You are a scientist

Embracing the Scientific Method in Software Testing

Christin Wiedemann, Ph.D.

Professional Quality Assurance (PQA) Ltd.
Personal Agenda

• Start a Scientific Testing Revolution

  ‣ Or at least promote skilled, professional testing
  ‣ Contribute to improving reputation and status of test
  ‣ Increase test credibility
  ‣ Boost testers’ self-confidence
• This talk is based on my personal interpretation of “science” and “testing”

› It’s not an objective account of facts
› It’s not the whole story – just the juicy parts
› Not trying to teach, but attempting to inspire
...to me, science is primarily physics...
...to me, “testing” equals “exploring” and “learning”...
The Scientific Method

Big Words:
- Deduction, Induction, Hypothesis, Theory

Trial & Error

Empirical Falsifiability

Credibility

Critical thinking

Using the Scientific Method

Revision

Philosophy of Science
• “There is no such thing as the scientific method.”
  ‣ J. B. Conant, 1947

• There is no such thing as the way to test
What is the Scientific Method?

• The Scientific Method
  ‣ “Scientific method refers to a body of techniques for investigating phenomena, acquiring new knowledge, or correcting and integrating previous knowledge.”

The Scientific Method

- "Scientific method refers to a body of techniques for investigating phenomena, acquiring new knowledge, or correcting and integrating previous knowledge."

The Scientific Method

- Test
- Test
- Regression test / Retest
- Design tests
The Scientific Method

• Collection of techniques for acquiring knowledge

  ‣ In order to be considered scientific, a method must be based on gathering *empirical* and *measurable* evidence

  ‣ Empirical means acquired through *observation* or *experimentation*

  ‣ Measurable means that the *physical properties* can be *quantified*
• Uh…what?

  ‣ Unless your claims are based on observations you have made, scientists will sneer at you

  ‣ Making things up without experimental evidence is using science fiction, not the scientific method
• Back to testing, please

  ‣ To test is to *experiment* and *observe* the results

  ‣ Test results are *empirical* evidence

  ‣ To be *testable*, it must be *measurable*

  ‣ Testing is the gathering of empirical and measurable evidence

  ‣ *Good testing adheres to the scientific method*
Basic concepts

• Introducing concepts

  ‣ **Deduction**: The process drawing a specific conclusion from general statements.

  ‣ **Induction**: The process of drawing general conclusions based on specific examples.
Basic concepts

• Deduction

1. All Swedes are blonde
2. Christin is Swedish
3. Therefore, Christin is blonde
Basic concepts

- Deduction

1. All Swedes are blonde
2. Christin is Swedish
3. Therefore, Christin is blonde

General statement

Specific example
Basic concepts

- Deduction

```
Class: Swedes
Attribute: Blonde
```

Christin
• **Induction**

1. All Swedes I have ever met are blonde
2. Therefore, all Swedes are blonde
Basic concepts

• Induction

1. All Swedes I have ever met are blonde
2. Therefore, all Swedes are blonde

The conclusion can be false even though the premises are true!
• Deduction
  1. All software malfunctions when it runs out of memory
  2. Notepad is a software
  3. Therefore, Notepad malfunctions when it runs out of memory

• Induction
  1. All software I have ever used has malfunctioned when it has run out of memory
  2. Therefore, all software malfunctions when it runs out of memory
Basic concepts

• Deduction
  1. All software malfunctions when it runs out of memory
  2. Notepad is a software
  3. Therefore, Notepad malfunctions when it runs out of memory

• Induction
  1. All software I have ever used has malfunctioned when it has run out of memory
  2. Therefore, all software malfunctions when it runs out of memory
Basic concepts

• Hypothesis
  ‣ Proposed explanation for phenomenon

• Theory
  ‣ Set of principles that explain and predict phenomena

1. Create a hypothesis by proposing an explanation
2. Test the accuracy of the hypothesis through observation and experiments
3. After corroboration, a hypothesis can be accepted as a theory
What is Science?

- An attempt to explain the many phenomena of nature
- Search for unity, simplicity, order
- An open, free-thinking, public activity
- A way to knowledge
- Testable knowledge
The Scientific Method

• Hypothetico-Deductive Model

1. Collect data through observation
2. Suggest a hypothesis
3. Predict a consequence of the hypothesis
4. Experiments to corroborate hypothesis

- If experiments confirm the hypothesis, go back to step 2.
- If experiments do not confirm the hypothesis, the hypothesis is rejected.
The Scientific Method

- Hypothetico-Deductive Model

1. Collect data through observation
2. Suggest a hypothesis
3. Predict a consequence of the hypothesis
4. Experiments to corroborate hypothesis

- Yes
- No

Observe mouse eating cheddar cheese
Mice like *all* sorts of cheese
Mice will eat Swiss cheese too
Give mouse Swiss cheese

Examples:
- Observe mouse eating cheddar cheese
- Give mouse Swiss cheese

Mice like *all* sorts of cheese
The Scientific Method

• Hypothetico-Deductive Model

Collect data through observation

Suggest a hypothesis

Predict a consequence of the hypothesis

Experiments to corroborate hypothesis

Yes

No

Mice like goat cheese

Give mouse goat cheese
The Scientific Method

1. Collect data through observation
2. Suggest a hypothesis
3. Predict a consequence of the hypothesis
4. Experiments to corroborate hypothesis

- Yes
  - Mice like cheese with no holes in it
- No
  - Mice will eat Colby cheese too
  - Give mouse Colby cheese

Suggest a hypothesis: Mice will eat Colby cheese too
Collect data through observation: Mice like cheese with no holes in it
Experiments to corroborate hypothesis: Give mouse Colby cheese
Predict a consequence of the hypothesis: Mice will eat Colby cheese too
The Scientific Method

• Hypothetico-Deductive Model

1. Collect data through observation
2. Suggest a hypothesis
3. Predict a consequence of the hypothesis
4. Experiments to corroborate hypothesis
5. Learn product
6. Identify set of risks
7. Predict failure
8. Execute test to reveal failure

Flowchart:
- Yes
- No

Christin Wiedemann – PQA
The Scientific Method

• No Checklists
  ‣ Scientists don’t conduct research using a checklist
  ‣ The method leaves plenty of room for interpretation and modification
  ‣ The Scientific Method is not A Law
The Scientific Method

• “The Scientific Method: A heuristic path; a suggestive, but not rigorously followed.”
  ‣ Bob Widing
• Everyone is Welcome!
  ‣ Being human means unconsciously using the scientific method on a daily basis

  ‣ In its simplest form it can be seen as trial & error
• Not an Exclusive Club
  ‣ The core of the scientific method is asking *questions* and trying to find *answers*
  ‣ Not only for scientists!
Why the Scientific Method?

- If it's so natural and common, why is it still called the scientific method?
- Scientists apply it more rigorously than non-scientists
- Scientists strive to
  - Avoid bias
  - Be objective
  - Use a standardized approach*
- Shouldn't testers too?

* Not the same thing as a "best practice"!
- The Scientific Method is cyclic
  - Simplified: common uses of the Scientific Method involves trial and error
• **Trial & Error**
  ‣ Make a guess
  ‣ Try it
  ‣ If the initial guess is incorrect, revise it
  ‣ Try again
  ‣ Repeat

SEE, I’VE GOT A REALLY GOOD SYSTEM: IF I WANT TO SEND A YOUTUBE VIDEO TO SOMEONE, I GO TO FILE→SAVE, THEN IMPORT THE SAVED PAGE INTO WORD. THEN I GO TO “SHARE THIS DOCUMENT” AND UNDER “RECIPIENT” I PUT THE EMAIL OF THIS VIDEO EXTRACTION SERVICE...

I’LL OFTEN ENCOURAGE RELATIVES TO TRY TO SOLVE COMPUTER PROBLEMS THEMSELVES BY TRIAL AND ERROR. HOWEVER, I’VE LEARNED AN IMPORTANT LESSON: IF THEY SAY THEY’VE SOLVED THEIR PROBLEM, NEVER ASK HOW.
• **Trial & Error**
  ‣ Make a guess – *Propose a hypothesis and make a prediction*
  ‣ Try it – *Run an experiment and evaluate result*
  ‣ If the initial guess is incorrect, revise it – *Make changes to hypothesis*
  ‣ Try again – *Run new experiment and evaluate new result*
  ‣ Repeat
• Example

1. Light doesn't turn on when flicking the switch.
2. It’s probably the bulb that is broken.
3. Change the bulb.
4. Flick the switch.
5. Light still doesn’t come on!
6. Go back to step 2 and repeat until light comes on.
Trial & Error – Example

- Hypothetico-Deductive Model

1. Light doesn't turn on when flicking the switch.
2. It’s probably the bulb that is broken.  
   - Hypothesis
3. Change the bulb.  
   - Prediction
4. Flick the switch.  
5. Light still doesn’t come on!  
   - Evaluation
6. Go back to step 2 and repeat until light comes on.
• Trial & Error

  ‣ Make a guess – Propose risk and predict potential failure
  ‣ Try it – Execute test and evaluate result
  ‣ If the initial guess is incorrect, revise it – No bugs? Revise risks and corresponding failures
  ‣ Try again – Execute new tests and evaluate results
  ‣ Repeat
The Objective of Science

• The scientific method involves
  ‣ Making observations
  ‣ Collecting data

• Those activities are just steps on the way, though

• The *objective* is to analyze the data to *learn* more about, and better understand, the world around us
The Objective of Testing

• In testing we
  ‣ Execute tests and observe how the software responds
  ‣ Record and store test results

• However, the goal of testing is not the execution or the artifacts

• But what learn about the product
Empirical Falsifiability

• Contradicting a hypothesis through observation or outcome of experiment

  ‣ Applying Empirical Falsifiability: Finding ways to show that the hypothesis is wrong
• Contradicting a hypothesis through observation or outcome of experiment

  ‣ Scientists propose a hypothesis:
     • Rocks float on water

  ‣ Scientists design experiments to test hypothesis:
     • Throw rock in water

  ‣ Scientists evaluate result of experiment:
     • Rock sank – the hypothesis was wrong
• Contradicting a hypothesis through observation of our outcome of experiment

  ‣ Testers propose a hypothesis:
    • The software works

  ‣ Testers design tests to...test the hypothesis:
    • How could the software possibly fail?

  ‣ Testers *evaluate* the result of the test:
    • The software failed – the hypothesis was wrong
Empirical Falsifiability

• “Science cannot discover truth, but it is an excellent means of discovering error.”
  ‣ Kenneth Boulding

• We can never prove something is true
  ‣ Only find evidence that *contradicts* the hypothesis
  ‣ We can never prove that software is *working* – only find proof of ways in which it *fails*
Empirical Falsifiability

- My research – Ultra-High Energy Neutrinos

  - Hypothesis: According to hypothesis A, the neutrino count at Earth from supernovae is $N$ per km$^2$

  - If I measure $N$ per km$^2$ I still cannot say that the hypothesis is correct

  - But I measured $M$ per km$^2$, $M<N$, and therefore I could say that probably the hypothesis is wrong

http://www.icecube.wisc.edu/
• Falsifying a theory is not that easy
  ‣ Design theories to be more easily falsifiable

• Finding bugs is not always easy either
  ‣ Design tests that are easy to evaluate
Ockham’s Razor

• “Entities should not be multiplied unnecessarily”
  ‣ Opt for the simplest explanation

• Used as a heuristic in science to guide in the development of theoretical models

• A simple theory is better testable

• A simple theory applies to more cases and is more easily falsifiable

*The simplest theory is not always correct!*
• **Example**

1. Light doesn't turn on when flicking the switch.
2. It could be the switch, the wiring, alien influences,…or maybe the bulb is broken.
3. Start with the simplest explanation – change the bulb.

It’s easy to falsify the hypothesis “the bulb is broken”

It’s more complicated to falsify the hypothesis “alien influences”
Characteristics of Science

- Credible
- Curiosity-driven
- Critical
- Impartial
- Dynamic
• How closely does the work adhere to principles such as the Scientific Method?

  ‣ Credibility assessed through peer review
• **Peers**
  
  ‣ People who are equal in some respect, e.g. same level of education and similar work tasks

• **Peer review**
  
  ‣ Scrutiny by peers
Test Credibility

• What renders test credibility?

  ‣ Who assesses credibility?
Test Credibility

• **Peer**
  - Fellow testers
  - Other team members

• **Peer review**
  - Of test ideas, test design, test reports
  - Improves quality of work
  - Gives credibility
  - Is a chance to learn from each other
• The goal of science is *knowledge*

• What is the goal of testing?

• The goal of testing **should** be *information*, i.e. knowledge
Critical Thinking

- Question assumptions
  - Sub-conscious assumptions are your enemy
  - Identify the assumptions you are making
  - Examine the assumptions
  - Discern hidden values
  - Consider the context
• Question assumptions

› Which triangle is larger – one with sides measuring 200, 300 and 400 cm or one with sides measuring 300, 400 and 700 cm?*

* Infuriating Lateral Thinking Puzzles, Paul Sloane & Des MacHale
• Question assumptions

- Which triangle is larger – one with sides measuring 200, 300 and 400 cm or one with sides measuring 300, 400 and 700 cm?*

- The first one! The second is a line and has area zero…

- Common assumption is that both suggestions are valid triangles.

* Infuriating Lateral Thinking Puzzles, Paul Sloane & Des MacHale
Critical Thinking

• How do we decide what to believe?

  ‣ What we know about the issue – evidence
  ‣ How compatible something is to our own world view
  ‣ Reliability of the source
  ‣ What are the consequences of accepting respectively refuting?
Critical Thinking

• Impact on testing
  ‣ What we know about the issue – evidence
  ‣ How compatible something is to our own world view
  ‣ Reliability of the source
  ‣ What are the consequences of accepting respectively refuting?

Test results

Expected result – bias

Who tested? How was is tested?

What are the consequences of reporting bugs?
Critical Thinking

• Question your approach

• Science continuously challenges and questions methods, techniques and core beliefs

• Do you question your own approach to testing?
• **Science is impartial**
  ‣ Science can study the effects of deforestation
  ‣ But science does not say that deforestation is bad

  ‣ Truth be told...science does not care
• **Is testing impartial?**
  ‣ Testers can study the effects of bugs
  ‣ Should testers *rate* bugs? Should testers say bugs are bad?
  ‣ Truth be told...maybe sometimes testers should care less
"Science is not static knowledge; it is the dynamic process of exploring the world and seeking to obtain a trustworthy understanding of it."

- Scientific Methods, Richard D. Jarrard

Continuous revision!
Published online 22 September 2011 | Nature | doi:10.1038/news.2011.554
Updated online: 23 September 2011

News

Particles break light-speed limit

Neutrino results challenge cornerstone of modern physics.

Geoff Brumfiel

An Italian experiment has unveiled evidence that fundamental particles known as neutrinos can travel faster than light. Other researchers are cautious about the result, but if it stands further scrutiny, the finding would overturn the most fundamental rule of modern physics — that nothing travels faster than 299,792,458 metres per second.

The experiment is called OPERA (Oscillation Project with Emulsion-tracking Apparatus), and lies 1,400 metres underground in the Gran Sasso National Laboratory in Italy. It is designed to study a beam of neutrinos coming from CERN, Europe's premier

Has OPERA found super-speedy neutrinos?

CERN
Using the Scientific Method

• OPERA (Oscillation Project with Emulsion-tRacking Apparatus) is an Italian experiment 1,400 metres underground

• It is designed to study a beam of neutrinos coming from CERN

• CERN, (Conseil Européen pour la Recherche Nucléaire), located 730 kilometres away near Geneva, Switzerland, is a high-energy physics laboratory

• Neutrinos are fundamental particles

Photo courtesy of CERN
• **So what?**
  ‣ A corner stone of the special theory of relativity is that nothing travels faster than the speed of light in vacuum
  
  ‣ Theory of special relativity forms the foundation for all modern physics
  
  ‣ Was Einstein wrong?
  
  ‣ If true, it would challenge everything we believe to know
Using the Scientific Method

- **Peer review**
  - The results were scrutinized by the scientific community

- **Questioning**
  - Other scientists questioned the results as well as the methods

- **Openness**
  - Questioning and scrutiny was made possible because the scientists working on the OPERA experiment were open with their data and methods
Neutrino experiment replicates faster-than-light finding

Latest data show the subatomic particles continue to break the speed limit.

**Eugenie Samuel Reich**

18 November 2011

Physicists have replicated the finding that the subatomic particles called neutrinos seem to travel faster than light. It is a remarkable confirmation of a stunning result, yet most in the field remain sceptical that the ultimate cosmic speed limit has truly been broken.

The collaboration behind the experiment, called OPERA (Oscillation Project with Emulsion-tracking Apparatus), made headlines in September with its claim that a beam of neutrinos made the 730-kilometre journey from CERN, Europe’s particle-physics lab near Geneva in Switzerland, to the Gran Sasso National Laboratory near L’Aquila, Italy, faster than the speed of light. The result defies Albert Einstein’s special theory of relativity, which states that this cannot happen.
Using the Scientific Method

- Independent replication by ICARUS
- Published in Nature, March 16, 2012

*NATURE | NEWS*

**Neutrinos not faster than light**

ICARUS experiment contradicts controversial claim.

Geoff Brumfiel

16 March 2012  |  Corrected: 19 March 2012

Neutrinos obey nature’s speed limit, according to new results from an Italian experiment. The finding, posted to the preprint server arXiv.org, contradicts a rival claim that neutrinos could travel faster than the speed of light.

Neutrinos are tiny, electrically neutral particles produced in nuclear reactions. Last September, an experiment called OPERA turned up evidence that neutrinos travel faster than the speed of light (see
Who was wrong?

- OPERA identified two potential error sources
  - Bad connection between GPS unit and computer caused incorrect results
  - Neutrinos still obey the speed limit
Using the Scientific Method

• The Scientific Method once again saved the day
  ‣ Scrutiny through peer-review
  ‣ Questioning
  ‣ Openness
  ‣ Nothing is holy

• The difference between science and religion is that science is open for re-evaluation of fundamental theories
Using the Scientific Method

• "Every scientific statement must remain tentative forever."
  ‣ Karl Popper

• “A tester is someone who knows that things can be different.”
  ‣ Jerry Weinberg
• Cooking from recipe
  ‣ List of ingredients and instructions

  ‣ But, as you cook you taste and modify the dish

  ‣ Significant changes to the recipe might be saved as notes in the margin
• Cooking from recipe

Hypothesis (Recipe)

Prediction (Will taste yummy)

Taste

Evaluation (Yummy or yucky)

Revise

Cook
• Testing

Test idea
Test evaluation
Test execution
• Richard Feynman
  † “Philosophy of science is about as useful to scientists as ornithology is to birds.”

• What separates science from non-science?
• Reality is objective and consistent
• Rational explanations exist for elements of the real world
• Non-science
  ‣ Non-science, or pseudoscience
  ‣ Pseudo as in *false*

• How do we recognise pseudoscience?
  ‣ By learning about the real thing – understanding the nature of science

• How do we recognise bad testing?
  ‣ By learning to do good testing – never stop honing your skills
Philosophy of Science

“For something to count as science it must be possible to imagine an observation that would make us reject the hypothesis.”

- Karl Popper
• **Science versus Non-science**
  ‣ Science aims at falsifiable claims
  ‣ Non-science makes claims that cannot be verified

• **Good testing versus bad testing**
  ‣ Good testing investigates potential risks to find failures
  ‣ Bad testing claims to show that the product is working
• **What is non-science?**
  
  ‣ When any of the elements of the Scientific Method are left out:
    • Goal
    • Model
    • Data
    • Evaluation
    • Revision

  ‣ Typically, “Revision” is left out
  ‣ There are no attempts to replicate
Philosophy of Science

• **What is bad testing?**
  ‣ Goal: No clear objective – why are we testing?
  ‣ Model: Poorly chosen or incorrect
  ‣ Data: Poor or limited data
  ‣ Evaluation: No questioning, being biased
  ‣ Revision: Follow a rigid process rather than revising your approach based on what you learn as you test
• **Technology is not science**
  
  ‣ Science can be used to develop better technology
  ‣ Technology is an application of what science taught us
  ‣ Using a computer to analyze data doesn’t make the analysis any more scientific

  ‣ …and using fancy tools doesn’t by itself make the testing any better…
Scientists don't have exclusive rights to the scientific method

Lab coat and goggles are not mandatory accessories
• Embrace this
  ‣ Hypothesis → Prediction → Observation → Evaluation → Revision
  ‣ Knowledge as the objective
  ‣ Peer-review
  ‣ Falsifiable claims
  ‣ Criticism
  ‣ Dynamism
  ‣ Uncertainty
• There is a science to testing
  ‣ Take test seriously – if testers don’t, no one else will either
  ‣ Try to bring a little bit of science into your testing
• I love to talk testing
  › christin.wiedemann@pqa.ca
  › www.pqa.ca
  › christintesting.wordpress.com
  › Twitter: c_wiedemann