

A photograph of a modern university building with a large glass facade and a copper-colored frame. In the foreground, there are wide concrete steps leading up to the building, and a row of flagpoles with various flags. Two bicycles are parked in a rack in the lower right. The sky is blue with some clouds.

innopolis
UNIVERSITY

Scientific and engineering research

S. M. Ahsan Kazmi

Course Outline

Week 1

Introduction

Week 2

Scientific and engineering research

Week 3

Introduction & Literature Review

Week 4

Methodology & Results

Week 5

Proposal Presentation

Week 6

Research Ethics

Week 7

Reviewing

Course Outline

Week 8

LR Presentation 1

Week 9

LR Presentation 2

Week 10

Guest Lecture 1

Week 11

Guest Lecture 2

Week 12

Guest Lecture 3/Presentation skills

Week 13 - 15

Final Presentations

Where can be found?

Primary Literature

- journals
- magazines
- conference/workshop proceedings
- research monographs
- technical reports
- informal media - unpublished papers, internet, talk

Tertiary Literature

- indexes
- on-lines databases

Secondary Literature

- encyclopedias
- dictionaries
- textbooks
- survey articles
- reviews



Goals for the Class

Become educated consumers (all of you)

- understand the research process
- know where to look for information
- be able to evaluate research results

Become competent producers (at least some of you)

- writing
- presentation
- research methods (focus on experimental)

Get an overview of your field

Why understand research?

Your career depends on your ability to

- find
- understand
- evaluate
- make decisions based on Information



Summary

- *Objectives of the course and your tasks*
- *Comments on research and research evaluation*
- **Sources of information in Computer Science**
- How to evaluate a research paper

How to access information in Computer science research

- ... as anywhere else



- Online journals available from scientific databases

- Books and journals available in the library

- E-journals



Search in

>> [Catalogue \(OPAC\)](#)

>> [Databases](#)

>> [E-Journals](#)

>> [Subject Portals](#)



[Reserve Collections 2005/2006](#)

Search for course materials



[Library Courses and Research Guides](#)

Get tips, tricks and tools



[RefWorks](#)

Manage your references



[CitationLinker](#)

Search for known articles



[Electronic Journals Library \(EZB\)](#)

OTHER SUPPLIERS



[EBSCO HOST - Electronic Journals Service](#)



[Science Direct - Web Editions](#)



[Springer Science+Business, formerly Kluwer online](#)



[JSTOR](#)



[Springer Link](#)



[ACM Digital Library](#)



[Ingenta](#)



[Blackwell](#)



[Cambridge : subscribed Journals](#)

E-Journals

- ACM digital library (<http://portal.acm.org/portal.cfm>)
- Springer Link
 - EMSE (Journal of Empirical Software Engineering, <http://www.springerlink.com>)
- IEEE Explore (<http://ieeexplore.ieee.org>) and IEEE Computer Society, IEEE digital library (<http://www.computer.org>)

Some important organizations for the Computer Science community

- [ACM: The Association for Computing Machinery](http://www.acm.org/)
<http://www.acm.org/>

the first society in computing

- [IEEE Computer Society](http://www.ieee.org/)

With over 100,000 members, the IEEE Computer Society is the world's leading organization of computer professionals. Founded in 1946, it is the largest of the 36 societies of the Institute of Electrical and Electronics Engineers (IEEE).

- [USENIX Association \(Advanced Computing Systems Association\)](http://www.usenix.org/)
<http://www.usenix.org/>

USENIX, the Advanced Computing Systems Association brings together the community of engineers, system administrators, scientists, and technicians working on the cutting edge of the computing world.

- [SAGE: System Administrators' Guild](http://www.sage.org/)
<http://www.sage.org/>

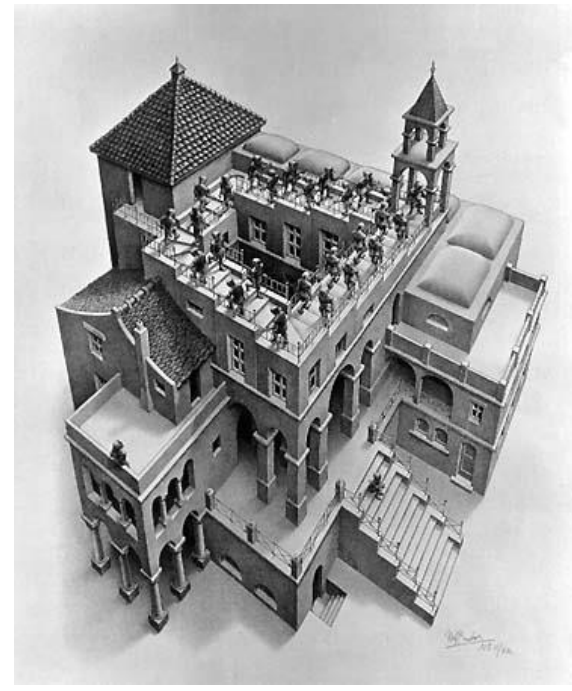
- ...

Today's Lecture

Scientific and engineering research

What is science?

The word "science" means "knowledge." It comes from the Latin word, "scire," "to know." The baseline definition of "science," then, is human knowledge.

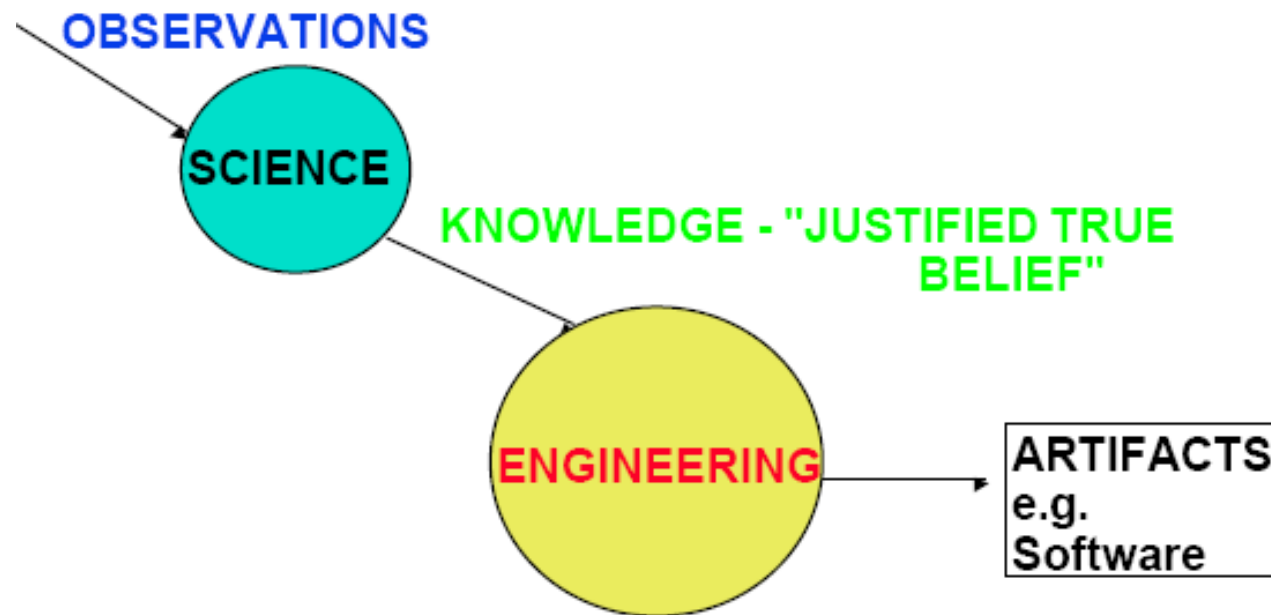


Important Issues

- Define science and explain how it works
- Define engineering and explain how it works
- Explain the relationship between science and engineering

Science and engineering

A Model Relating Science And Engineering



- ENGINEERS STRIVE TO BUILD BETTER ARTIFACTS

Science

- Definition 1: The goal is to establish the **truth** of statements about the world

What (if anything) is wrong with this definition?

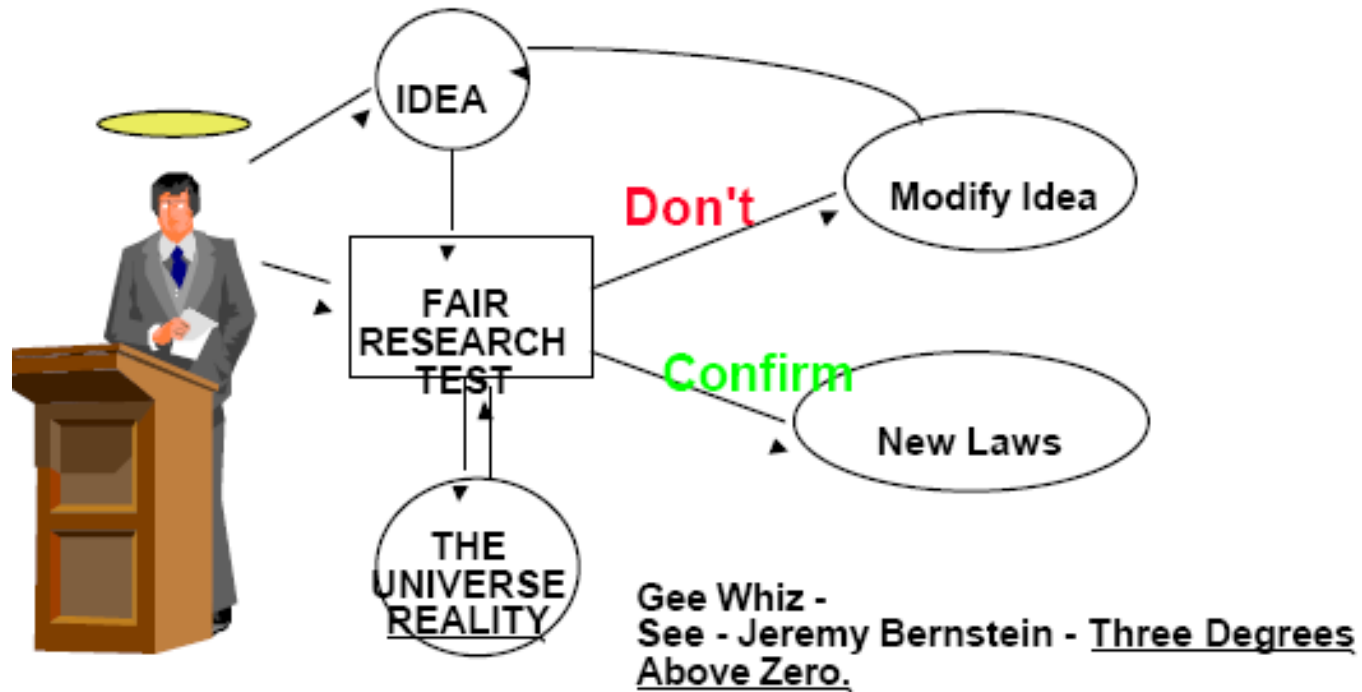
You cannot establish the absolute truth or falsity of a statement based on induction

- Definition 2: The goal of science is to establish the **probability of truth** of a statement based on induction

But science isn't really about just statements – it is about theories

Utopia

"The Scientific Method is an objective search by unbiased scientists for true knowledge."



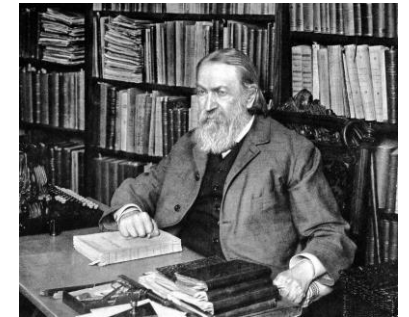
Reality

"The scientific arena is ... intellectual warfare fought by scientists who are passionate partisans of one theoretical viewpoint or another"



Scientific Warfare

- CREATIONISTS VERSUS DARWINIANS
- Mach versus Boltzmann
 - Boltzmann's ideas have not been accepted in the scientific community ... this was one of the reasons for his suicide
- Einstein versus Bohr
- EMACS users versus everyone
- See:
 - James Watson - The Double Helix
 - Martin Gardner - SCIENCE - GOOD, BAD, & BOGUS



Utopia: some rules of science

- Open communication of ideas and results in the literature
- 50% of published papers are never cited
- 50% of psych papers have incorrect data
- politics is involved in reviewing
- papers are sometimes faked

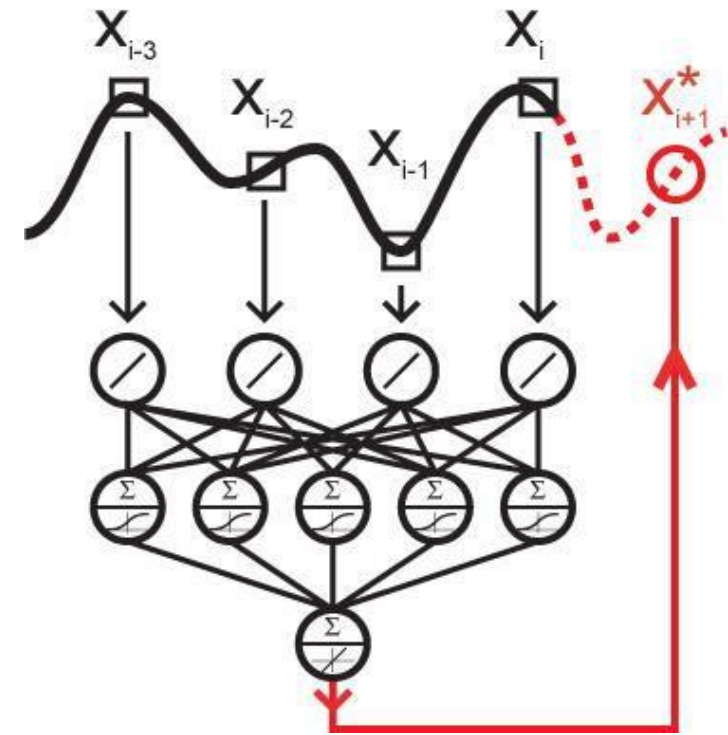
- See -

William Broad - Betrayers of the truth

Appeal to the evidence

A theory must be justified by

- test of a prediction
 - e.g. regression testing will reduce the number of bugs in a system
- solution to a problem
- other



Mutability

- the ability to alter previous theoretical positions to accommodate new empirical results
- but... contradictory evidence is sometimes squashed

Replication

- Results should be replicable - even by an adversary

Some paradigms in Computer Science

- Organizations - Field Studies
- Human Behaviour - Field Studies & Controlled Experiments
- Algorithm Analysis - Math Simulation
- Design - Controlled Experiments?
- Programming - Controlled Experiments?
- Compilers - Math

Evolution of science

- During Normal Science (most of the time)

- knowledge is accumulated
- problems are solved

But

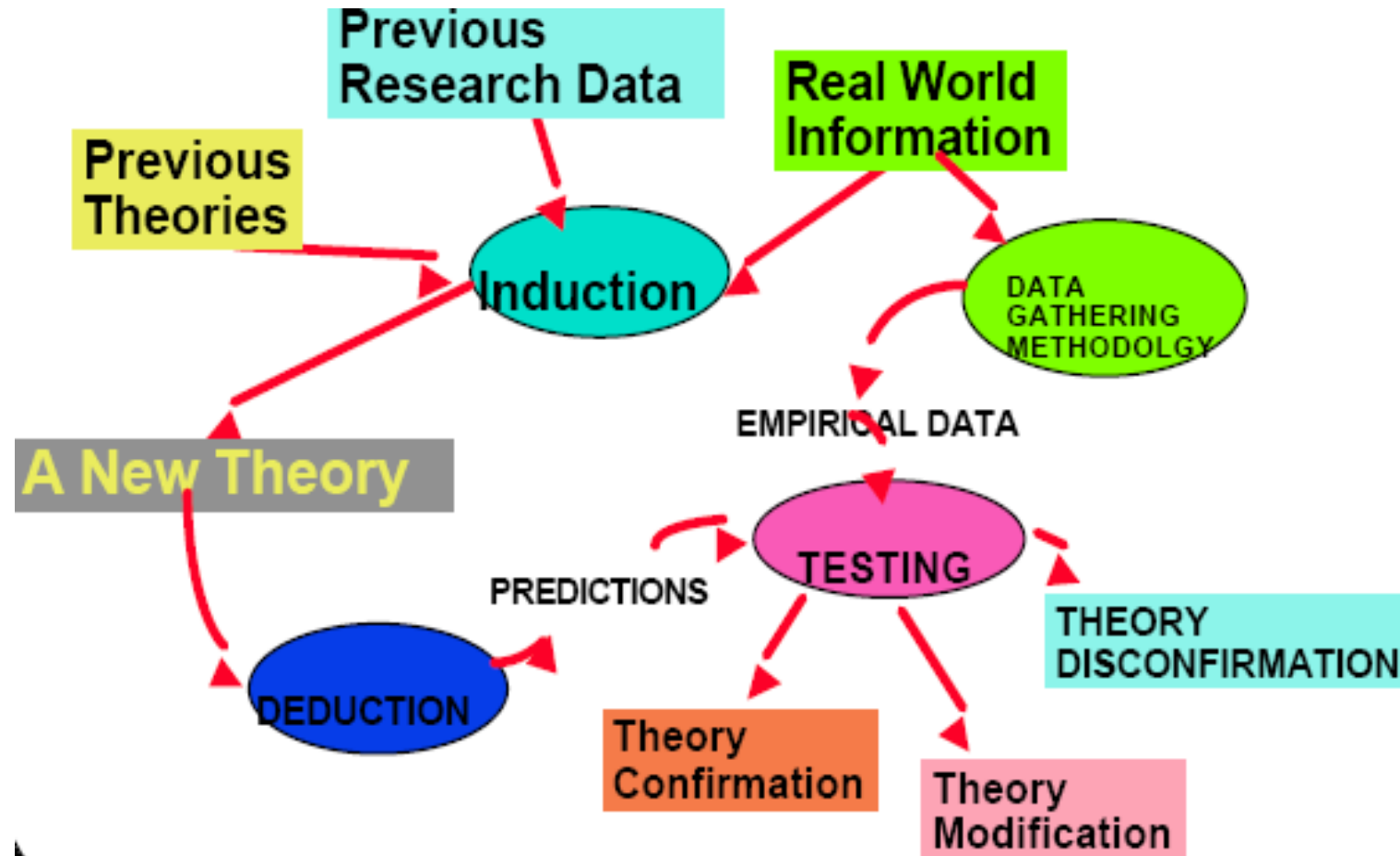
- some problems cannot be solved within the paradigm
- empirical results are found that cannot be fit into any paradigm
- This leads to DISSATISFACTION and sometimes to REVOLUTION
 - e.g. COPERNICUS



Is Computer Science in a normal science stage at present?

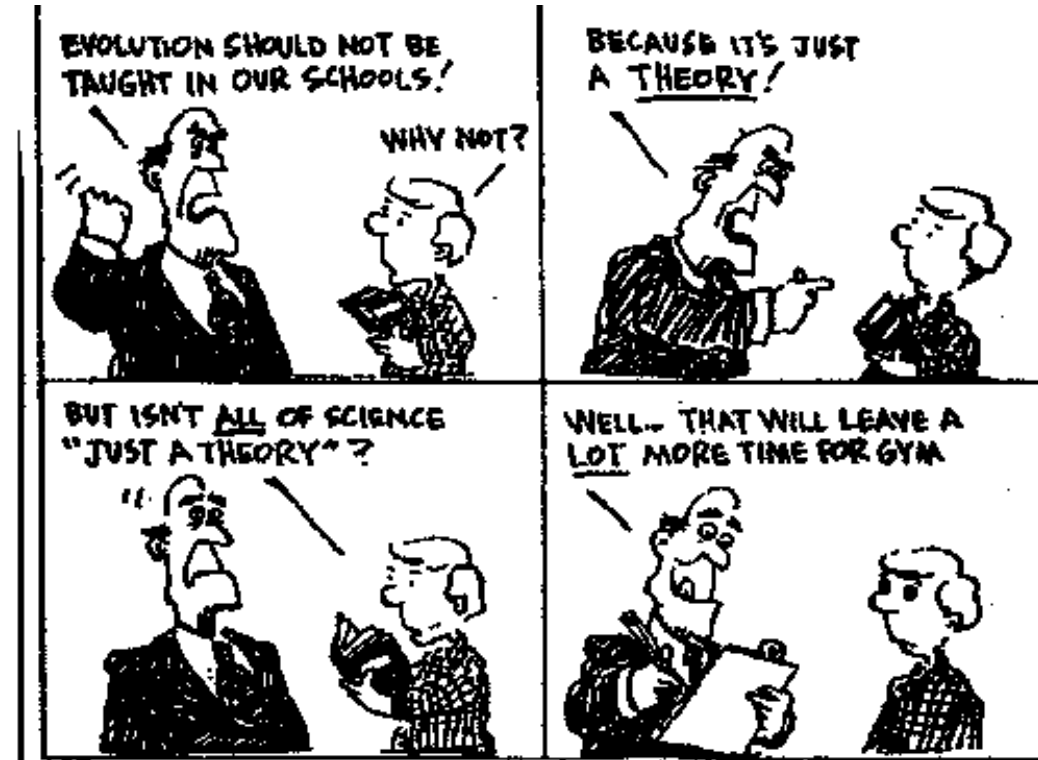
- Have there been computer science revolutions in the past?
- If computer science is currently in a normal state, are there any anomalies lurking around?
- See Walter F. Tichy, Should Computer Scientists Experiment More?

A Model Of The Scientific Process (Kemeny, Holt)



What is a Theory

- No consensus – as usual!



Vague
Verbal
Ideas

FREUD

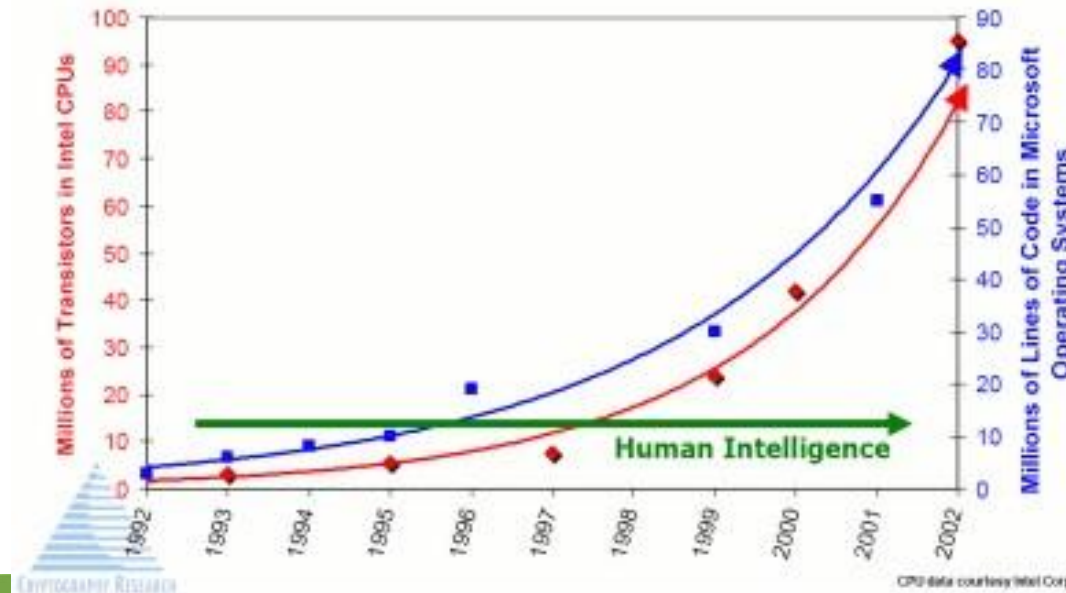
Rigorous
Math
Equations

NEWTON

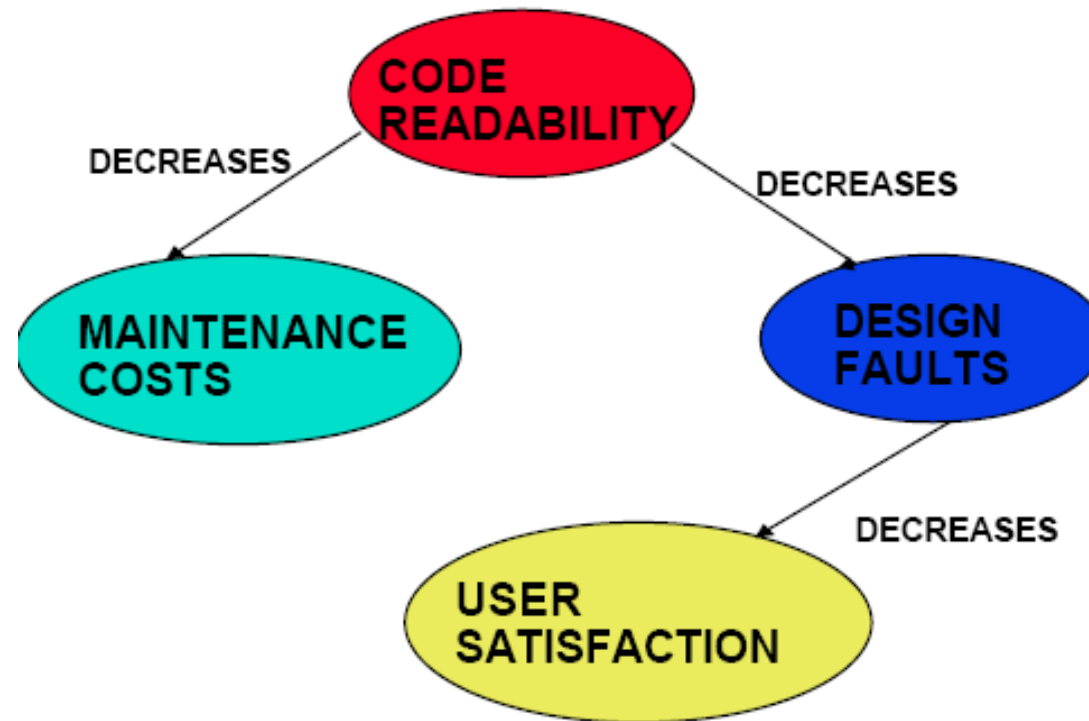
THEORY

A definition of theory (Holt)

- "Theories specify one or more relationships among two or more variables."
- Is this a theory according to this definition?
 - "The number of bugs in a piece of code goes up as NCSL goes up"

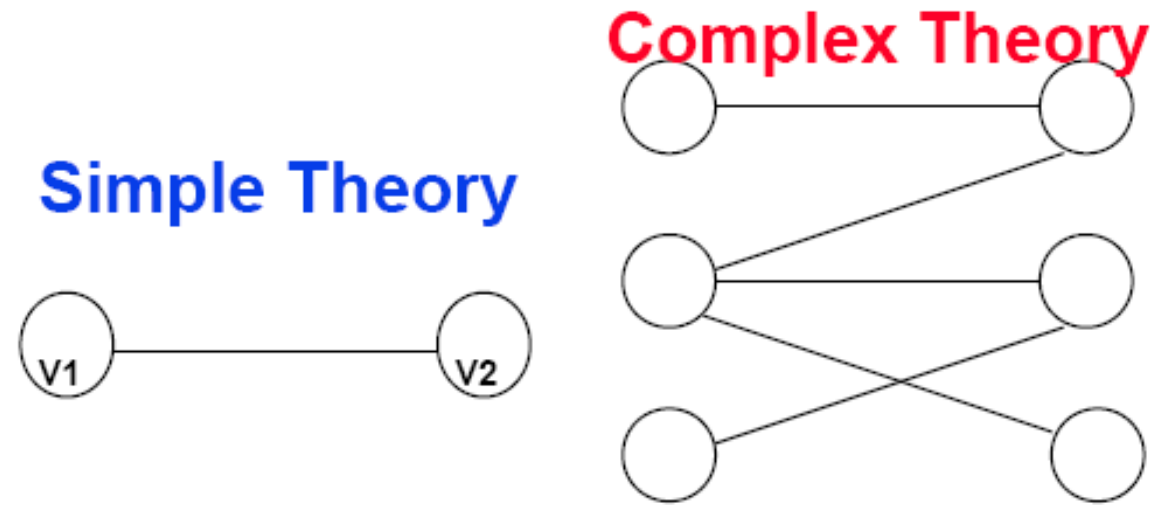


A THEORY



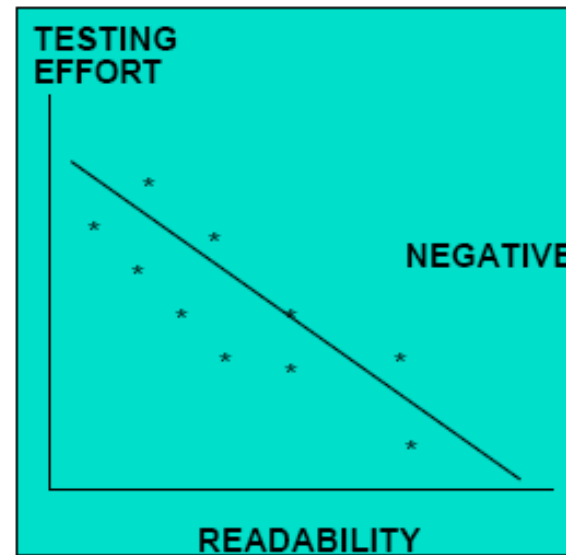
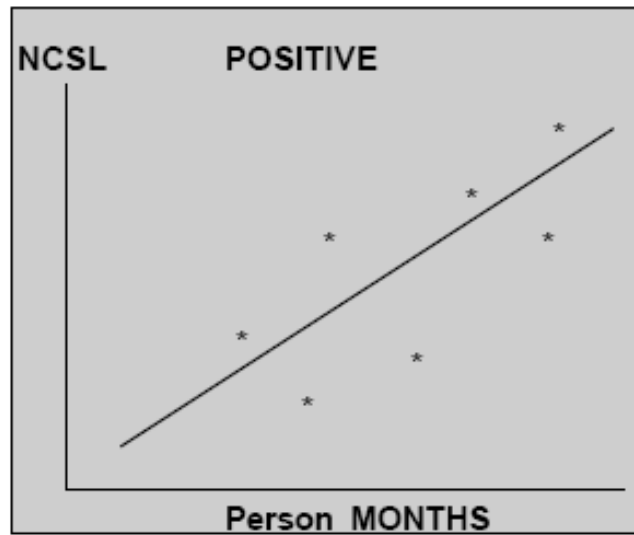
Complexity of a theory

- Complexity - # of variables in the theory, and the number of relations



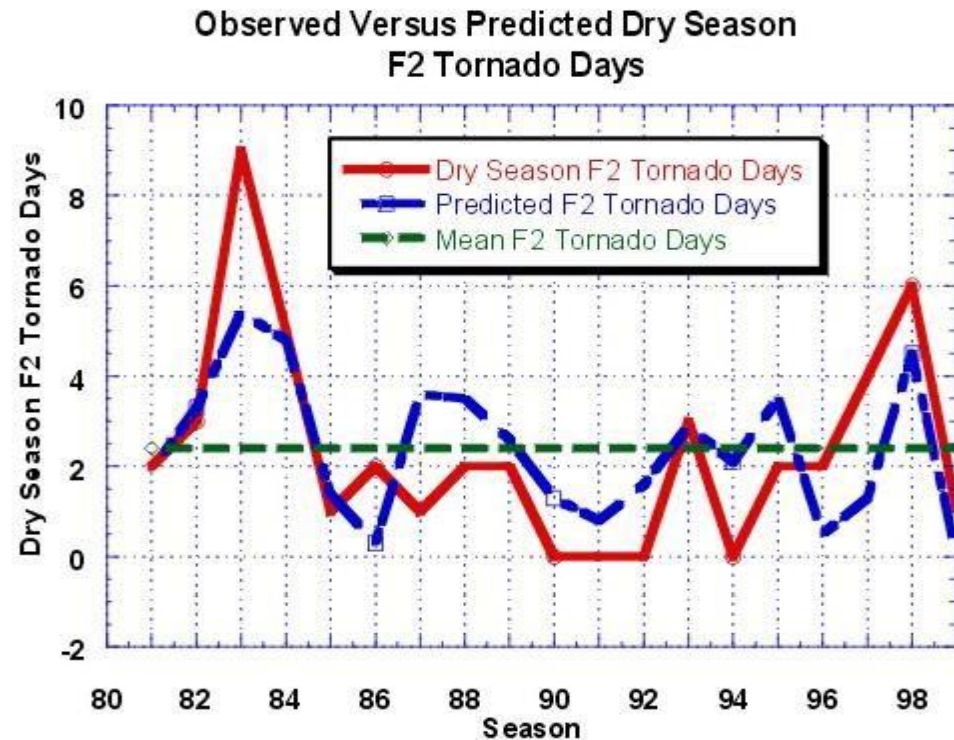
Direction of relationship

- Does a variable facilitate(positive) or inhibit(negative) another?



Criteria for a theory

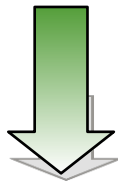
- Internal Consistency
- Explanation
- Prediction
- Control
- Scope



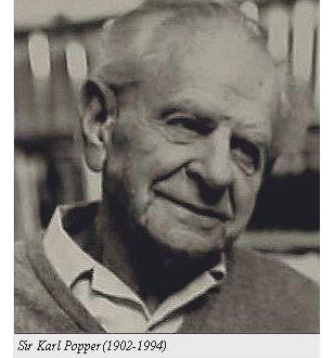
Internal consistency

- A theory must not contradict itself because it must be testable, i.e. falsifiable

- "Adding more personnel to a software project will increase productivity, or it won't"



- Such a theory can never be falsified!



Sir Karl Popper (1902-1994)

Karl Popper

Explanation

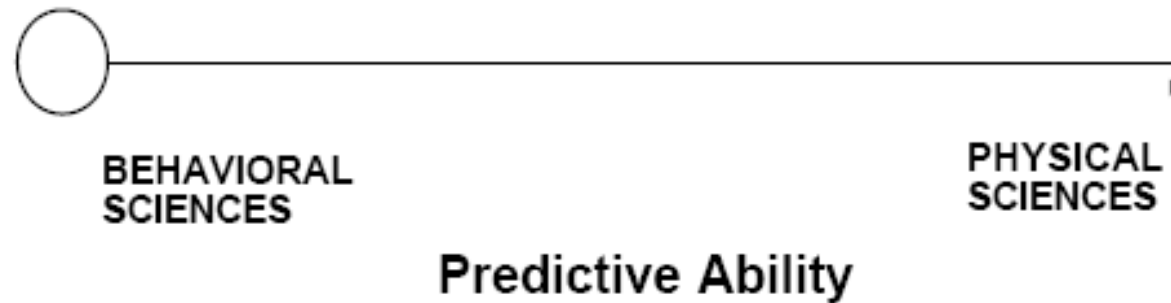
- A good theory should plausibly explain the phenomena
- This is a pretty vague notion
 - What does plausible mean?
 - Plausible to whom?
- "The position of the planets affects human behavior."
- "If you don't let PROGRAMMERS COMPILE THEIR CODE, you'll get fewer BUGS"

Prediction

- Can the theory predict future events?
- Theory - If you increase testing time, the number of remaining faults will decrease

Predictive Ability

- All things being equal, the more specific the prediction, the better the theory



Control

- A good theory allows us to control the phenomena
- e.g. physical theory allows us to control the orbits of satellites

Scope

- The domain of phenomena covered by a theory
- More is better
 - Can there be a unified theory of software engineering?
 - What components would it need to have?
 - What types of phenomena would it need to account for?

Science and engineering

Simon Summary

REAL WORLD

ARTIFICIAL WORLD

(MAN MADE)

Analytic - Study the world
to see how it is

Synthetic - Create new things

Non-normative - declarative
- state the way things are

Normative - Imperative - state
goals, how things should be

Need a complete and
precise account of
phenomena

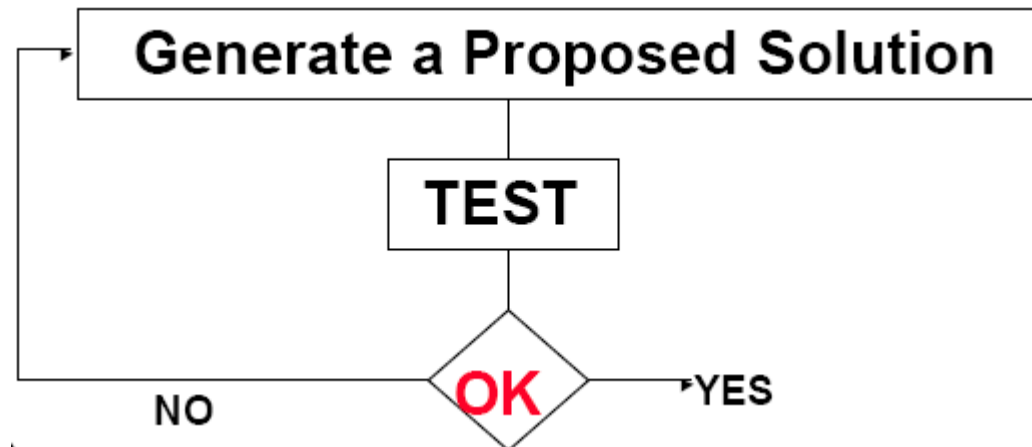
Partial information o.k. You
don't need to know quantum
mechanics to build a bridge

Hiding internal details and
focusing on interfaces
important

Design as search

Design is goal directed

- Therefore, we can do a search to move towards a goal
- Generate and Test



Model = theory?

- "To control the software development and maintenance processes, it is important to model* certain interesting factors(metrics), such as effort and defects, based on other available metrics

* model - a mathematical relationship among software metrics

Theoretical model

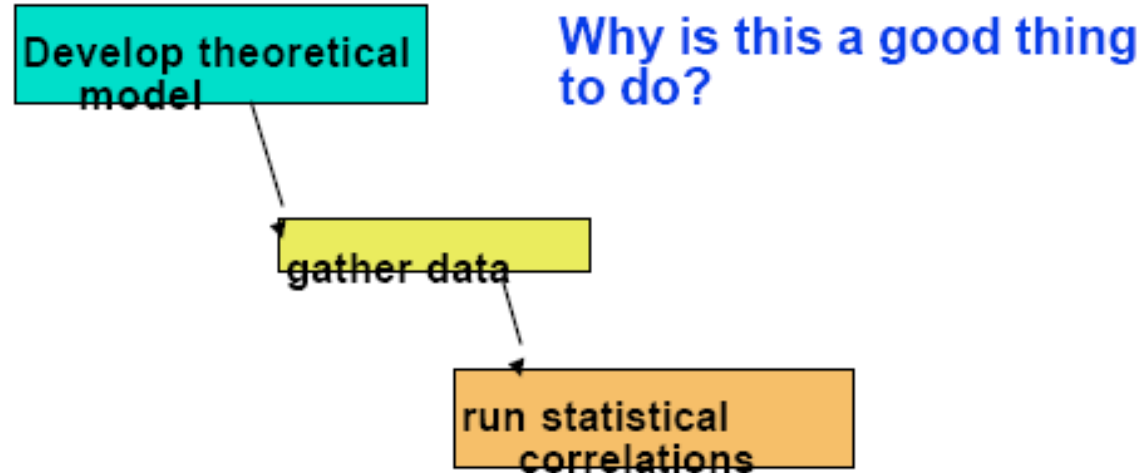
- A theory driven model is preferred:
 - guards against capitalizing on chance e.g. misuse of a correlation table
 - based on hypothesized relationship between variables
 - data independent
 - e.g. What factors will affect productivity expressed as new NCSL/time?

Data driven model

- Usually based on statistical analysis
 - may not be explanatory

Combined model

- Statistics constrained by theory



End of Lecture