Scientific thinking

Folded paper question:

What's the point of science?

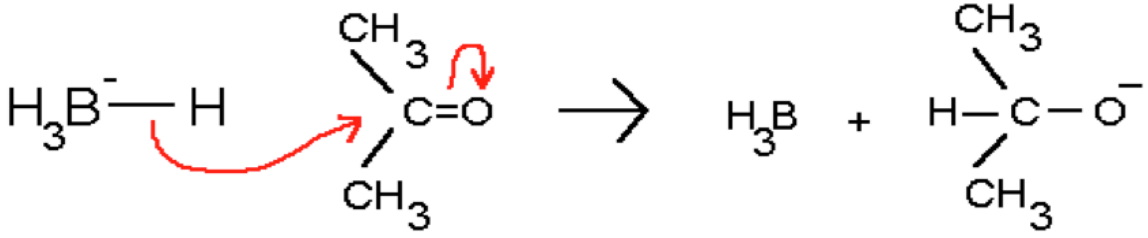
(Please answer in a *simple* sentence, fold your paper and give it to me)

Scientific understanding (for what are we aiming ?)

Physics:

$$\left(\frac{-h^2}{2m}\right)\frac{d^2\varphi}{dx^2} = E\varphi(x)$$

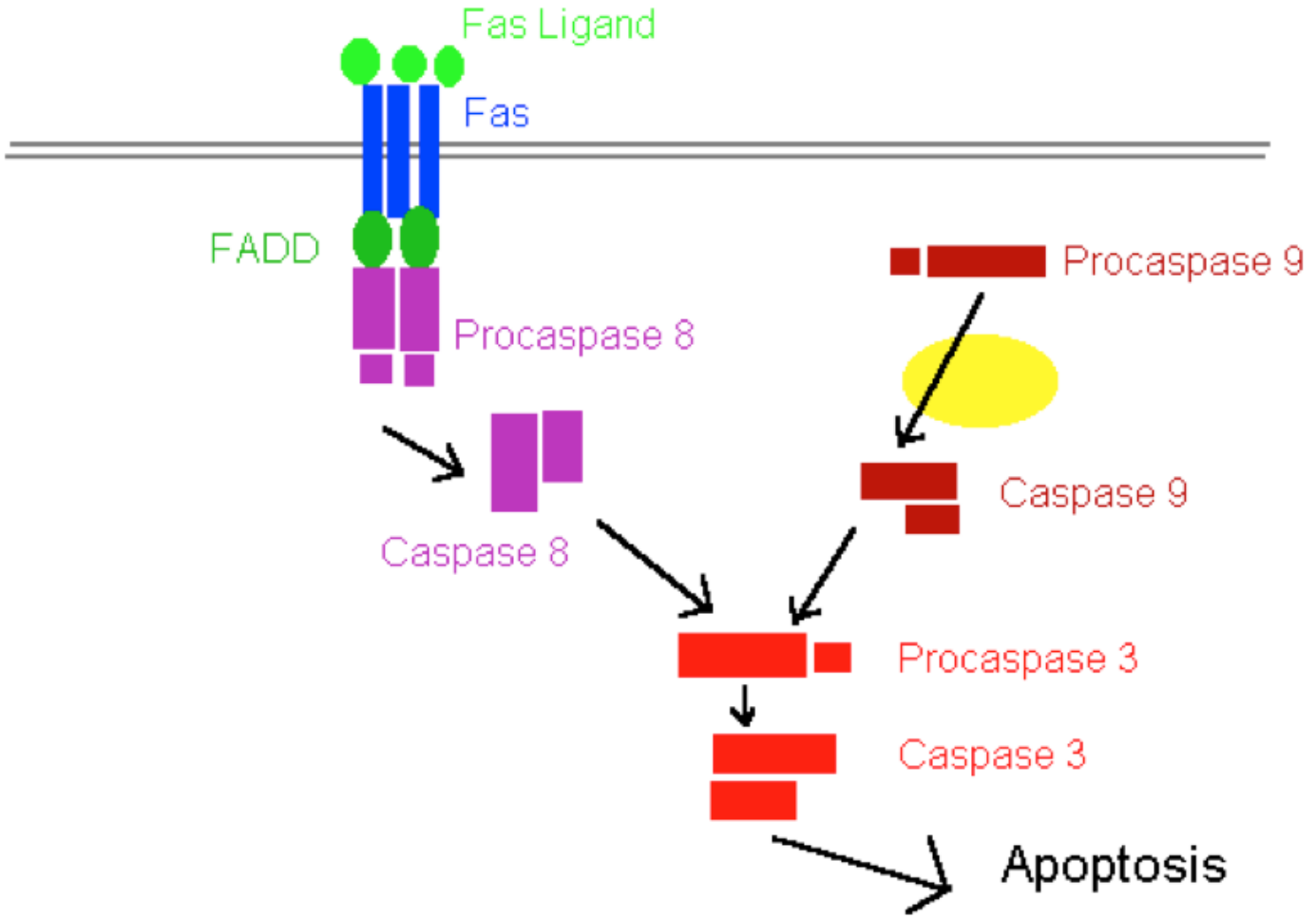
Chemistry:



How about us?

What constitutes 'understanding' in (biomedical) biology?

One answer:


Fas Ligand
Fas
FADD
Procaspase 8
Caspase 8
Procaspase 9
Caspase 9
Procaspase 3
Caspase 3
Apoptosis

I am not trying to get you to accept one answer.

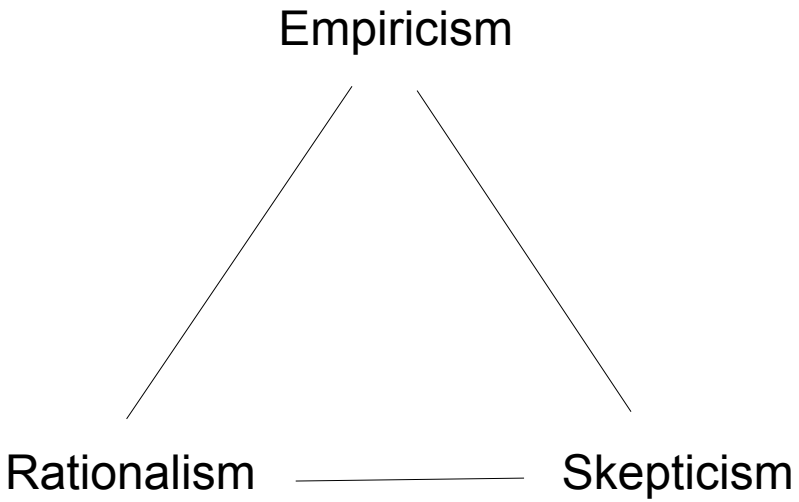
I am trying to get you to ASK YOURSELF what you think 'understanding' is in your field.

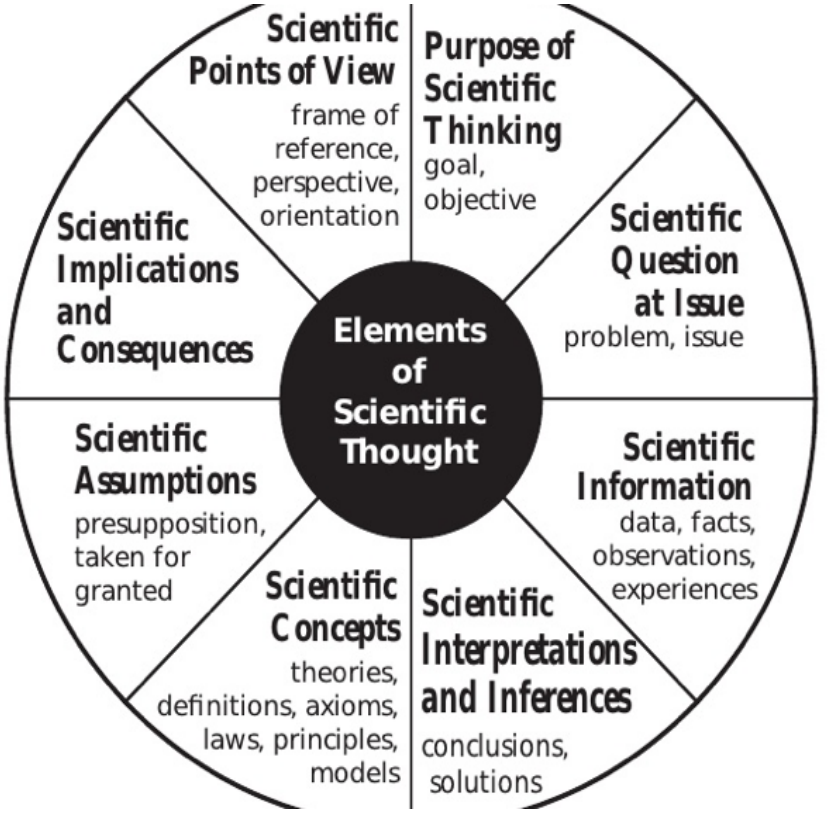
If you are not clear on your goal, you are unlikely to reach it.

Thinking scientifically

Class exercise

What distinguishes “scientific thinking” from other sorts?





Foundation for critical thinking, 2003.

Pathways to understanding

- Bacon
- Popper

Francis' Bacon's view:

•Explore and play

•Try to notice patterns in what happens

•Find general principles by induction

In vitro translation in Rabbit reticulocyte lysate:
•Cows use the same genetic code as rabbits
•Peas use the same genetic code as rabbits
•Yeast uses the same genetic code as rabbits

→ all organisms (on Earth) use the same genetic code

Induction can include interpolation/ extrapolation

Example: Johan Daniel Titius (1766) noticed that the orbit sizes of the planets obeyed the following pattern:

Orbit size(/AU) = $0.4 + 0.3 \times 2^m$

where m = -∞, 0, 1, 2 …

Titius

Planet	m	predicted	real
Mercury	-∞	0.4	0.4
Venus	0	0.7	0.7
Earth	1	1.0	1.0
Mars	2	1.6	1.5
	3	2.8	
Jupiter	4	5.2	5.2
Saturn	5	10	9.5
Uranus	6	19.6	19.2

Induction can include interpolation/ extrapolation

Example: Johan Daniel Titius (1766) noticed that the orbit sizes of the planets obeyed the following pattern:

Orbit size(/AU) = $0.4 + 0.3 \times 2^m$

where m = -∞, 0, 1, 2 …

Titius

Planet	m	predicted	real
Mercury	-∞	0.4	0.4
Venus	0	0.7	0.7
Earth	1	1.0	1.0
Mars	2	1.6	1.5
	3	2.8	
Jupiter	4	5.2	5.2
Saturn	5	10	9.5
Uranus	6	19.6	19.2

"But should the Lord Architect have left that space empty? Not at all."

Induction can include interpolation/ extrapolation

Example: Johan Daniel Titius (1766) noticed that the orbit sizes of the planets obeyed the following pattern:

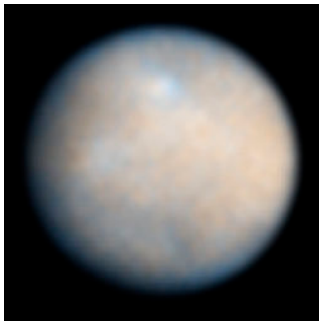
Orbit size(/AU) = $0.4 + 0.3 \times 2^m$

where m = -∞, 0, 1, 2 …

Titius

Planet	m	predicted	real
Mercury	-∞	0.4	0.4
Venus	0	0.7	0.7
Earth	1	1.0	1.0
Mars	2	1.6	1.5
Ceres et al.	3	2.8	2.8
Jupiter	4	5.2	5.2
Saturn	5	10	9.5
Uranus	6	19.6	19.2

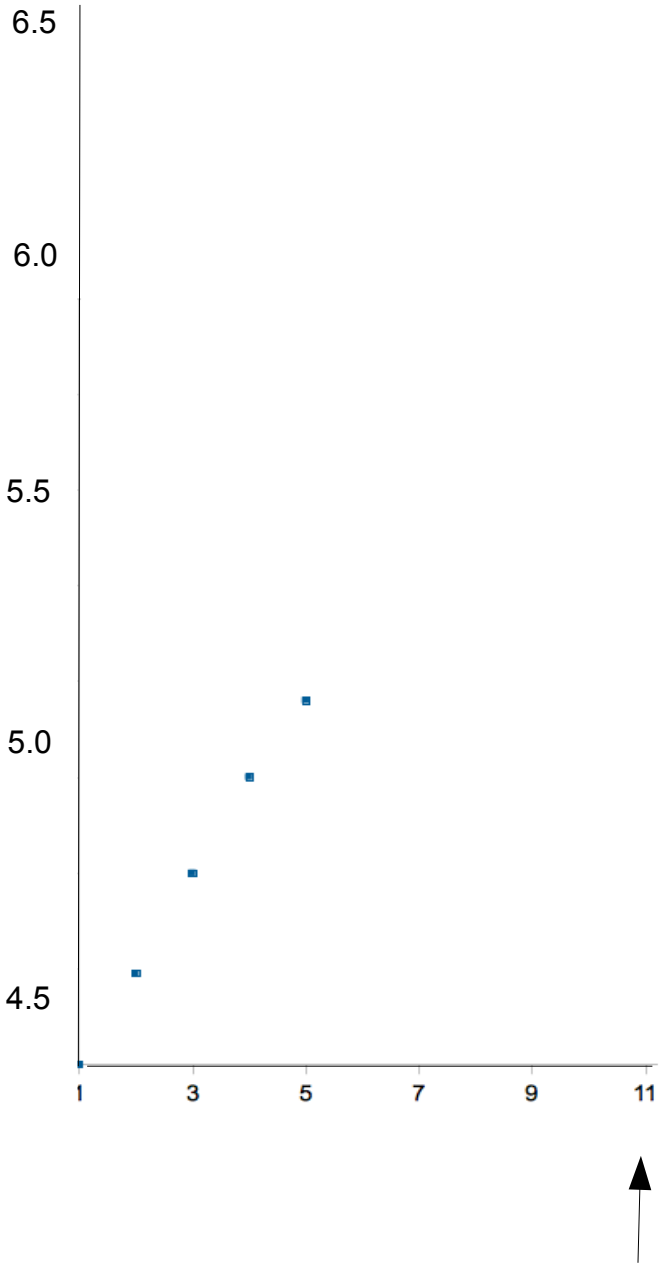
"But should the Lord Architect have left that space empty? Not at all."



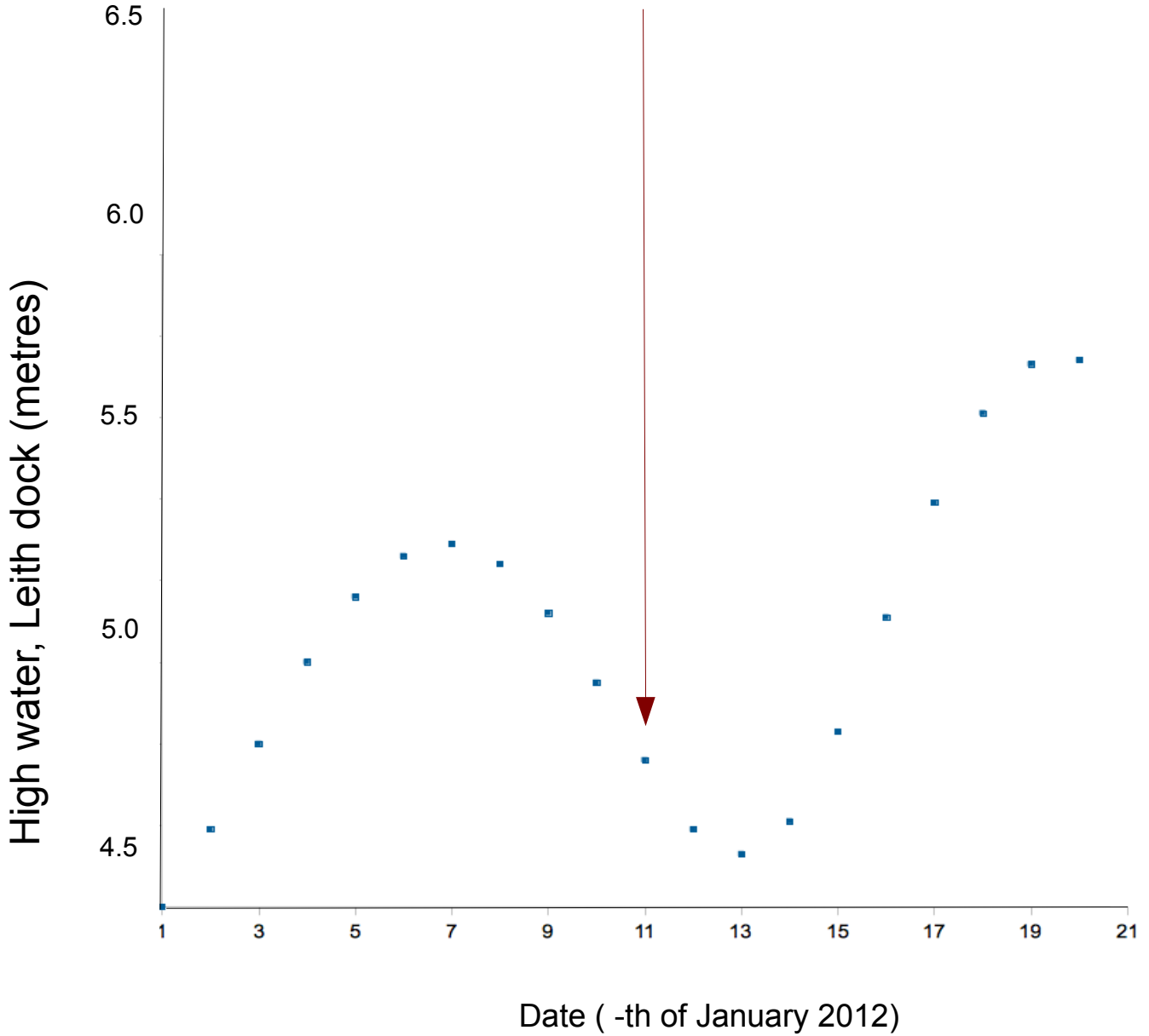
Guiseppe Piazzi

Ceres (dwarf planet, disc. 1801)
photo: Hubble Space Telescope

Extrapolation: what interval would you give for the y value when x is 11?

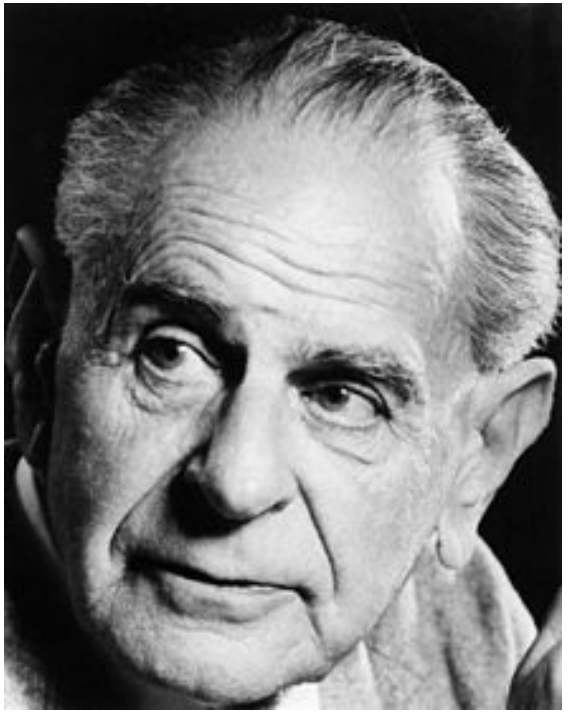


That's the problem with inductive reasoning....

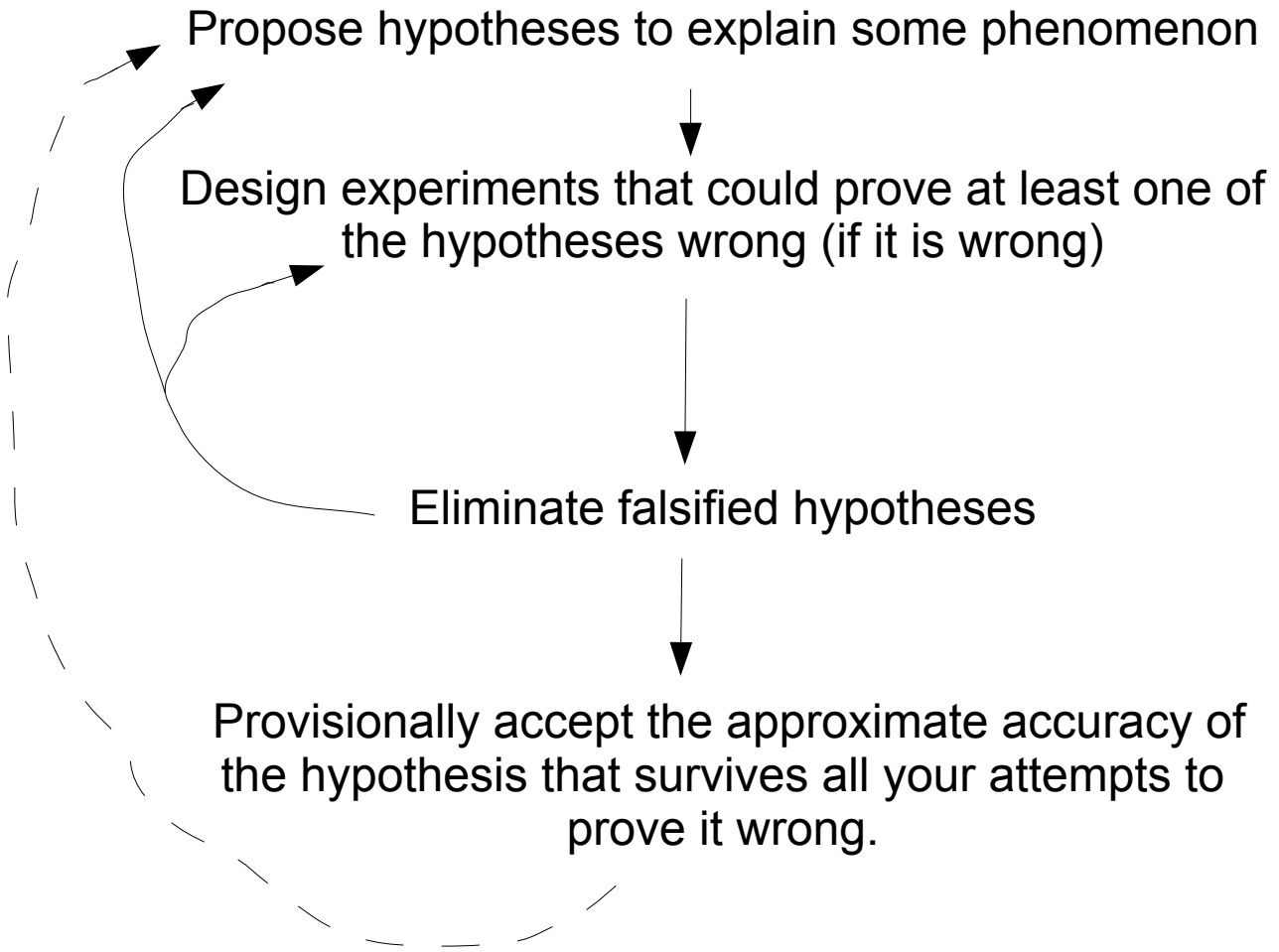


Moral: extrapolation is dangerous !!!

Conjecture and Refutation



Carl Popper



Class exercise:

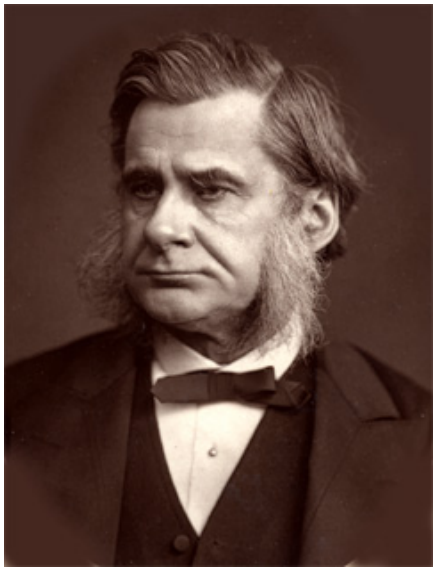
"Overweight mice die younger than lean ones"

- assume the fact in this sentence is true.

Suggest hypotheses to account for this fact.

How would you test them?

The great tragedy of science is the slaying of a beautiful hypothesis by an ugly fact.



Thomas Huxley

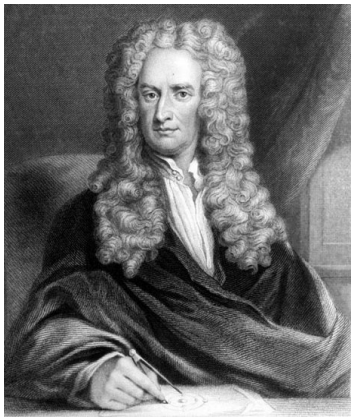
No hypothesis is safe from later refutation:

No hypothesis is safe from later refutation:

Epitaph on Newton:

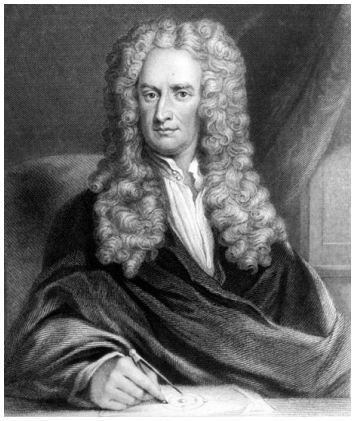
Nature and Nature's law lay hid in night:

God said, "Let Newton be!," and all was light.



No hypothesis is safe from later refutation:

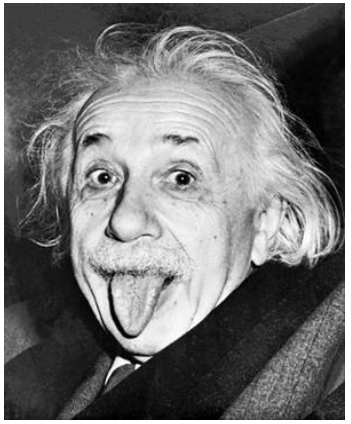
Epitaph on Newton:



Nature and Nature's law lay hid in night:

God said, "Let Newton be!," and all was light.

added by Sir John Collings:



It did not last: the Devil shouting "*Ho!*

Let Einstein be!" restored the status quo.

Coming up with hypotheses:

Imagination *Imagination is more important than knowledge...* (Einstein)

Critical examination of assumptions *Assumption is the mother of screw-ups.* (Angelo Donghia)

Pattern recognition

Here is a number sequence that obeys a rule:

2, 4, 6...

Please discover the rule, by proposing other 3-number sequences and ask me if they obey the rule or not. I will answer 'yes' or 'no'.

Here is a number sequence that obeys the a rule:

2, 4, 6...

Please discover the rule, by proposing other 3-number sequences and ask me if they obey the rule or not. I will answer 'yes' or 'no'.

(The rule was just that numbers had to be in ascending order)

The important point is that, although you all nodded sagely when we discussed conjecture and refutation and the value of proving hypotheses wrong, almost everyone *actually* only tries to confirm not refute.

With thanks to P C Watson for this game.

Imperfect data – probability and statistical reasoning

Estimation and confidence limits.

How many road vehicles are licensed in the UK (total population of people c. 60 million)?

Please choose two numbers, a lower limit and an upper limit, between which you feel 95% confident the actual number lies. Work on your own.

(For example, if you were asked to guess many people work in this building, you may say you are 95% certain the answer is between 90 and 120).

Also, how many lovers did Catherine the Great have?

Please write them on a slip of paper, and hand it in.

Estimation and confidence limits.

How many road vehicles are licensed in the UK (total population of people c. 60 million)?

Please choose two numbers, a lower limit and an upper limit, between which you feel 95% confident the actual number lies. Work on your own.

(For example, if you were asked to guess many people work in this building, you may say you are 95% certain the answer is between 90 and 120).

Also, how many lovers did Catherine the Great have?

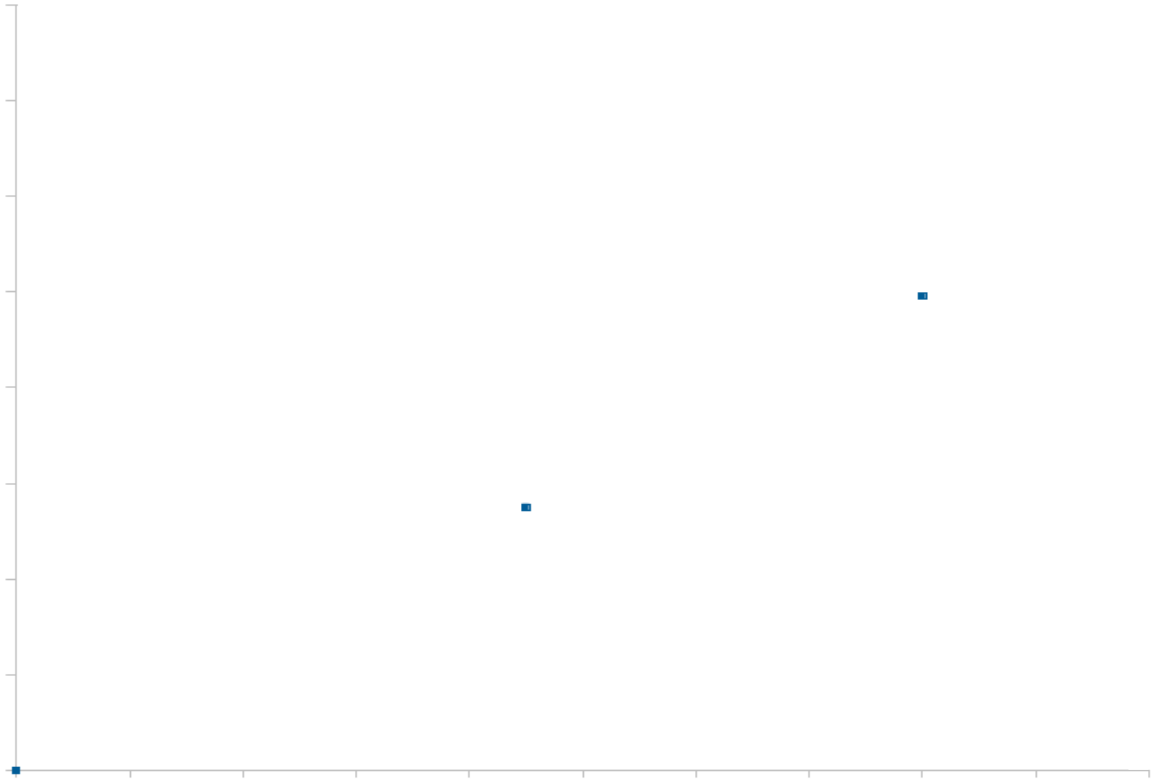
The actual figures are 33,000,000 and 12 respectively.

How well did your free choice of confidence limit reflect your actual knowledge?

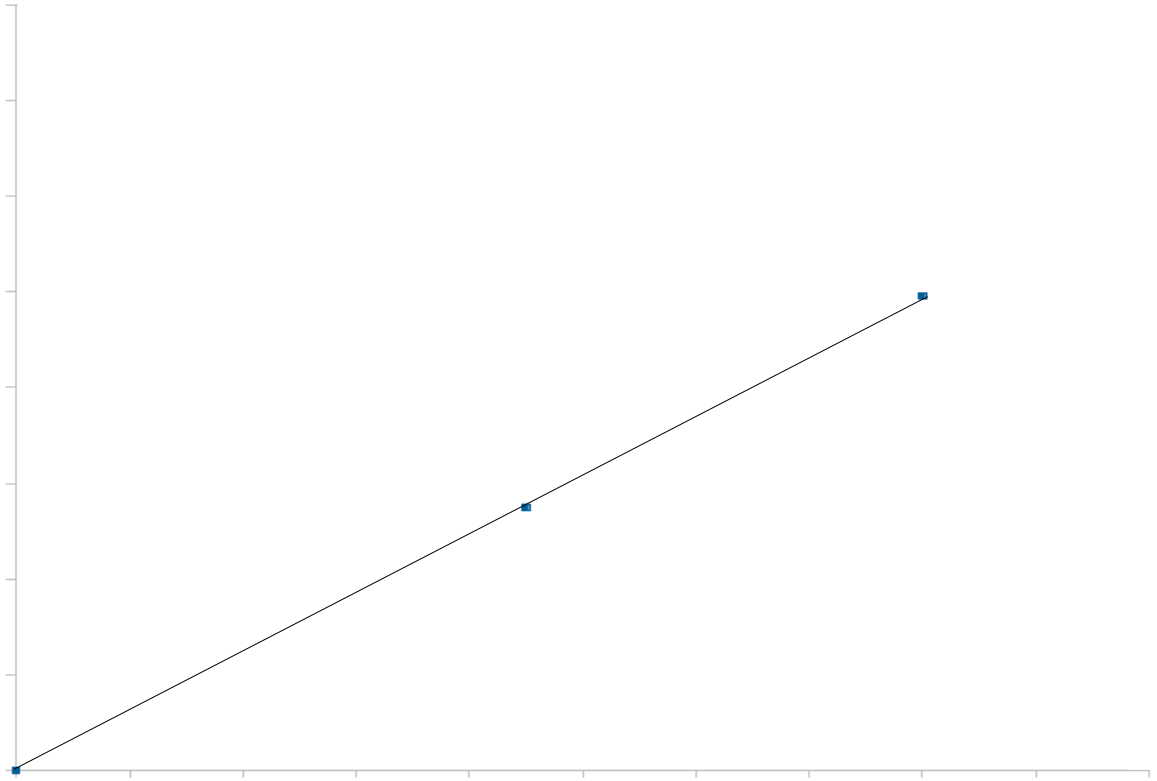
(In New York financial traders, about 30% of people playing this game set the limits confidently close together around a completely wrong number)

With thanks to Nicolas Nassim Taleb for this game

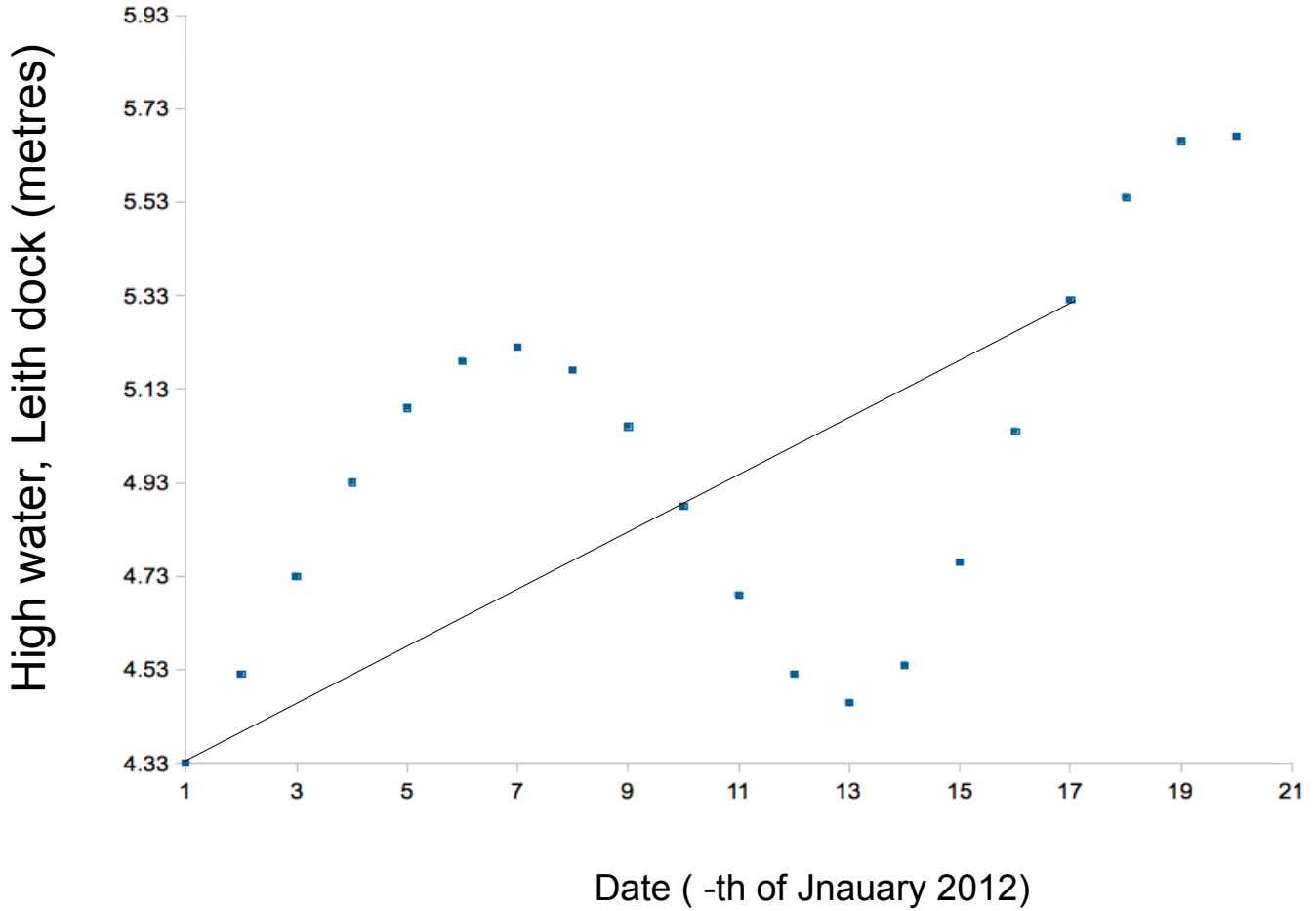
Interpolation



Interpolation

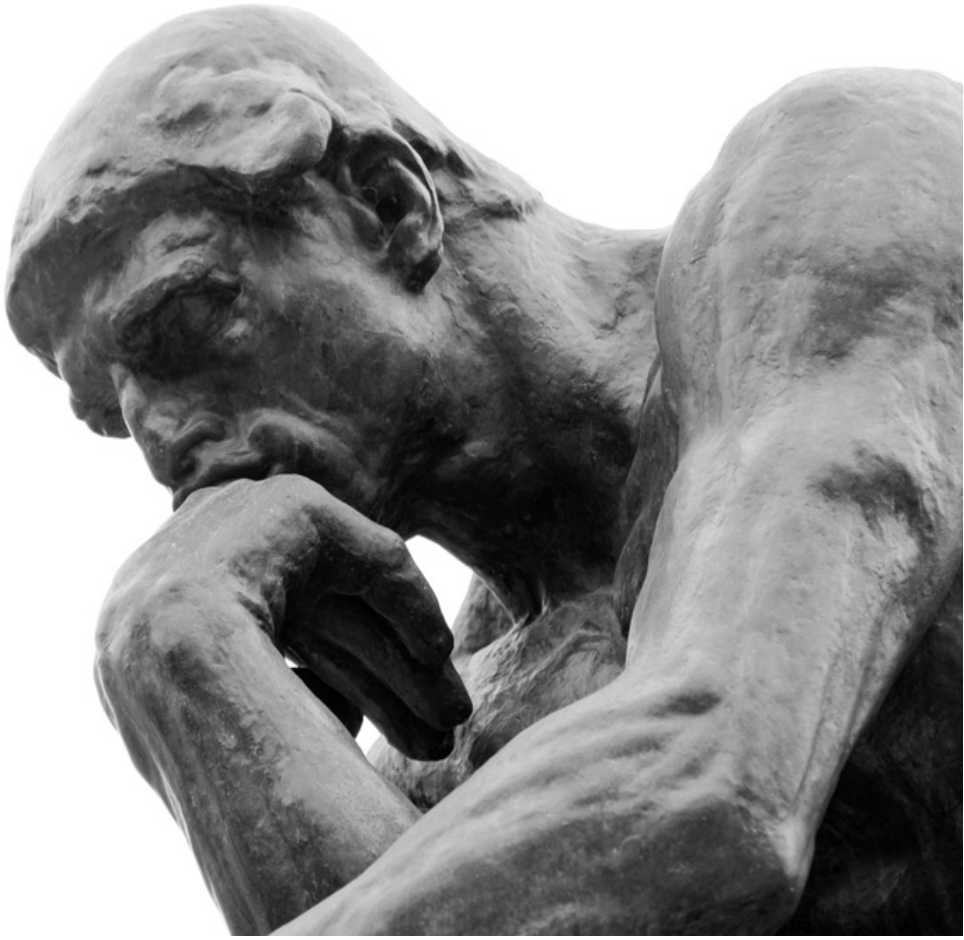


Interpolation



Interpolation is most dangerous when you space out your measurement points (especially when you space them regularly, as people tend to do)

Part II:



Here are two micro-stories: which strikes you as more probable?

1) Alice and Bob seemed happily married. Then she killed him.

2) Alice and Bob seemed happily married. Then she killed him because he had been cheating on her.

Here are two micro-stories: which strikes you as more believable?

1) Alice and Bob seemed happily married. Then she killed him.

2) Alice and Bob seemed happily married. Then she killed him because he had been cheating on her.

This probability includes all possible reasons for her killing him (accident, jealous rage etc) *including the infidelity*.

Here are two micro-stories: which strikes you as more believable?

1) Alice and Bob seemed happily married. Then she killed him.

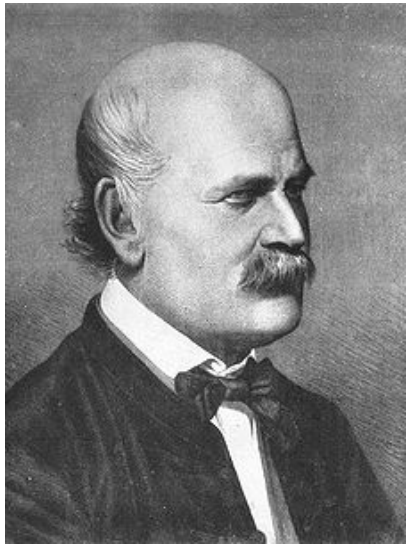
This probability includes all possible reasons for her killing him (accident, jealous rage etc) *including the infidelity.*

2) Alice and Bob seemed happily married. Then she killed him because he had been cheating on her.

We instinctively believe things more when we have an explanation for them.

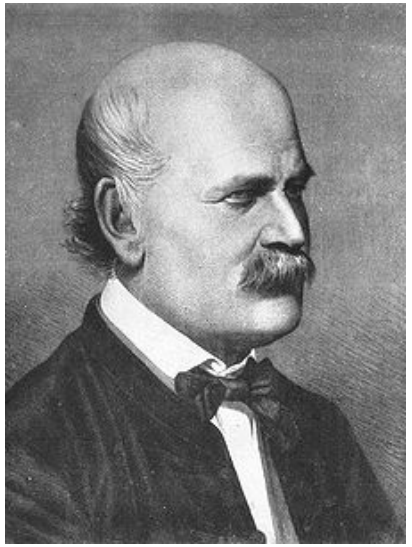
(most people, most of the time).

Ignaz Semmelweis



1847 – showed that washing hands between post-mortems and patients cut incidence of puerperal fever from 10-35% to < 1%.

Ignaz Semmelweis

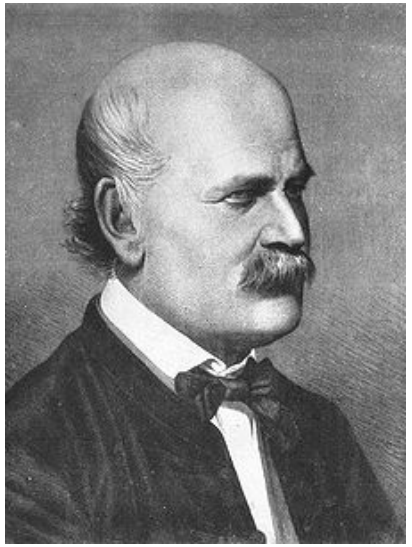


1847 – showed that washing hands between post-mortems and patients cut incidence of puerperal fever from 10-35% to < 1%.

Largely ignored by medical community because it made no sense. Puerperal fever remained common in most of Europe.

1865 – Semmelweis committed to asylum → died.

Ignaz Semmelweis



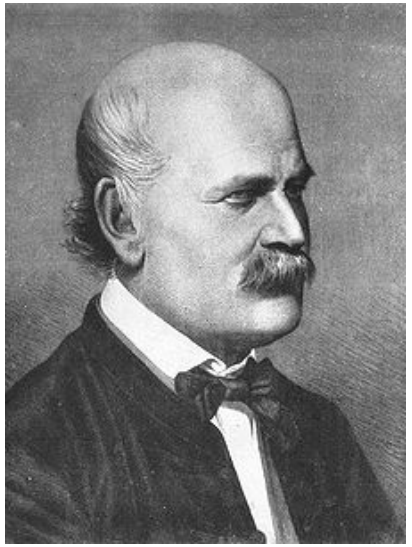
1847 – showed that washing hands between post-mortems and patients cut incidence of puerperal fever from 10-35% to < 1%.

Largely ignored by medical community because it made no sense. Puerperal fever remained common in most of Europe.

1865 – Semmelweis committed to asylum → died.

1878 – Pasteur publishes 'germ' theory of disease.

Ignaz Semmelweis



1847 – showed that washing hands between post-mortems and patients cut incidence of puerperal fever from 10-35% to < 1%.

Largely ignored by medical community because it made no sense. Puerperal fever remained common in most of Europe.

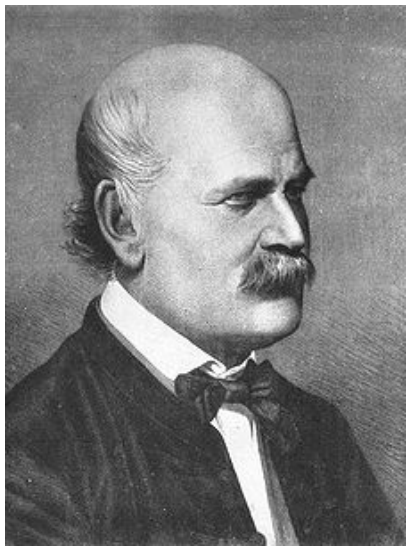
1865 – Semmelweis committed to asylum → died.

1878 – Pasteur publishes 'germ' theory of disease.

Now hand washing made sense and was adopted.

→ incidence of puerperal fever, and many other infections, dropped very quickly across Europe.

Ignaz Semmelweis



1847 – showed that washing hands between post-mortems and patients cut incidence of puerperal fever from 10-35% to < 1%.

Largely ignored by medical community because it made no sense. Puerperal fever remained common in most of Europe.

1865 – Semmelweis committed to asylum → died.

1878 – Pasteur publishes 'germ' theory of disease.

Now hand washing made sense and was adopted.

→ incidence of puerperal fever, and many other infections, dropped very quickly across Europe.

Bottom line – people only believed the evidence-based story when there was an explanation.

Here is a modern example: does the underlined phrase actually carry any scientific weight?

The Possible Harmful Biological Effects of Low-level Electromagnetic Fields of Frequencies up to 300 GHz
IET Position Statement - May 2006

- The absence of a plausible biophysical mechanism operating at environmental levels of exposure to power frequency EMFs remains a significant component in the balance of the evidence against health effects. Considerable research effort remains centred round the effect of magnetic fields on free radicals as a possible mechanism. It remains doubtful whether this mechanism could produce effects at the microtesla level implicated by the epidemiology.

Any other examples of this?

A simple coin-tossing game:

Unbiased coin, 50% probability of head or tail, keep tossing it

Ladies, you want HTT

Gentlemen, you want HTH

Over many plays of the game, who wins most? (the side that gets its combination 1st wins)

a) Ladies (ie HTT comes up, on average, faster than HTH)

b) Ladies and men score equally (HTT as fast as HTH)

c) Men score more (ie HTT comes up more slowly than HTH)

(each choose a, b or c, and pass me the paper)

A simple coin-tossing game:

Unbiased coin, 50% probability of head or tail, keep tossing it

Ladies, you want HTT

Gentlemen, you want HTH

Over many plays of the game, who wins most?

→ a) Ladies (ie HTT comes up, on average, faster than HTH)
b) Ladies and men score equally (HTT as fast as HTH)
c) Men score more (ie HTT comes up more slowly than HTH)

Reminder: ladies want HTT, gentlemen HTH

Imagine the game starts like this

H ….. ladies get excited, men get excited

Reminder: ladies want HTT, gentlemen HTH

Imagine the game starts like this

H ….. ladies get excited, men get excited

T ….. ladies get even more excited, men get even more excited

(so far, so symmetrical)

Reminder: ladies want HTT, gentlemen HTH

Imagine the game starts like this

H ….. ladies get excited, men get excited

T ….. ladies get even more excited, men get even more excited

(so far, so symmetrical)

Now, imagine what happens to each if they do *not* get what they want next...

Reminder: ladies want HTT, gentlemen HTH

Imagine the game starts like this

H ….. ladies get excited, men get excited

T ….. ladies get even more excited, men get even more excited

(so far, so symmetrical)

Now, imagine what happens to each if they do *not* get what they want next...

Men get T: sadly, they have to wait for the next H before getting excited again.

Reminder: ladies want HTT, gentlemen HTH

Imagine the game starts like this

H ….. ladies get excited, men get excited

T ….. ladies get even more excited, men get even more excited

(so far, so symmetrical)

Now, imagine what happens to each if they do *not* get what they want next...

Men get T: sadly, they have to wait for the next H before getting excited again.

Ladies get H: they lose this time but can at least be excited that they are *already* starting again with this H.

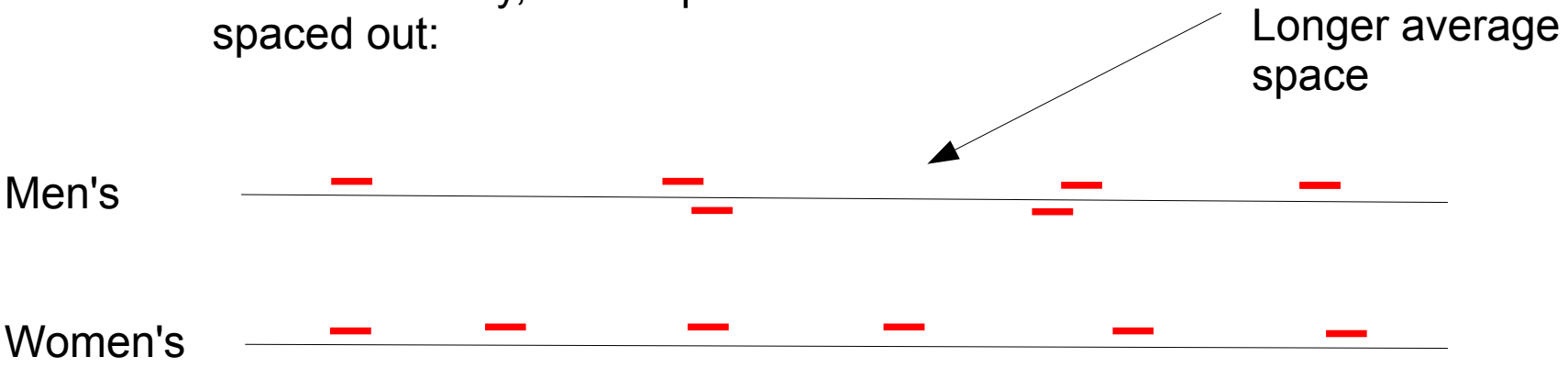
With thanks to Peter Donnelly, Univ Oxford, for this idea.

There's also another way of looking at it:

The men's pattern, HTH, overlaps itself.

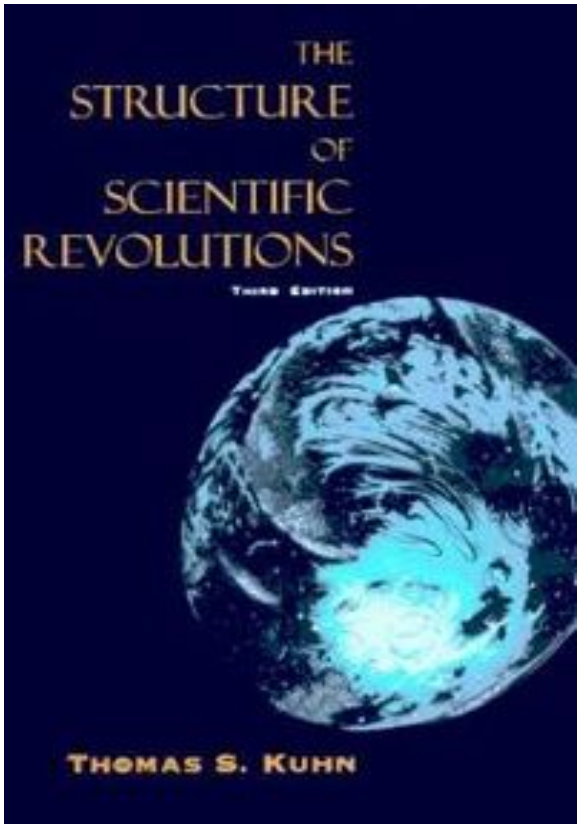
You could get TTTHTHTHTTT

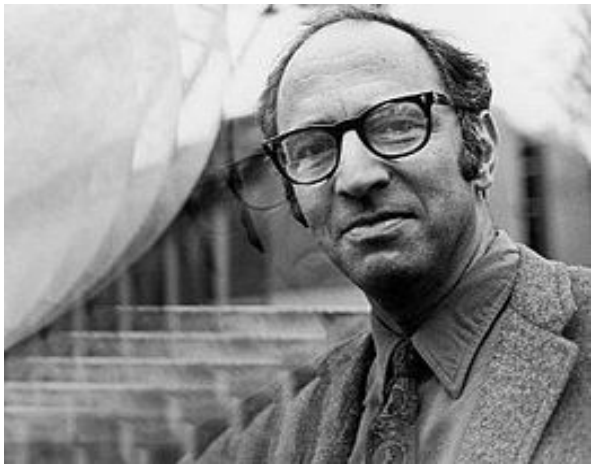
So, with an equal number incidence of a given triplet (in any 'reading frame') in a long sequence, since the men's sequence can self-overlap and cluster that way, this sequence must be more spaced out:



Bottom line: be VERY careful to check that 'obviously true' things about probability are actually true!

Part 3: the progress of science.



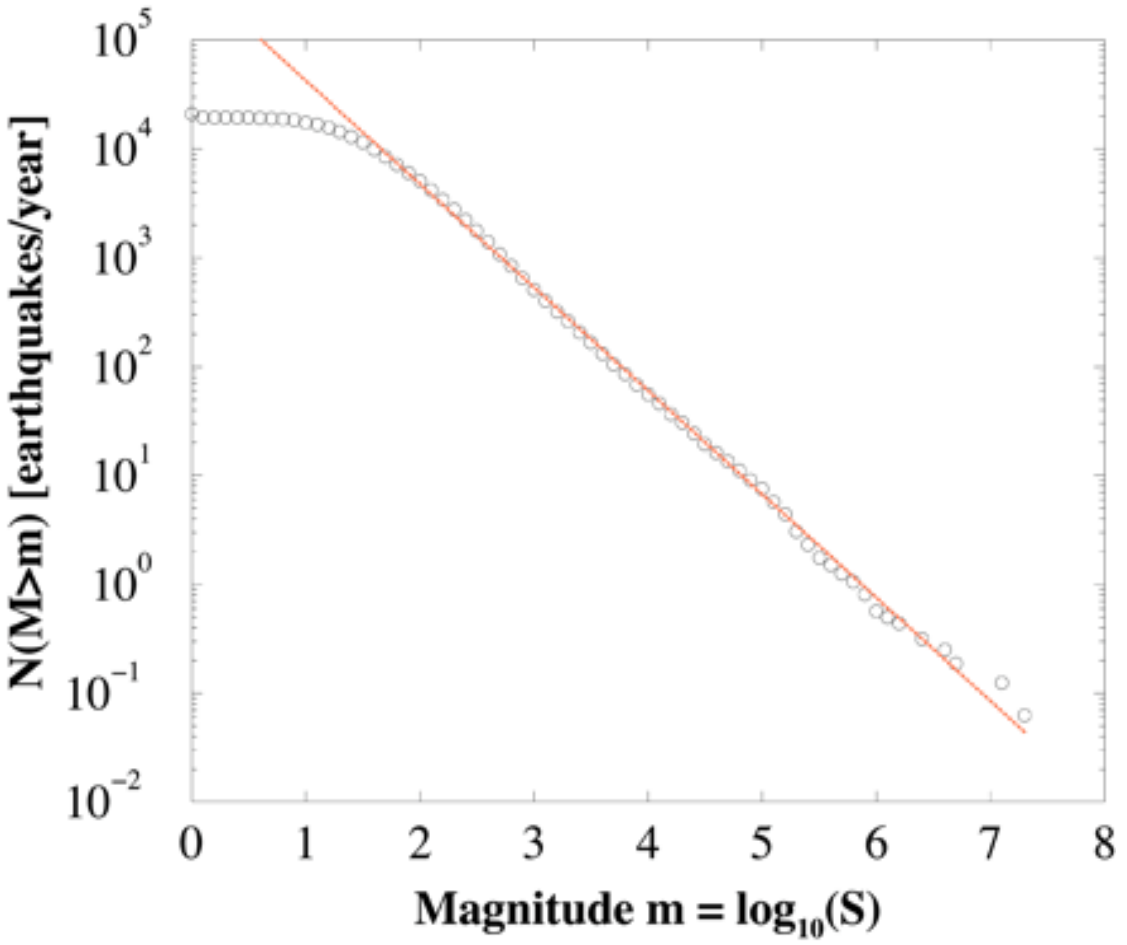


Thomas Kuhn

“Paradigm shifts”

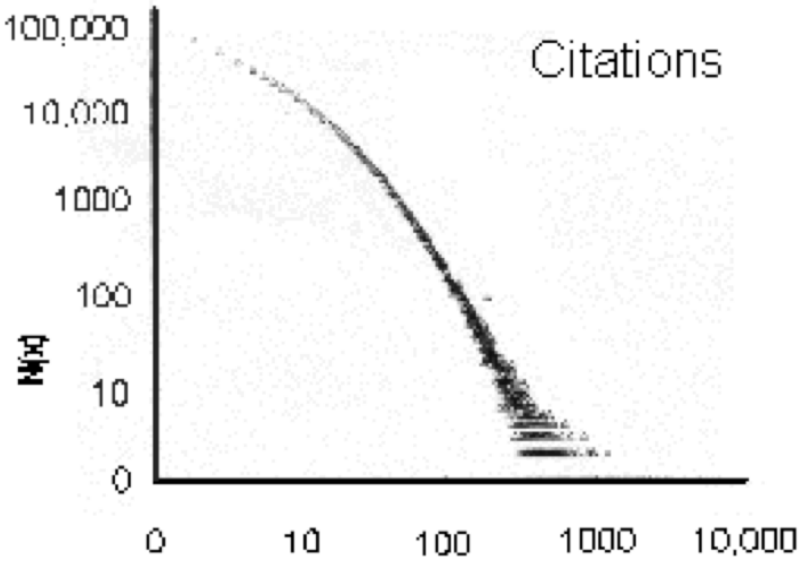
Examples?

Impact, and earthquakes



Impact, and earthquakes

Science:



Eur. Phys. J. B **27**, 445–447 (2002)
DOI: 10.1140/epjb/e2002-00176-y

THE EUROPEAN
PHYSICAL JOURNAL B

The individual success of musicians, like that of physicists, follows a stretched exponential distribution

J.A. Davies[a]

University of Edinburgh, Teviot Place, Edinburgh EH8 9AG, UK

Received 23 October 2001
Published online 25 June 2002 –

Abstract. Over the last five years or so, a number of studies have focussed on the distribution of 'success' in physics and other sciences; in these studies, 'success' is measured by the number of times a paper, or an author, is cited. The distribution of citations of individual papers approximates to a power-law [S. Redner, Eur. Phys. J. B **4**, 131 (1998)], while lifetime total citations of the 1120 most-cited physicists follows a stretched exponential [J. Laherrère, D. Sornette, Eur. Phys. J. B **2**, 525 (1998)]. Here, I examine the distribution of success in popular music, a field of creativity that has social structures very different from those of physics, and which is generally held to be controlled primarily by fashion. For this study, the lifetime total success of bands was measured by the total number of weeks they were in the weekly 'top 75' list of best-selling recordings. Like the lifetime success of physicists reported by Laherrere and Sornette, the success of the 6107 bands that appeared in the UK 'top 75' from 1950 until 2000 follows a stretched exponential of the form $P(x)\mathrm{d}x = c(x^{c-1}/x_0^c)\exp[-(x/x_0)^c]\mathrm{d}x$; for the music data, $c = 0.5$ and $x_0 = 9.37$.

PACS. 43.75.+a Music and musical instruments – 01.30.-y Physics literature and publications – 87.23.-n Ecology and evolution

→ success in science follows the same pattern as in rock music

The contribution of different organisms also follows this kind of distribution:

BMC Developmental Biology

BioMed Central

Open Access

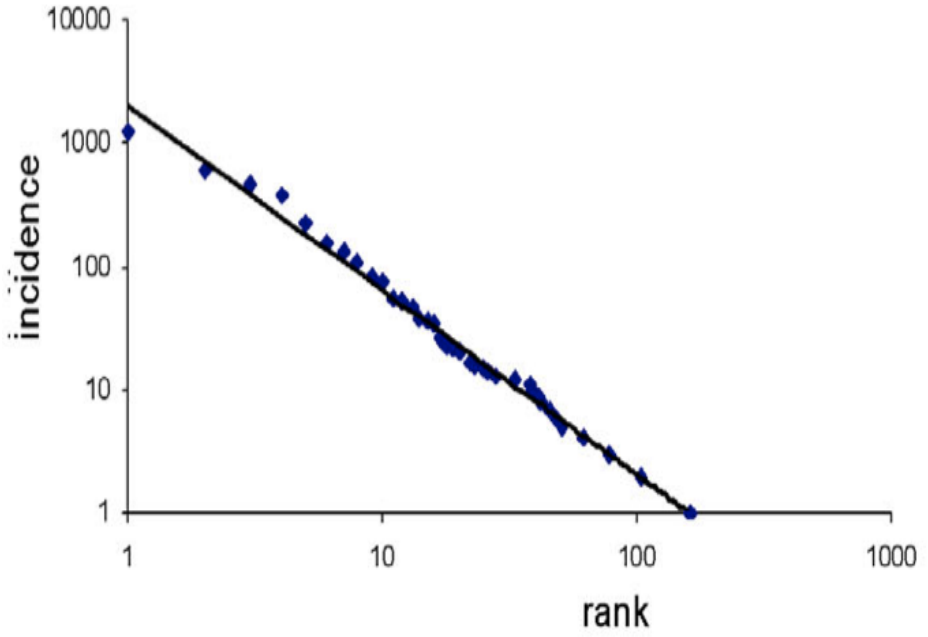
Developmental biologists' choice of subjects approximates to a power law, with no evidence for the existence of a special group of 'model organisms'

Jamie A Davies*

Address: University of Edinburgh Centre for Integrative Physiology, Edinburgh, UK

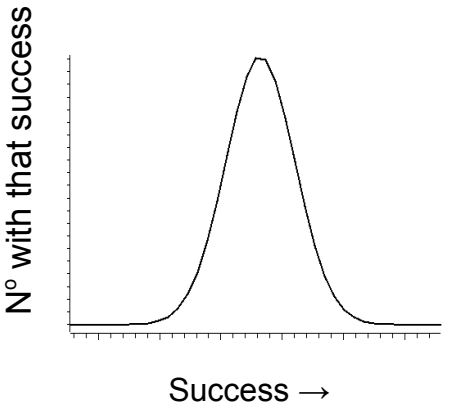
Email: Jamie A Davies* - jamie.davies@ed.ac.uk

* Corresponding author



Mediocristan: (Gaussian; non-scaleable)

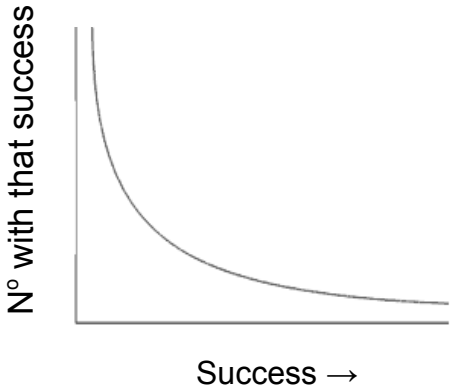
Baker,
Cab driver,
Miner,
Street performer,
Shoe-maker,
Surgeon...



Nassim Nicholas Taleb

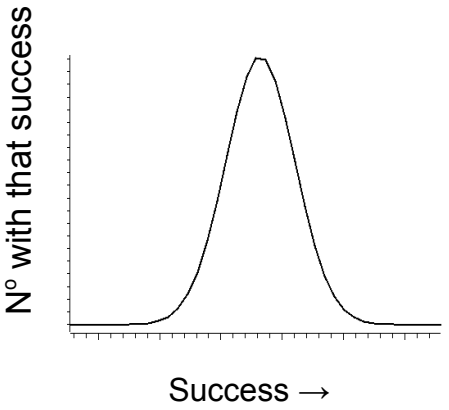
Extremistan: (Power law; scaleable)

Author,
Artist,
Software engineer,
Stock trader,
Rock musician,
Scientist,
Movie star...



Mediocristan: (Gaussian; non-scaleable)

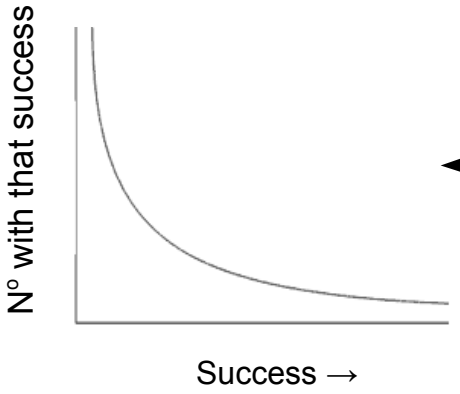
Baker,
Cab driver,
Miner,
Street performer,
Shoe-maker,
Surgeon...



Nassim Nicholas Taleb (Photo: Time)

Extremistan: (Power law; scaleable)

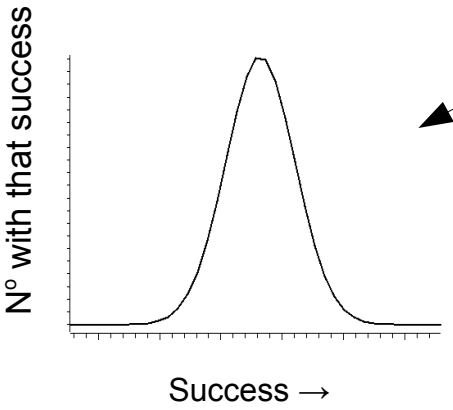
Author,
Artist,
Software engineer,
Stock trader,
Rock musician,
Scientist,
Movie star...



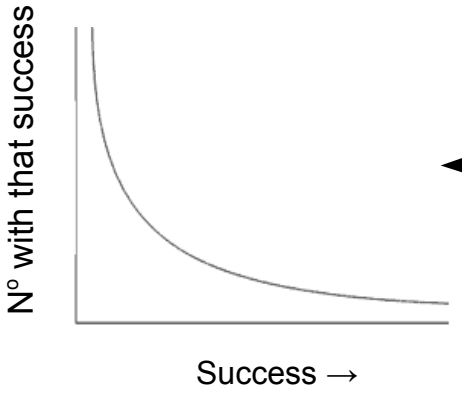
The impact of one particular loaf, journey, lump of coal, mime, boot or appendectomy is always a tiny fraction of the whole success.

The impact of one particular book, painting, application, trade, album, paper or film can dwarf the rest.

By the way, scientists often assume all distributions are like this.

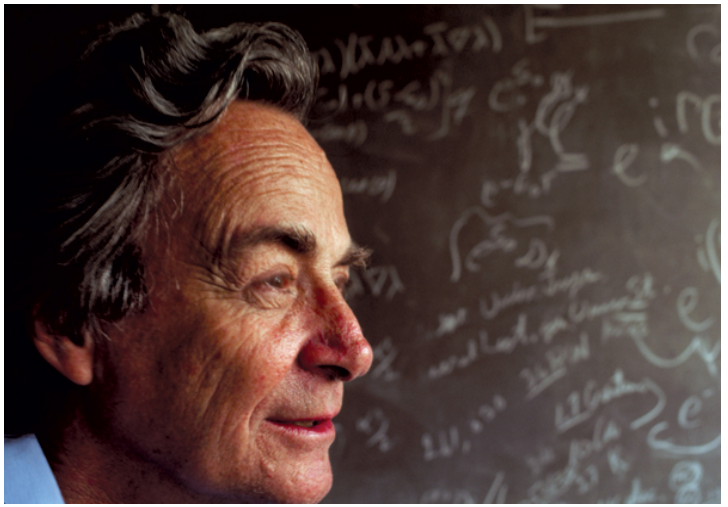


Many really interesting ones are like this, ESPECIALLY in biology because natural selection itself is like this.



→ Don't blindly assume Gaussian distribution or you will mess up.
Don't blindly assume your supervisor knows this.

Unscientific thinking



(A parable from Richard Feynmann)

Unscientific thinking

Many medical 'facts' are established by committee

How many times have you heard phrases such as 'most scientists believe that...' in the context of contentious areas (global warming, food safety, healthy eating....)?

Undergrad → Research transition: losing the fear of failure.

I have not failed. I've just found 10,000 ways that won't work. ~Thomas Edison

Try again. Fail again. Fail better. ~Samuel Beckett

Failure is only the opportunity to begin again more intelligently. ~Henry Ford

There is no failure. Only feedback. ~Robert Allen

Success is stumbling from failure to failure with no loss of enthusiasm. ~Winston Churchill

Failures are fingerposts on the road to achievement. ~C.S. Lewis

Success represents the 1% of your work which results from the 99% that is called failure. ~Soichiro Honda

Some final quotations to ponder.

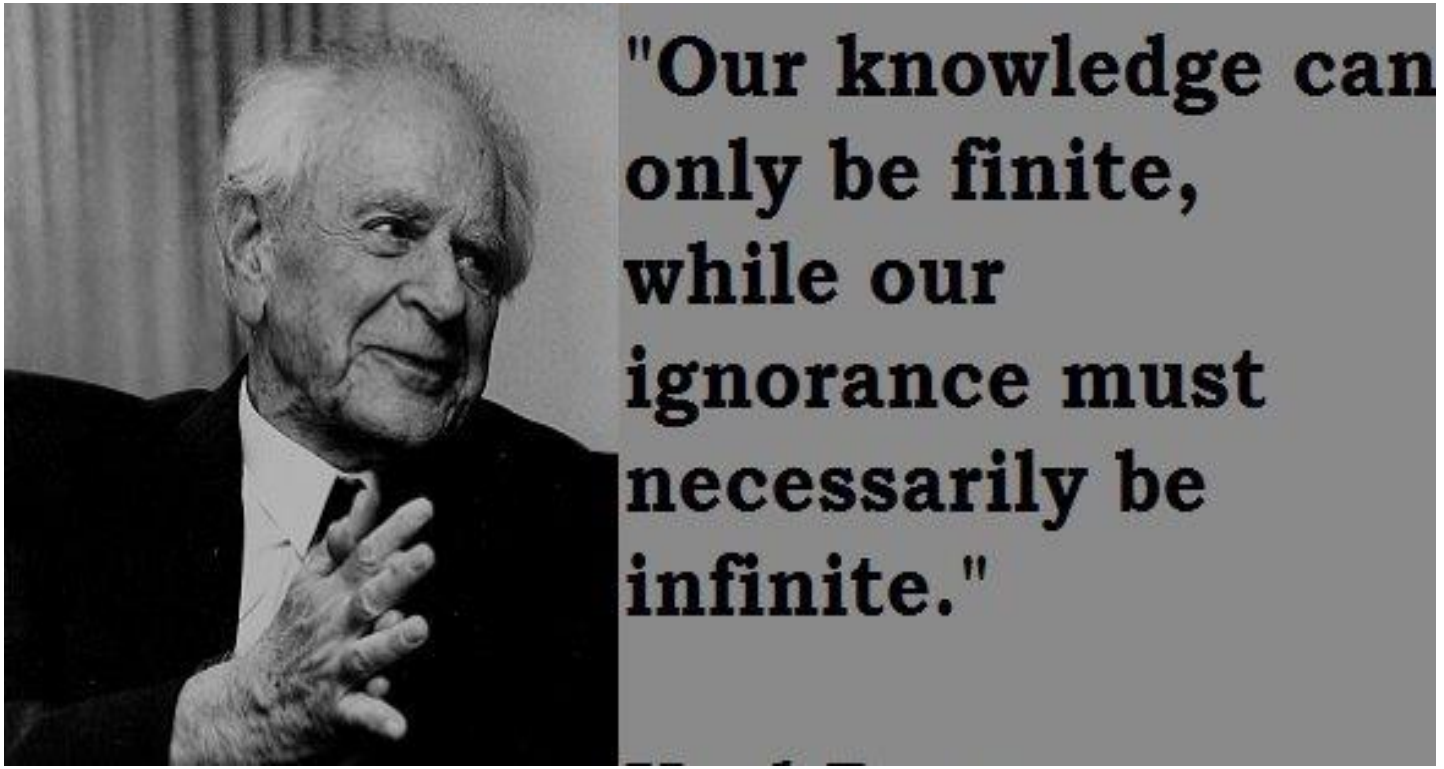
Steve Jones

“Science is the refuge of the mediocre … however pedestrian is what you do, you still add one more piece to the sum of knowledge”.

Speaking in *The Life Scientific*, Radio 4, 2012

Ada Lovelace

I never am really satisfied that I understand anything; because, understand it well as I may, my comprehension can only be an infinitesimal fraction of all I want to understand...



Karl Popper

Further reading: http://www.slideshare.net/guest0df09b/scientific-thinking-dc