Critical Thinking for Testers

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

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

Why Don't People Think Well?

Consider this problem:

A bat and ball cost $1.10.
The bat costs one dollar more than the ball.
How much does the ball cost?

Why Don't People Think Well?

"Steve is very shy and withdrawn, invariably helpful but with little interest in people or in the world of reality. A meek and tidy soul, he has a need for order and structure, and a passion for detail."

Is Steve more likely to be

a librarian?

a farmer?

Do you prefer A or B?

Imagine that the US is preparing for the outbreak of an unusual Asian disease, which is expected to kill 600 people. Two alternative programs to combat the disease have been proposed. Assume that the exact scientific estimates of the consequences of the programs are as follows.

Program A: If Program A is adopted, 200 people will be saved.

Program B: If Program B is adopted, there is 1/3 probability that 600 people will be saved, and 2/3 probability that no people will be saved.

Do you prefer C or D?

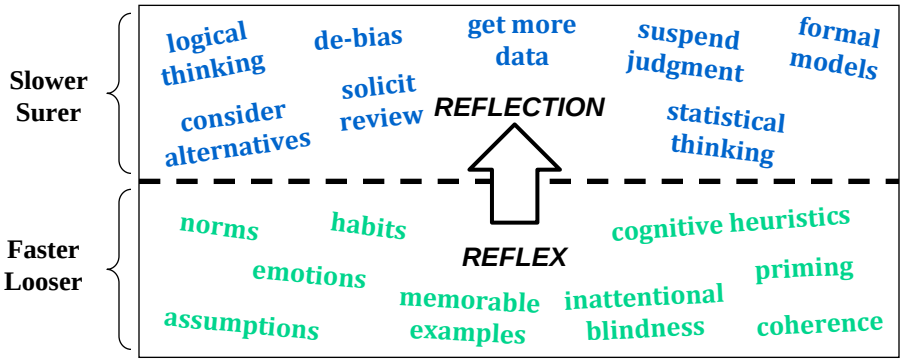
Imagine two more programs to combat the disease are proposed:

Program C: If Program C is adopted, 400 people will die.

Program D: If Program D is adopted, there is 1/3 probability that 600 people will be saved, and 2/3 probability that no people will be saved.

Reflex is IMPORTANT
But Critical Thinking is About Reflection

System 2



System 1

See Thinking Fast and Slow, by Daniel Kahneman

Themes

- Technology consists of complex and ephemeral relationships that can seem simple, fixed, objective, and dependable even when they aren't.
- Testers are people who ponder and probe complexity.
- Basic testing is a straightforward technical process.
- But, *excellent* testing is a difficult social and psychological process in addition to the technical stuff.

The Nature of Critical Thinking

- "Critical thinking is **purposeful, self-regulatory judgment** which results in interpretation, analysis, evaluation, and inference, as well as explanation of the evidential, conceptual, methodological, criteriological, or contextual considerations upon which that judgment is based." - *Critical Thinking: A Statement of Expert Consensus for Purposes of Educational Assessment and Instruction, Dr. Peter Facione*

(Critical thinking is, for the most part, about getting all the benefits of your "System 1" thinking reflexes while avoiding self-deception and other mistakes.)

Bolton's Definition of Critical Thinking

Critical Thinking
is thinking about thinking
with the aim of not getting fooled.

- Michael Bolton

The Nature of Critical Thinking

- We call it critical thinking whenever we systematically doubt something that the "signs" tell us is probably true. Working through the doubt gives us a better foundation for our beliefs.
- Critical thinking is a kind of **de-focusing** tactic, because it requires you to seek alternatives to what is already believed or what is being claimed.
- Critical thinking is also a kind of **focusing** tactic, because it requires you to analyze the specific reasoning behind beliefs and claims.

Why You Should Care

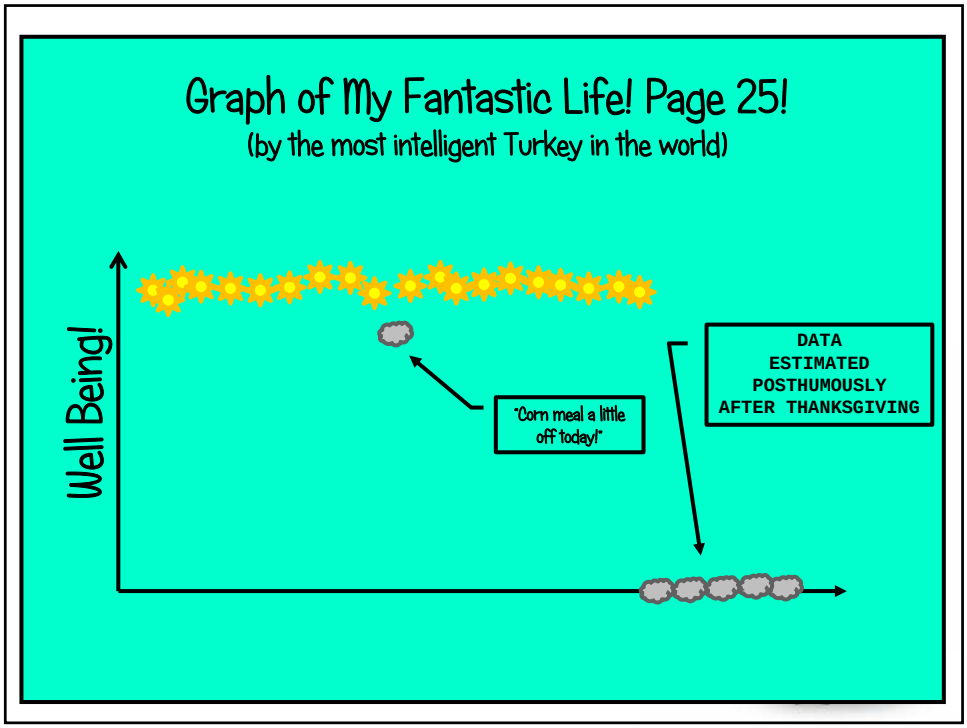
Technology is way more tricky than regular life.

But testers are not supposed to get tricked.

The Big Theme of This Workshop

A tester is someone who knows that things can be different.

- Jerry Weinberg



Don't Be A Turkey

- No experience of the past can LOGICALLY be projected into the future, because we have no experience OF the future.
- No big deal in a world of stable, simple patterns.
- **BUT SOFTWARE IS NOT STABLE OR SIMPLE.**
- **"PASSING" TESTS CANNOT PROVE SOFTWARE GOOD.**

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

How Do We Know What "Is"?

"We know what is because we see what is."

Let's be more specific...

We believe
we know what is because we see
what we interpret as signs that indicate
what is
based on our prior beliefs about the world.

How Do We Know What "Is"?

We see the signs!

"If I see X, then probably Y, because probably A, B, C, D, etc."

- THIS CAN FAIL:
  - Getting into a car– oops, not my car.
  - Bad driving– Why?
  - Bad work– Why?
  - Ignored people at my going away party– Why?
  - Couldn't find soap dispenser in restroom– Why?
  - Ordered orange juice at seafood restaurant– waitress misunderstood

Remember this, you testers!

What you see is NOT all there is.

Exercise: Calculator Test

"I was carrying a calculator.

I **dropped** it!

Perhaps it is **damaged**!

What might you do to test it?"

What prevents us
from asking questions?

What is an assumption?

What makes an assumption more dangerous?

1. **Foundational:** required to support critical plans and activities. (Changing the assumption would change important behavior.)
2. **Unlikely:** may conflict with other assumptions or evidence that you have. (The assumption is counter-intuitive, confusing, obsolete, or has a low probability of being true.)
3. **Blind:** regards a matter about which you have no evidence whatsoever.
4. **Controversial:** may conflict with assumptions or evidence held by others. (The assumption ignores controversy.)
5. **Impolitic:** expected to be declared, by social convention. (Failing to disclose the assumption violates law or local custom.)
6. **Volatile:** regards a matter that is subject to sudden or extreme change. (The assumption may be invalidated unexpectedly.)
7. **Unsustainable:** may be hard to maintain over a long period of time. (The assumption must be stable.)
8. **Premature:** regards a matter about which you don't yet need to assume.
9. **Narcotic:** any assumption that comes packaged with assurances of its own safety.
10. **Latent:** Otherwise critical assumptions that we have not yet identified and dealt with. (The act of managing assumptions can make them less critical.)

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

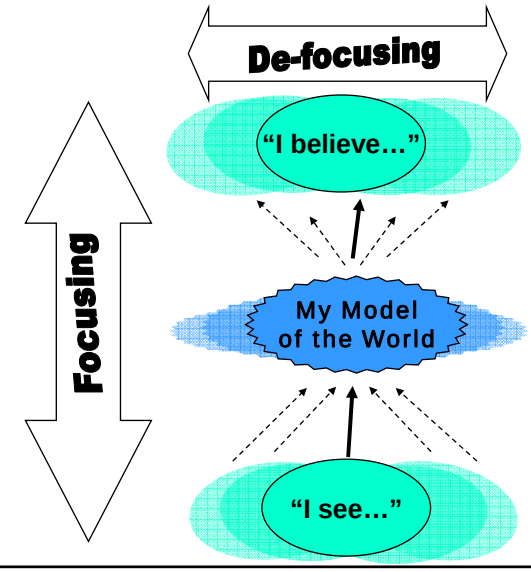
- **…that represents another idea, activity, or object…**

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

- **…whereby understanding the model may help you understand or manipulate what it represents.**

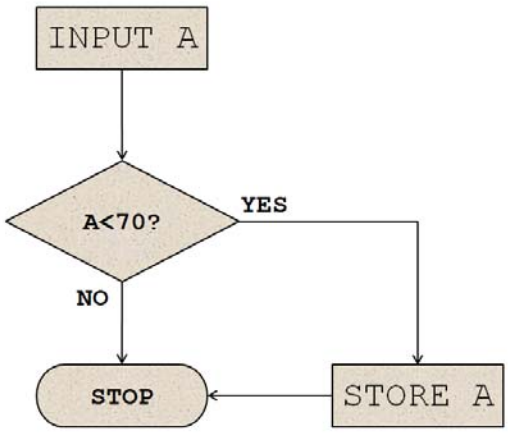
- A map helps navigate across a terrain.
- 2+2=4 is a model for adding two apples to a basket that already has two apples.
- Atmospheric models help predict where hurricanes will go.
- A fashion model helps understand how clothing would look on actual humans.
- Your beliefs about what you test are a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

Models Link Observation & Inference



- Testers must distinguish observation from inference!
- Our mental models form the link between them
- Defocusing is lateral thinking.
- Focusing is logical (or "vertical") thinking.

How many test case are needed to test the product represented by this flowchart?



Testing against requirements is all about modeling.

How do you test this?

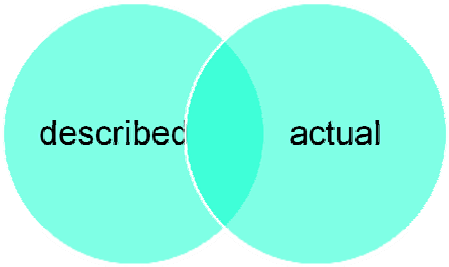
"The system shall operate at an input voltage range of nominal 100 - 250 VAC."

Poor answer:

"Try it with an input voltage in the range of 100-250."

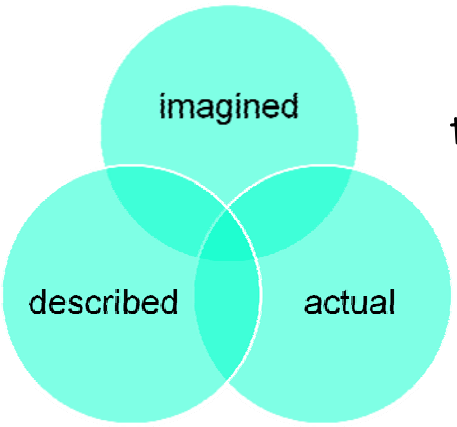
This is what people think you do

"Compare the product to its specification"



This is more like what you really do

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



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

"Compare the actual product to a description of it"

This is what you find…

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

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

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

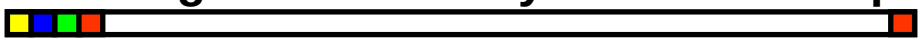
The designer **INTENDS** the product to be Firefox compatible, **SAYS SO**, and **IT IS**.

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

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

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

How to Think Critically: Slowing down to let System 2 wake up



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)

So?

- The truth may not matter, or may matter much more than you think. (poor understanding of risk)

To What Do We Apply Critical Thinking?

Huh?

Really?

So?

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

"Huh?"
Critical Thinking About Words

- 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."

Example: 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."

Example: Generating Interpretations

- Selectively emphasize each word in a statement; also consider alternative meanings.

MARY had a little lamb.
Mary **HAD** a little lamb.
Mary had **A** little lamb.
Mary had a **LITTLE** lamb.
Mary had a little **LAMB**.

"Really?" Critical Thinking About Interpretation

What did you see or hear (smell, taste, touch) that made you believe...?

"So?"
Critical Thinking About Risk

"When the user presses a button on the touchscreen, the system shall respond within 300 milliseconds."

How would you test this?

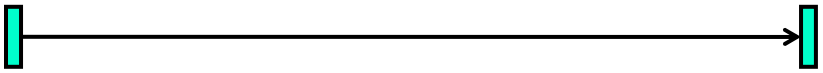
Heuristic Model:
The Four-Part Risk Story

Someone may be hurt or annoyed
because of something that might go wrong while operating the product,
due to some vulnerability in the product
that is exploited by some threat.

- **Victim.** Someone that experiences the impact of a problem. Ultimately no bug can be important unless it victimizes a human.
- **Problem:** Something the product does that we wish it wouldn't do.
- **Vulnerability:** Something about the product that causes or allows it to exhibit a problem, under certain conditions.
- **Threat:** Some condition or input external to the product that, were it to occur, would trigger a problem in a vulnerable product.

Critical Thinking About Projects

- You will have five weeks to test the product:



5 weeks

Safety Language ("epistemic modality")

- A precise, circumspect style of speaking and writing, intended to clarify the difference between observation and inference
- Informed by a determination to suspend conclusions, certainty, and judgment
  - All conclusions are conclusions *for now*
  - Certainty isn't available
  - Judgment is always uncertain, and decisions about quality are based on politics and emotions.
- Emphasizes appropriate subjectivity
- *A form of tester self-defense*

Why Use Safety Language?

- Helps to defend credibility and reputation
- Precision and accuracy for our clients
- Requires and helps to sharpen critical thinking
- A qualifier circles back to you and changes your thinking.
- Helps to prevent critical thinking errors
- Fundamental attribution error
- Cause-and-effect correlation
- Lumping errors (assimilation bias)
- Confirmation bias

The logical language of *test framing* is a form of safety language. Words like “if”, “or”, “else”, “unless”, and so forth establish context and preserve appropriate levels of uncertainty.

See http://www.developsense.com/blog/2010/09/test-framing/

Risks With Safety Language

- To some, it sounds non-committal.
- Done well, it prohibits you from being pinned down, which some people will want to do.
- Places responsibility for decisions in the hands of those who should be making them; many find this uncomfortable.
- When you use safety language, you are sending a social message that may have political and emotional overtones.
- Skillful use of safety language depends on knowing when *not* to use it.

Some Verbal Heuristics: "A vs. THE"

Whatever is making a difference is probably not the only thing.

- Example: "A problem…" instead of "THE problem…"
- Using "A" instead of "THE" helps us to avoid several kinds of critical thinking errors
  - single path of causation
  - confusing correlation and causation
  - single level of explanation

Some Verbal Heuristics: "Unless…"

Whatever is true sometimes may not be true always.

- When someone asks a question based on a false or incomplete premise, try adding "unless…" to the premise
- When someone offers a Grand Truth about testing, try appending "unless…" or "except in the case of…"—or try countering with "What if..?"

Some Verbal Heuristics: "And Also…"

Whatever is happening, something else may ALSO be happening.

- The product gives the correct result! Yay!
- …It also may be silently deleting system files.

Some Verbal Heuristics: "So far" and "Not yet"

Whatever is true now may not be true for long.

- The product works… so far.
- We haven't seen it fail… yet.
- No customer has complained… yet.
- Remember: There is no test for ALWAYS.

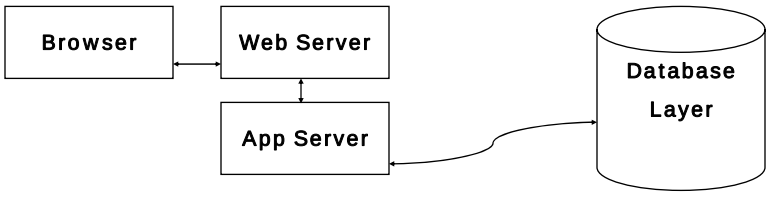
Some Verbal Heuristics: "Compared to what?"

Whatever you're comparing, you need something to compare it to.

- Software is too expensive!
- Testing is taking too long!
- We don't have enough information!

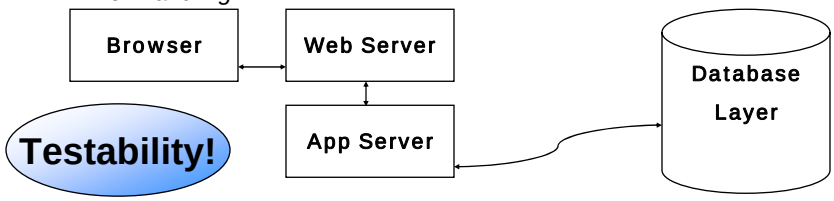
Critical Thinking About Diagrams *Analysis*

- [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?*

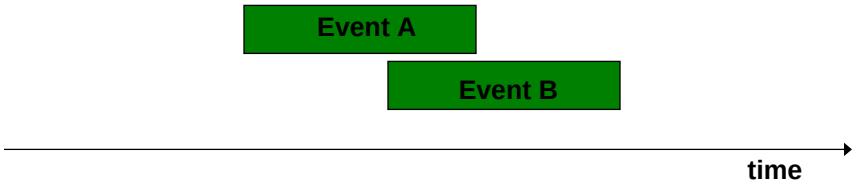


Guideword Heuristics for Diagram Analysis

- Boxes
  - *Interfaces (testable)*
  - *Missing/Drop-out*
  - *Extra/Interfering/Transient*
  - *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*



Critical Thinking About Timing



- You want to test the interaction between two potentially overlapping events.
- What would you do to test them?

Critical Thinking About Practices: What does "best practice" mean?

Someone believes you might suffer unless you do this practice

- **Someone:** Who is it? What do they know?
- **Believes:** What specifically is the basis of their belief?
- **You:** Is *their* belief applicable to *you*?
- **Might:** How *likely* is the suffering to occur?
- **Suffer:** So what? Maybe it's worth it?
- **Unless:** Really? There's no alternative?
- **You do this practice:** What does it mean to "do" it? What does it cost? What are the side effects? What if you do it badly? What if you do something else really well?

Beware of…

- **Numbers:** "We cut test time by 94%."
- **Documentation:** "You must have a written plan."
- **Judgments:** "That project was *chaotic*. This project was a *success*."
- **Behavior Claims:** "Our testers follow test plans."
- **Terminology:** Exactly what is a "test plan?"
- **Contempt for Current Practice:** CMM Level 1 (initial) vs. CMM level 2 (repeatable)
- **Unqualified Claims:** "A subjective and unquantifiable requirement is not testable."

Look For…

- **Context:** "This practice is useful when you want the power of creative testing but you need high accountability, too."
- **People:** "The test manager must be enthusiastic and a real hands-on leader or this won't work very well."
- **Skill:** "This practice requires the ability to tell a complete story about testing: coverage, techniques, and evaluation methods."
- **Learning Curve:** "It took a good three months for the testers to get good at producing test session reports."
- **Caveats:** "The metrics are useless unless the test manager holds daily debriefings."
- **Alternatives:** "If you don't need the metrics, you ditch the daily debriefings and the specifically formatted reports."
- **Agendas:** "I run a testing business, specializing in exploratory testing."

Some Common Beliefs About Testing Apply some critical thinking!

- Every test must have an expected, predicted result.
- Effective testing requires complete, clear, consistent, and unambiguous specifications.
- Bugs found earlier cost less to fix than bugs found later.
- Testers are the quality gatekeepers for a product.
- Repeated tests are fundamentally more valuable.
- You can't manage what you can't measure.
- Testing at boundary values is the best way to find bugs.

Some Common Beliefs About Testing
Apply some critical thinking!

- Test documentation is needed to deflect legal liability.
- The more bugs testers find before release, the better the testing effort.
- Rigorous planning is essential for good testing.
- Exploratory testing is unstructured testing, and is therefore unreliable.
- Adopting best practices will guarantee that we do a good job of testing.
- Step by step instructions are necessary to make testing a repeatable process.

Some Common Thinking Errors

- Reification Error
  - giving a name to a concept, and then believing it has an objective existence in the world
  - ascribing material attributes to mental constructs—"that product has quality"
  - mistaking relationships for things—"its purpose **is**…"
  - purpose and quality are *relationships*, not attributes; they depend on the person
  - how can we count ideas? how can we quantify relationships?

MY NOMINEE FOR TESTING'S BIGGEST PROBLEM!

Some Common Thinking Errors

- Fundamental Attribution Error
  - “it always works that way”; “he’s a jerk”
  - failure to recognize that circumstance and context play a part in behaviour and effects
- The Similarity-Uniqueness Paradox
  - “all companies are like ours”; “no companies are like ours”
  - failure to consider that *everything* incorporates similarities *and* differences
- Missing multiple paths of causation
  - “A causes B” (even though C and D are also required)

Some Common Thinking Errors

- Assuming that effects are linear with causes
  - “If we have 20% more traffic, throughput will slow by 20%”
  - this kind of error ignores non-linearity and feedback loops—c.f. general systems
- Reactivity Bias
  - the act of observing affects the observed
  - a.k.a. “Heisenbugs”, the Hawthorne Effect
- The Probabilistic Fallacy
  - confusing unpredictability and randomness
  - after the third hurricane hits Florida, is it time to relax?

Some Common Thinking Errors

- Binary Thinking Error / False Dilemmas
  - “all manual tests are bad”; “that idea never works”
  - failure to consider gray areas; belief that something is either entirely something or entirely not
- Unidirectional Thinking
  - expresses itself in testing as a belief that “the application works”
  - failure to consider the opposite: what if the application *fails*?
  - to find problems, we need to be able to imagine that they *might* exist

Some Common Thinking Errors

- Availability Bias
  - the tendency to favor prominent or vivid instances in making a decision or evaluation
  - example: people are afraid to fly, yet automobiles are far more dangerous per passenger mile
  - to a tech support person (or to some testers), the product always seems completely broken
  - spectacular failures often get more attention than grinding little bugs
- Confusing concurrence with correlation
  - “A and B happen at the same time; they must be related”

Some Common Thinking Errors

- Nominal Fallacies
  - believing that we know something well because we can name it
    - “equivalence classes”
  - believing that we don’t know something because we don’t have a name for it at our fingertips
    - “the principle of concomitant variation”; “inattentional blindness”
- Evaluative Bias of Language
  - failure to recognize the spin of word choices
  - …or an attempt to game it
  - “our product is full-featured; theirs is bloated”

Some Common Thinking Errors

- Selectivity Bias
  - choosing data (beforehand) that fits your preconceptions or mission
  - ignoring data that doesn’t fit
- Assimilation Bias
  - modifying the data or observation (afterwards) to fit the model
  - grouping distinct things under one conceptual umbrella
  - Jerry Weinberg refers to this as “lumping”
  - for testers, the risk is in identifying setup, pinpointing, investigating, reporting, and fixing as “testing”

Some Common Thinking Errors

- Narrative Bias
  - a.k.a “post hoc, ergo propter hoc”
  - explaining causation after the facts are in
- The Ludic Fallacy
  - confusing complex human activities with random, roll-of-the-dice games
  - “Our project has a two-in-three chance of success”
- Confusing correlation with causation
  - “When I change A, B changes; therefore A must be causing B”

Some Common Thinking Errors

- Automation bias
  - people have a tendency to believe in results from an automated process out of all proportion to validity
- Formatting bias
  - Things are more credible when they’re on a nicely formatted spreadsheet or document
- Survivorship bias
  - we record and remember results from projects (or people) who survived
  - “The sailors survived because they prayed to Neptune.”
    - What about the sailors who prayed and died anyway?
  - “The bug rate for our successful projects was 0.2%”
    - What was the bug rate for projects that were cancelled?

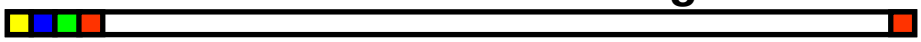
Books on Testing and Critical Thinking

- Baron, Jonathan. Thinking and Deciding. 4th ed. Cambridge University Press, 2007.
- Collins, Harry M., and Martin Kusch. The Shape of Actions: What Humans and Machines Can Do. The MIT Press, 1999.
- Collins, H. M, and T. J Pinch. The Golem?: What Everyone Should Know About Science. Cambridge [England]; New York, NY, USA: Cambridge University Press, 1994.
- Collins, Harry. Tacit and Explicit Knowledge. University Of Chicago Press, 2010.
- Friedl, Jeffrey E.F. Mastering Regular Expressions. Third ed. O'Reilly Media, 2006.
- Gigerenzer, Gerd. Gut Feelings: The Intelligence of the Unconscious. Reprint. Penguin (Non-Classics), 2008.
- Gigerenzer, Gerd, Peter M. Todd, and ABC Research Group. Simple Heuristics That Make Us Smart. 1st ed. Oxford University Press, USA, 2000.

Books on Testing and Critical Thinking

- Gladwell, Malcolm. Blink: The Power of Thinking Without Thinking. Back Bay Books, 2007.
- Huff, Darrell. How to Lie with Statistics. W. W. Norton & Company, 1993.
- Kahneman, Daniel. Thinking, Fast and Slow. Penguin, 2011.
- Kahneman, Daniel, Paul Slovic, and Amos Tversky, eds. Judgment Under Uncertainty: Heuristics and Biases. 1st ed. Cambridge University Press, 1982.
- Kaner, Cem, James Bach, and Bret Pettichord. Lessons Learned in Software Testing. 1st ed. Wiley, 2001.
- Kaner, Cem, and Walter P. Bond. 'Software Engineering Metrics: What Do They Measure and How Do We Know?', 2004. http://www.kaner.com/pdfs/metrics2004.pdf.
- Kirk, Jerome, and Marc L Miller. Reliability and Validity in Qualitative Research. Beverly Hills: Sage Publications, 1986. http://www.amazon.com/Reliability-Validity-Qualitative-Research-Methods/dp/0803924704.

Books on Testing and Critical Thinking



- Levy, David A. Tools of Critical Thinking: Metathoughts for Psychology. 2nd ed. Waveland Pr Inc, 2009.
- Popper, Karl. Conjectures and Refutations: The Growth of Scientific Knowledge. 2nd ed. Routledge, 2002.
- Schneier, Bruce. ‘Schneier on Security: Teaching the Security Mindset’, n.d. http://www.schneier.com/blog/archives/2012/06/teaching_the_se.html.
- Simon, Herbert A. The Sciences of the Artificial - 3rd Edition. third ed. The MIT Press, 1996.
- Taleb, Nassim Nicholas. The Black Swan: Second Edition: The Impact of the Highly Improbable: With a New Section: ‘On Robustness and Fragility’. 2nd ed. Random House Trade Paperbacks, 2010.
- Watts, Duncan J. Everything Is Obvious: *Once You Know the Answer. Crown Business, 2011.
- Weinberg, Gerald M. An Introduction to General Systems Thinking. 25 Anv. Dorset House, 2001.
- ———. Perfect Software and Other Illusions About Testing. Dorset House, 2008.
- Weinberg, Gerald M., and Daniela Weinberg. General Principles of Systems Design. Dorset House, 1988.