issues can be discussed
 - new risks can be identified
 - coaching and mentoring can happen

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

Short: 60 minutes (+-15)

Normal: 90 minutes (+-15)

Long: 120 minutes (+-15)

- 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.

Reviewable Results



A test session sheet that can be scanned by a Perl script for compilation elsewhere

- 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

```
CHARTER
-----
Analyze MapMaker's View menu functionality and
report on areas of potential risk.

#AREAS
OS | Windows 2000
Menu | View
Strategy | Function Testing
Strategy | Functional Analysis

START
-----
5/30/00 03:20 pm

TESTER
-----
Jonathan Bach

TASK BREAKDOWN
-----

#DURATION
short

#TEST DESIGN AND EXECUTION
65

#BUG INVESTIGATION AND REPORTING
25

#SESSION SETUP
20
```

Debriefing



Assessment begins with observation

- The manager or test lead reviews the session sheet to assure that (s)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 and mentoring happens.

Session-Based Test Management



- For more information on SBTM, see <http://www.satisfice.com/sbtm>

KEY IDEA



How do you effectively report your work?

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

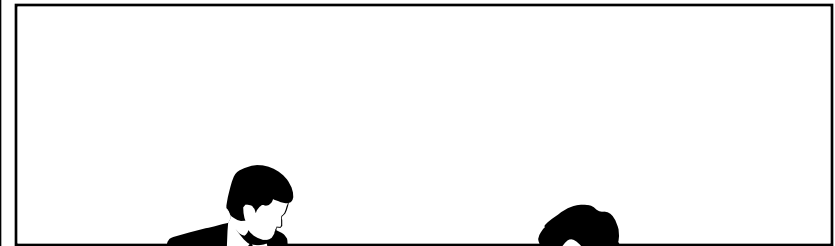
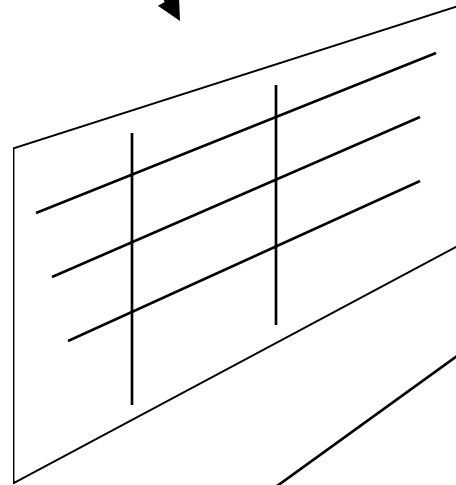
The Dashboard Concept



Large dedicated whiteboard
“Do Not Erase”

Project conference room










Project status meeting



Testing Dashboard

Updated:
2/21

Build:
38

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

Product Area



| Area |
|----------------------|
| file/edit |
| view |
| insert |
| format |
| tools |
| slideshow |
| online help |
| clipart |
| converters |
| install |
| compatibility |
| general GUI |

- 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.

Test Effort



| | |
|----------------|--------------------------------------------------------|
| None | Not testing; not planning to test. |
| Start | No testing yet, but expect to start soon. |
| Low | Regression or spot testing only; maintaining coverage. |
| High | Focused testing effort; increasing coverage. |
| Pause | Temporarily ceased testing, though area is testable. |
| Blocked | Can't effectively test, due to blocking problem. |
| Ship | Going through final tests and signoff procedure. |

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



| | |
|-----------|--------------------------------------------------------------------------------------------------|
| 0 | We don't have good information about this area. |
| 1 | Sanity Check: major functions & simple data. |
| 1+ | More than sanity, but many functions not tested. |
| 2 | Common & Critical: all functions touched; common & critical tests executed. |
| 2+ | Some data, state, or error coverage beyond level 2. |
| 3 | Complex Cases: strong data, state, extreme, exceptional, error, stress or long-sequence testing. |

Test Coverage

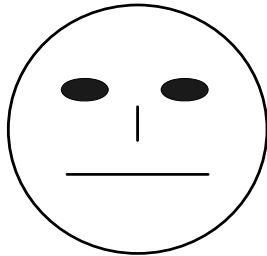


- Color green if coverage level is acceptable for ship, otherwise color black.
- Level 1 and 2 focus on functional requirements and capabilities: *can* this product work at all?
- Level 2 may span 50%-90% code coverage.
- Level 2+ and 3 focus on information to judge performance, reliability, compatibility, 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.”

Quality Assessment



“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.”



“We know of problems that are possible showstoppers, or we suspect that there are important problems not yet discovered.”



“We know of problems in this area that definitely stop ship or interrupt testing.”

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???



“We shape our tools;
thereafter they shape us.”

-- Marshall McLuhan