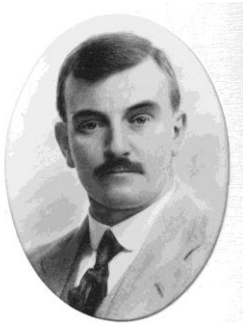
SAIMM
THE SOUTHERN AFRICAN INSTITUTE
OF MINING AND METALLURGY

Truth and Error in Scientific Publishing

Rodney Jones

13 August 2015

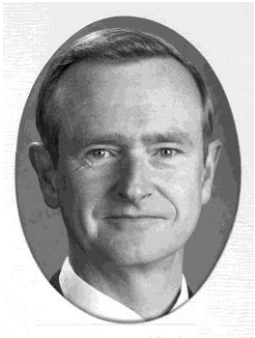


1914 – 1915
GH Stanley

1961 – 1962
DG Maxwell

1971 – 1972
DD Howat

1975 – 1976
RE Robinson



1980 – 1981
PR Jochens

1978 – 1979
MC Atmore

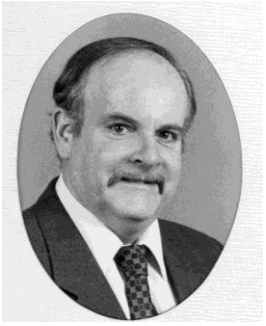
CAPACI OCCASIO
FOUNDED 1894

SAIMM
THE SOUTHERN AFRICAN INSTITUTE
OF MINING AND METALLURGY

MINTEK

1983 – 1984
RP King

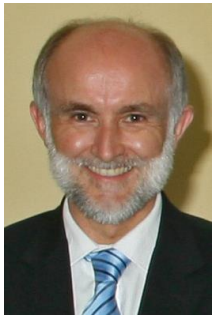
1985 – 1986
HE James



1996 – 1997
NA Barcza

2002 – 2003
SJ Ramokgopa

2008 – 2009
AM Garbers-Craig



2015 – 2016
RT Jones

SAIMM
THE SOUTHERN AFRICAN INSTITUTE
OF MINING AND METALLURGY

Truth and Error in Scientific Publishing

Rodney Jones

13 August 2015

Introduction

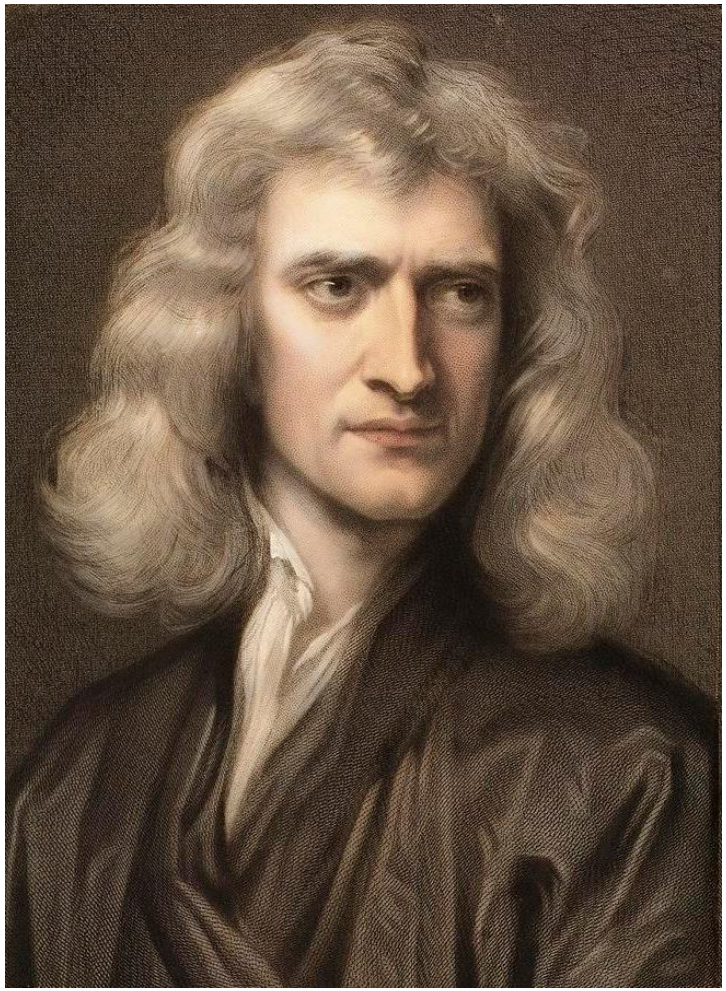
- Some current controversies:
  - Climate change
  - Diet

- These topics rely on scientific evidence to get closer to the truth
- Scientific progress relies on the publication of ideas and experimental results that can be replicated, tested, and improved over time

Truth and Knowledge

- The wisest of men is he who knows that he knows nothing – Socrates
- All men by nature desire knowledge – Aristotle
- If you would be a real seeker after truth, it is necessary that at least once in your life you doubt, as far as possible, all things – Rene Descartes
- All truths are easy to understand once they are discovered; the point is to discover them – Galileo Galilei
- In questions of science, the authority of a thousand is not worth the humble reasoning of a single individual – Galileo Galilei
- All truth passes through three stages. First, it is ridiculed. Second, it is violently opposed. Third, it is accepted as being self-evident. – Arthur Schopenhauer

Newton’s ‘undiscovered ocean of truth’

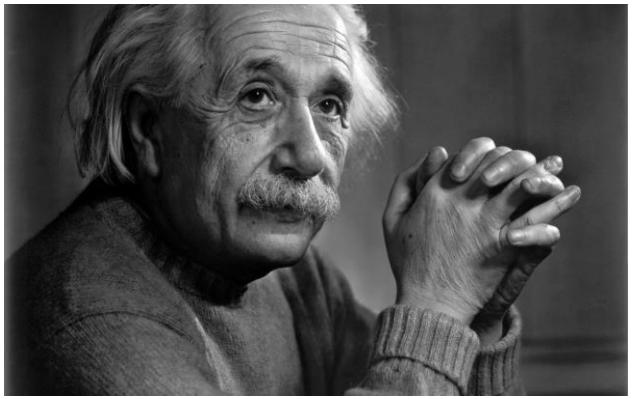


- Isaac Newton said: "I do not know what I may appear to the world, but to myself I seem to have been only like a boy playing on the seashore, and diverting myself in now and then finding a smoother pebble or a prettier shell than ordinary, whilst the great ocean of truth lay all undiscovered before me."

Epistemology (the study of knowledge)

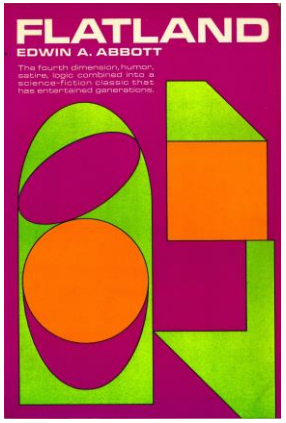
- Knowledge is justified true belief (not merely a lucky guess)
- Knowledge of the truth can be obtained by *a priori* reasoning, or by empirical experience

- "Science without epistemology is - insofar as it is thinkable at all - primitive and muddled. However, no sooner has the epistemologist, who is seeking a clear system, fought his way through such a system, than he is inclined to interpret the thought-content of science in the sense of his system and to reject whatever does not fit into his system. The scientist, however, cannot afford to carry his striving for epistemological systematic that far. ... He therefore must appear to the systematic epistemologist as an unscrupulous opportunist." (Einstein, 1949)



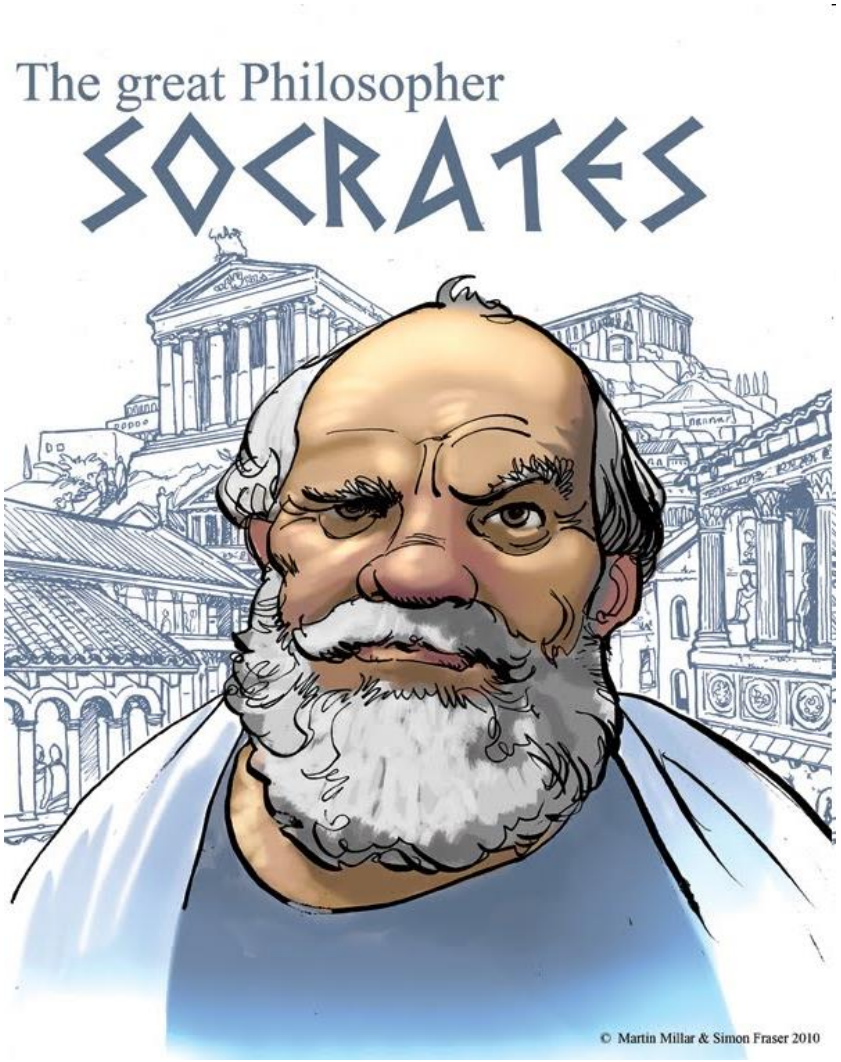
What is truth?

- **Solipsism** – denial of reality
- **Radical scepticism** – only direct knowledge of our senses
- **Relativism** – depends on point of view
  - "Everything we hear is an opinion, not a fact. Everything we see is a perspective, not the truth." – Marcus Aurelius
- **Instrumentalism** – quantities can be measured, even if we have no way of knowing whether theoretical entities actually exist
- **Fallabilism** – can know things even though we are sometimes wrong
- **Empiricism** – based on observation or experience
- **Rationalism** – based on reason
- **Dogmatism** – certain of their truth
  - "Certitude is not the test of certainty" – Justice Oliver Wendell Holmes, Jr

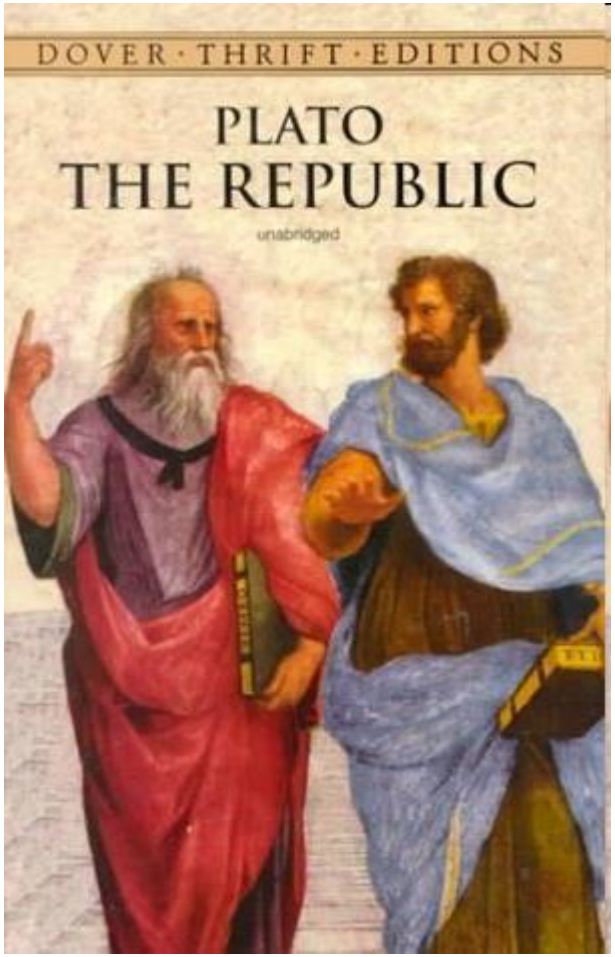


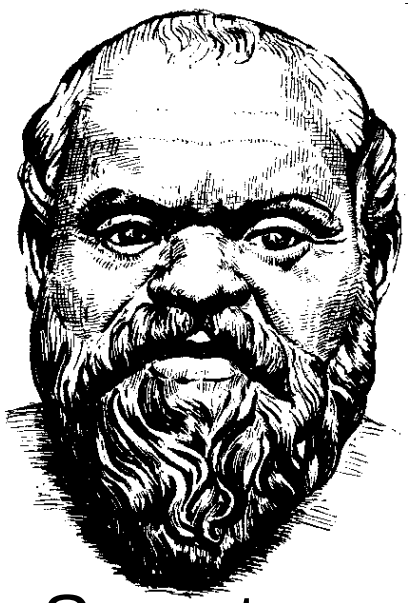
Socrates

- Socrates lived from 469 – 399 BC
- The Socratic method of teaching involved asking questions
- The great philosopher, Socrates, did not record his own words
- Plato recorded the teachings of Socrates

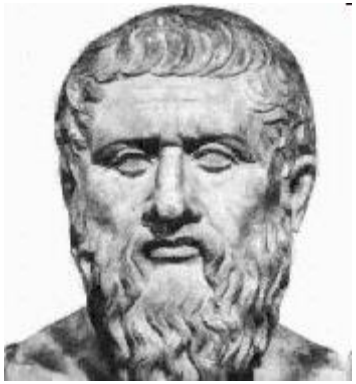


Socrates and Plato

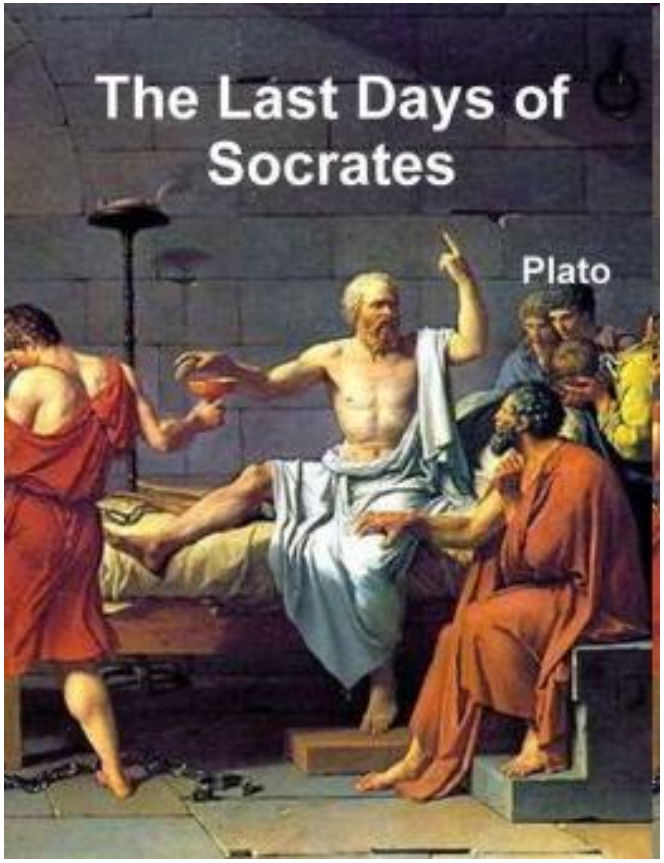




Socrates



Plato

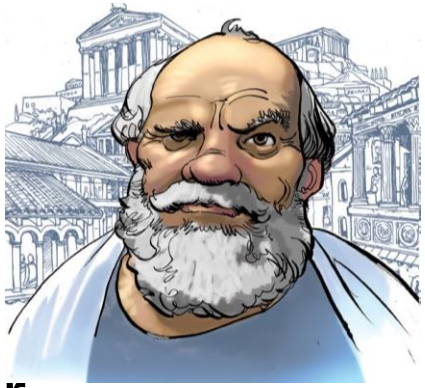


Socratic Method

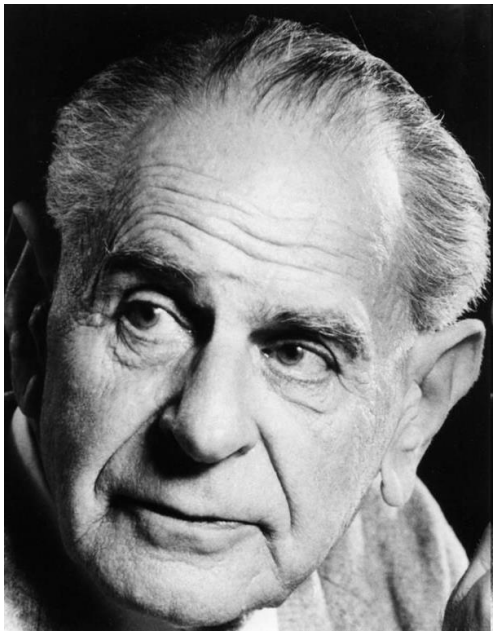
- The Socratic method led to the development of the scientific method

Scientific method and attitude

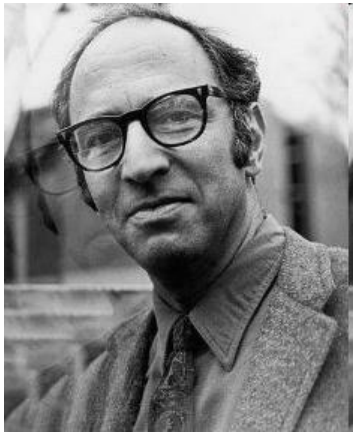
- Socrates



- Karl Popper
  - falsification and falsifiability
    - observation, hypothesis formulation, testing



- Thomas Kuhn
  - scientific revolutions by paradigm shifts
    (*e.g.*, Darwin, Marx, Einstein)



Scientific method and attitude

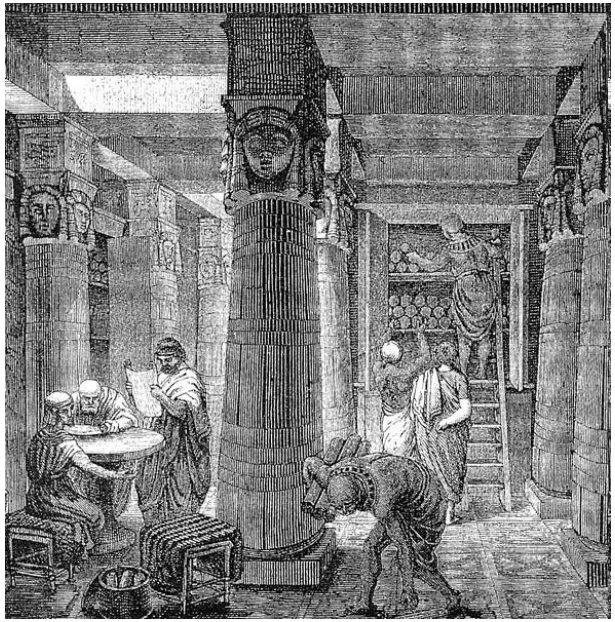
- Observations of the world lead to the recognition of patterns (through inductive reasoning) which lead to interesting questions about why the pattern might occur
- Hypotheses are formed in such a way that they are framed as testable questions. The scientific method cannot be applied to untestable, unfalsifiable questions.
- Find a number of possible explanations and think up ways of testing which ones might be wrong
- Data is gathered (and repeated) to test the prediction. The hypothesis is then rejected, accepted, or refined and re-tested.
- Experiments should be replicated reproducibly
- Use control groups and double-blind protocols where appropriate
- If two ideas explain the data equally well, the simpler one is preferred. Occam's razor suggests that the simpler theory with fewer (or less onerous) unproved assumptions is probably the most appropriate one.
- General theories then gather together the hypotheses that are consistent with all current data. They remain provisional and tentative.

Scientific method and attitude

- Scientific theories cannot be proven (only corroborated), but can be shown to be beyond reasonable doubt
- The sun will rise tomorrow. But this does not imply that it will rise ten billion years from now (by which time its fuel will most likely have been exhausted).

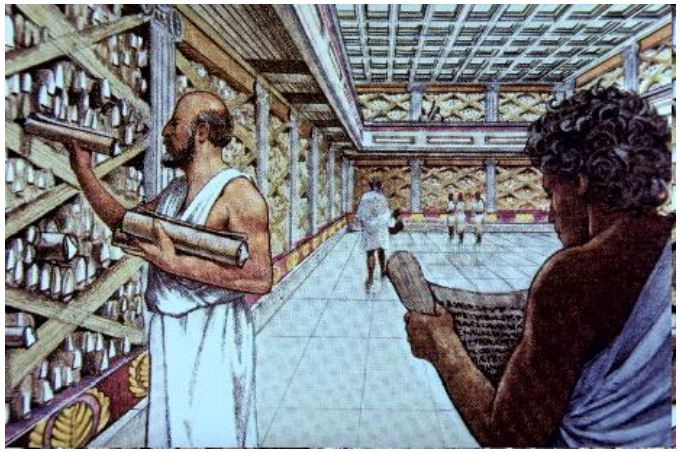
Library of Alexandria

- From ~280 BC to a few hundred years later

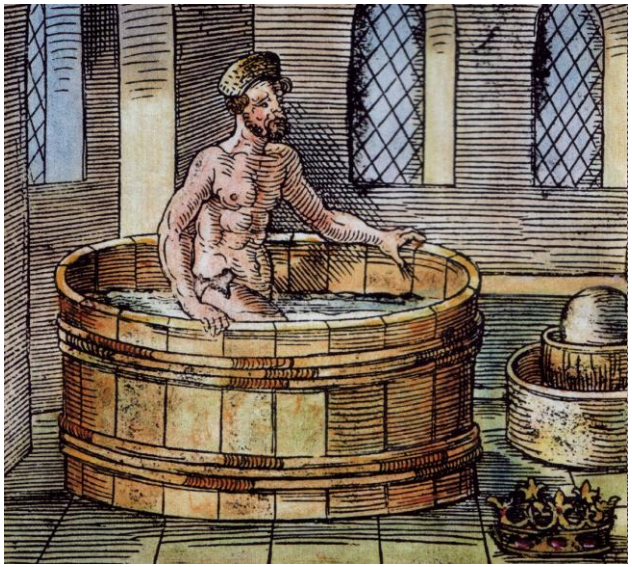


Library of Alexandria

- The Library of Alexandria was the centre of collected knowledge in the ancient world
- Scholars such as Euclid and Archimedes are said to have written, studied, and experimented there

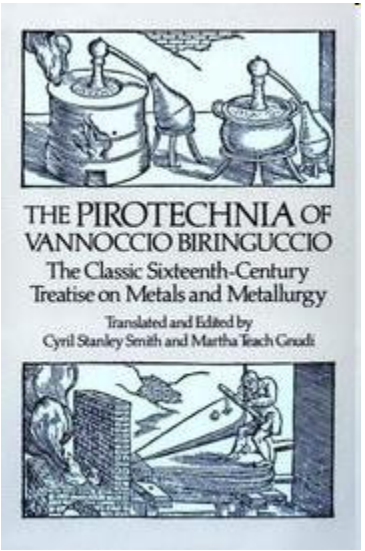


Euclid



Archimedes

De la Pirotechnia – Vannoccio Biringuccio – 1540



- The tradition of scientific and technical literature in the field of metallurgy started a long time ago

- De la Pirotechnia, by Vannoccio Biringuccio, was the first printed book on metallurgy to have been published in Europe
- Written in Italian and published in Venice in 1540

De la Pirotechnia – Vannoccio Biringuccio – 1540

DE LA PIRO-
TECHNIA.

LIBRI, X. DOVE AMPIAMEN
te si tratta non solo di ogni sorte & di,
uersita di Miniere, ma anchora quan
to si ricerca intorno à la prattica di
quelle cose di quel che si appartiene
à l'arte de la fusione ouer gitto de me
talli come d'ogni altra cosa simile à
questa. Composti per il. S. Vanoc-
cio Biringuccio Sennese.

Con Priuilegio Apostolico & de la
Cesarea Maesta & del Illustris. Sena
to Veneto, M D XL.

L. S. DELA P.

do generato in pietra spognosa dela sorte che ho detto disopra, si ma-
cina con pistoni benissimo, o con mulino da vitulera che la schiacci, &
dipoi si laua. Ma seglie in pietra ben cõmista, sonno alcuni che fanno
vna stanzetta piccola con vna volta abbotte, ouero a catino nõ di mol
ta grandezza, ma benissimo intonicato, & disotto fanno vno spazzo
che penda alquanto in verso il mezzo, nettamente murato, & dacito
dentro alla grossezza del muro, fanno vno o piu fornelli in tal modo
adiutati, che le bocche per lequali s'ha da dar fuocho venghin fuor de
la stanza & sopra alli fornelli vi muran dentro tanti pignati quãti ne
possan cõtenere, & ghiempino di detta miniera in poluere, o redutta in
mo'tri pezzetti. Et dipoi infra la volta dela stanza & lo spazzo fanno
vno infrascato di frasche darbori verdi & serran bene la fine stretta &
lentrata che niente ne possa respirare, & dipoi dan fuocho alli forni, &
cosi il mercurio sentendo il caldo del fuocho come suo cõtrario vnol
fuggire, & euaporando saglie & esce fuor dela boccha de vasi, qualeg
sentire certa freschezza che porgen le foglie di quelle frasche a esso cõ
forme corre e esce & vi sattacha sopra. Per il che dipoi quando il pra-
tico artifice pensa o crede che dela sua materia che misse ne pignati la
sustantia del mercurio sia tutta vscita, lassa spegnare il fuocho & istuto
benissimo refreddare, & dipoi entra i detta stanza, & anchor chel mer
curio per la sua ponderosita da perse dele frasche doue e attaccato ca-
schi buona parte nelo spazzo scrullano le predette frasche, & quel che
non fusse caschato il fan cascare, & dipoi nettamente dalo spazzo il
ricoglieno, & p questo modo van cõtinuando p fin che hãno miniera,

Alcuni altri sonno che lo istraggano con manchi o trauaglio cõ pignatte
di terra grãdi cõmesse che cõmettino le bocche luna nel latra, & empi
ta la pignatta maggiore di miniera vi metteno sopra vn pian d'arena,
ouer di cennere staccata, accioche sagliẽdo largẽto viuo spento dal fuo
cho nel laltra pignatta sopra a posta tornando indrieto congionto in
goccie

DEL LARGENTO VIVO 25

goccie come acqua caschi sopra a tal cennari, ouer arene, & fredda che
sara la pignatta. dipoi lauandolo facilmente doue il sia si ricupera.

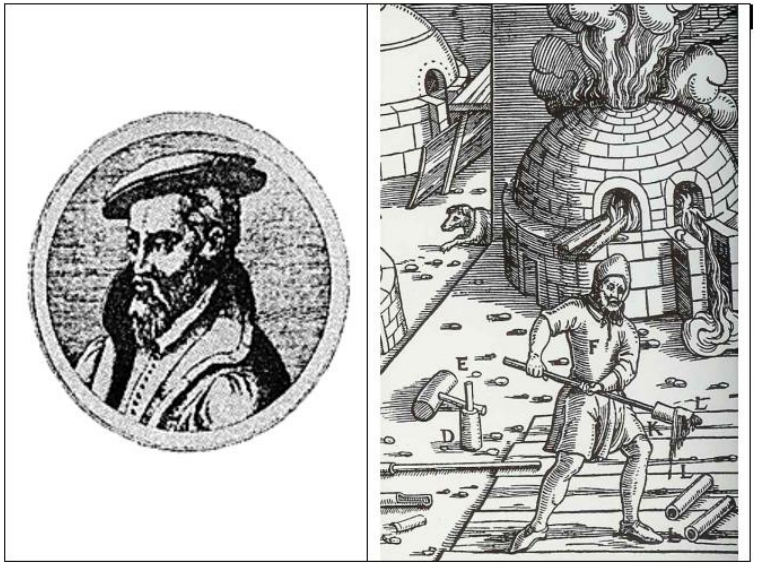
Alcuni altri sonno che in scambio di pignatti fan certi vasi di terra lar-
ghi in boccha, & stretti in fondo come son le forme daffar gli zucchari,
alquali fanno vn coperchio cõmesso grosso vn dito o piu di terra & il
vetriano dentro, & questo tal vaso empieno di miniera pesta o trita be
nissimo, & sopra copreno de vn dito o due di cennare staccata, & so-
pra col coperchio serrano benissimo il vaso legandolo, ouero cõ qual-
che cosa graue che gli calchi disopra con aggrauãdolo, & dipoi mette
fuocho al fornello done sonno adiutati a star dentro detti vasi, & cosi
per sentire il caldo il mercurio esce dela miniera & saglie per voler eua
porare, & percotendo nelli coperchi, cascha infra le cennari come nel
laltro modo sopradetto, da lequali come hauete intesо, lauandole o cõ
staccio fitto staccandolo tuto si ricupera.

Alcuni altri m'han detto hauer veduto mettere in scambio di tal co-
perchio & cennare vn vaso simile a quello che si chiama campana da

D

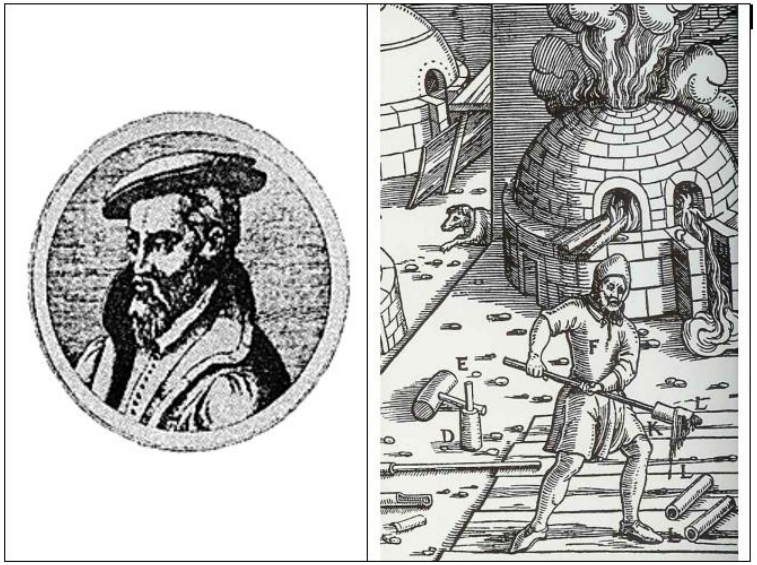
- Provides details on mining practice, the extraction and refining of numerous metals, and the production of alloys such as brass

De Re Metallica – Georgius Agricola – 1556



- Georgius Agricola (1494-1555) wrote ‘of things metallic’ or ‘on the nature of metals’
- Travelled extensively and wrote about mining methods and metallurgical processes that were in place in the Middle Ages
- Original Latin version was published in 1556, the year after Agricola died

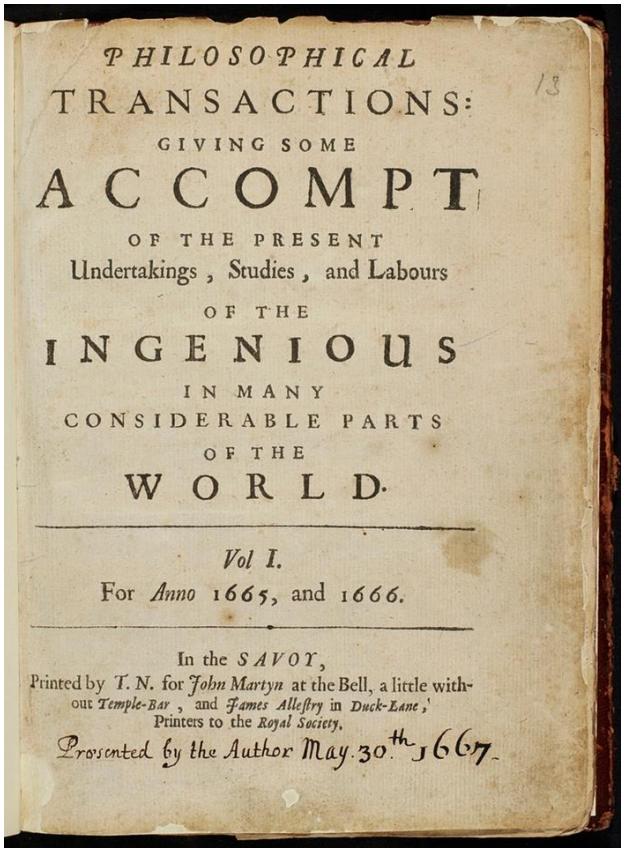
De Re Metallica – Georgius Agricola – 1556



- The English version of De Re Metallica was published in 1912 - translated into English by Herbert Hoover, a mining engineer and later President of the United States, and his wife Lou Henry Hoover, a geologist and classicist
- “I have omitted all those things which I have not myself seen, or have not read or heard of from persons upon whom I can rely. That which I have neither seen, nor carefully considered after reading or hearing of, I have not written about.”

Philosophical Transactions of the Royal Society

- Scientific publishing of journal papers has been in existence for 350 years
- The first issue of Philosophical Transactions of the Royal Society was published 350 years ago, under the guidance of Henry Oldenburg

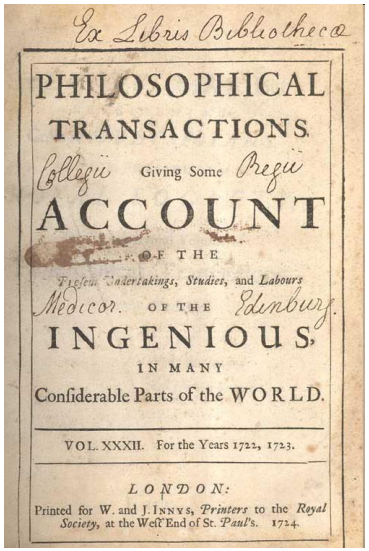


PHILOSOPHICAL
TRANSACTIONS:
GIVING SOME
ACCOMPT
OF THE PRESENT
Undertakings, Studies, and Labours
OF THE
INGENIOUS
IN MANY
CONSIDERABLE PARTS
OF THE
WORLD.

Vol I.
For *Anno* 1665, and 1666.

In the *SAVOY*,
Printed by *T. N.* for *John Martyn* at the Bell, a little without *Temple-Bar*, and *James Allestry* in *Duck-Lane*,
Printers to the *Royal Society*.
Prosented by the Author May. 30th 1667

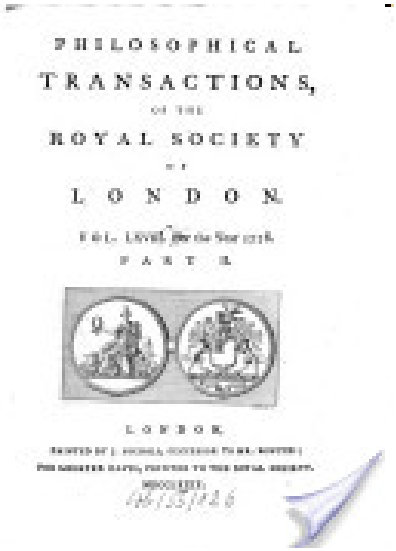
Philosophical Transactions of the Royal Society


Ex Libris Bibliotheca

PHILOSOPHICAL
TRANSACTIONS
Giving Some
ACCOUNT
OF THE
Ingenious
IN MANY
Considerable Parts of the WORLD.

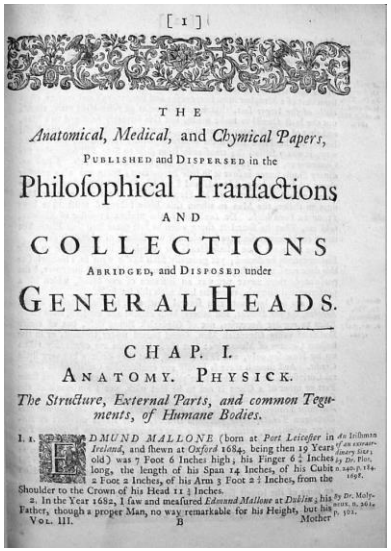
VOL. XXXII. For the Years 1722, 1723.

LONDON:
Printed for W. and J. Innys, *Printers* to the *Royal Society*, at the West End of St. *Paul's*. 1724.


PHILOSOPHICAL
TRANSACTIONS,
OF THE
ROYAL SOCIETY
OF
LONDON.

PART II.

LONDON.


[1]

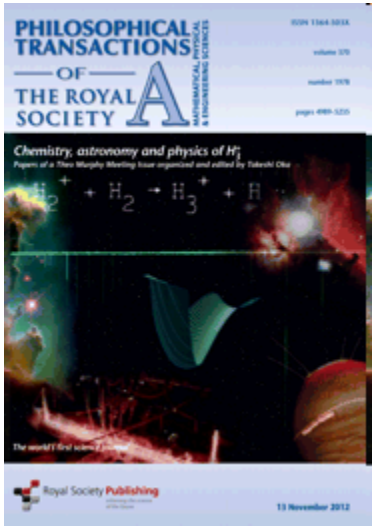
THE
Anatomical, Medical, and *Chymical Papers,*
PUBLISHED and DISPERSED in the
Philosophical Transactions
AND
COLLECTIONS
ABRIDGED, and DISPOSED under
GENERAL HEADS.

CHAP. I.
ANATOMY. PHYSICK.
The Structure, External Parts, and common Teguments, of Humane Bodies.

1. EDMUND MALLONE (born at *Port Leicester* in
Ireland, and shewn at *Oxford* 1684, being then 19 Years
old) was 7 Foot 6 Inches high; his Finger 6 ½ Inches
long, the length of his Span 14 Inches, of his Cubit
a Foot 2 Inches, of his Arm 3 Foot 2 ½ Inches, from the
Shoulder to the Crown of his Head 11 ½ Inches.
2. In the Year 1682, I saw and measured *Edmund Mallone* at *Dublin*, his
Father, though a proper Man, no way remarkable for his Height, but his
Vol. III. B Mother

- The world's oldest and longest-running scientific journal, the Philosophical Transactions of the Royal Society, was first published in March 1665, in London
- Henry Oldenburg – Secretary of the Royal Society and first Editor of the publication – ensured that it was 'licensed by the council of the society, being first reviewed by some of the members of the same', thus making it the first ever reviewed journal

Philosophical Transactions of the Royal Society



- The Royal Society announced in October 2011 that they had made the historical archives of the Philosophical Transactions (over 60 000 scientific papers) permanently free to online access from anywhere in the world

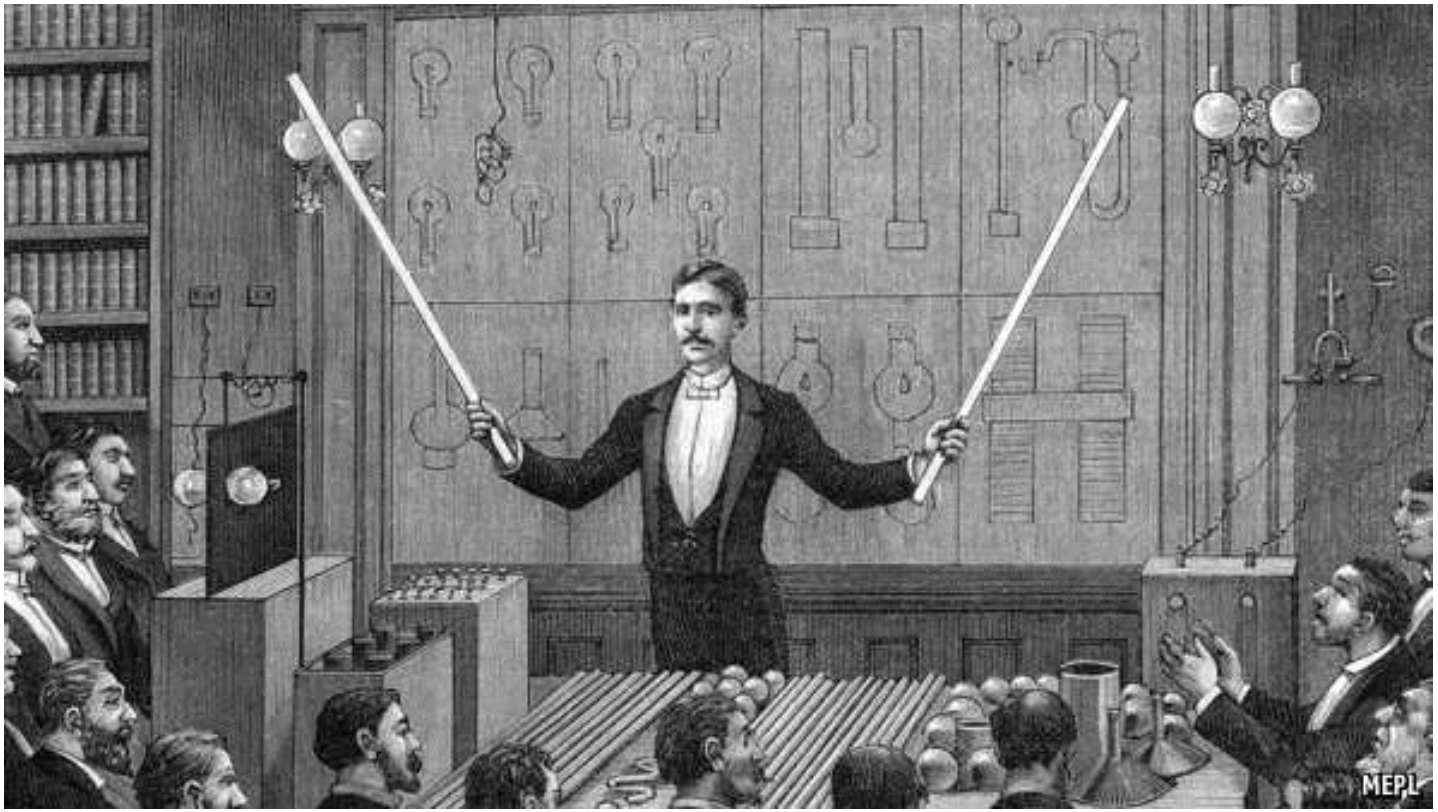
- All of the historical archival papers (published more than 70 years ago) from the Philosophical Transactions are now freely available on their website

- Current publications are available via delayed open access where older articles (12 months for biological sciences, and 24 months for physical sciences) are made freely available

THE ROYAL SOCIETY

Early days of scientific societies

- In the early days, regular meetings were held to discuss science and run experiments. The reading of scientific papers took place, and publication of papers and discussions was done to record the proceedings of meetings. Meetings often included rather robust debate.
- The picture below shows Nikolai Tesla giving a demonstration in 1893



http://www.economist.com/news/science-and-technology/21656134-electronics-has-already-cut-data-cord-can-it-now-cut-power-cord?fsrc=nlw|hig|26-06-2015|

Notes from 1894

- “As Chemists and Metallurgists of the Rand you have before you much useful and interesting work, and it remains with you, gentlemen, by publication or diffusion of accurate scientific information, by exposure of pseudo-scientific frauds, …, to claim as a right the recognition of your proper status in relation to this community”
- “I heard a rumour about a certain company getting an actual extraction of 125 per cent. from concentrates”
- “Such results as I have instanced are obviously imagined, or as chemists say, ‘cooked’”

  – Mr William Bettel, The President’s Inaugural Address, Chemical and Metallurgical Society of South Africa, 19 May 1894, pp.9,10,11

Robust discussion in 1895

- “As for Mr Suckling’s process, I fail to see the novelty or usefulness of the method. The use of a blast (pressure) instead of a draught of air (suction) is, to my mind, a retrograde movement, and absurd from its manifest conditions.”
  - Mr William Bettel, Monthly meeting, Chemical and Metallurgical Society of South Africa, 16 February 1895, p.93.

- “I have examined the Sulman process from both theoretical and practical points of view, and I can only come to the conclusion that it is not a bromination process, neither is it a cyanide process, but that it is a very bad oxidation process, consequently useless.”
  - Mr Schlunde, Monthly meeting, Chemical and Metallurgical Society of South Africa, 18 May 1895, p125.

The changing nature of scientific societies

- In today's world, there is a plethora of publications, and it is close to impossible for anyone to keep up with the vast flow of information. International conferences with hundreds of presentations have taken the place of the local meetings that used to discuss a single paper or experiment. In this frenetic environment, it is essential that researchers are able to trust the material they read.

- Learned societies nowadays exist to promote an academic discipline or profession, and are mostly not-for-profit organizations. They typically hold conferences for the presentation and discussion of new research results, and publish or sponsor academic journals in their discipline. The system of peer review (significantly flawed, but the best we have) is used to maintain standards and to improve the quality of papers.

- Some learned societies continue to publish journals themselves, while others have contracted this job to commercial publishing companies

- In recent times, the Internet and the World Wide Web devised by Tim Berners-Lee have transformed the dissemination of knowledge, a capacity once exclusive to publishers

Internet and information explosion

> Between the birth of the world and 2003, there were five exabytes of information created. We [now] create five exabytes every two days.
>
> - Google CEO Eric Schmidt, 2010

Google

Electronic publishing

- Recent developments in electronic publishing on websites make it possible to disseminate information more widely and cost effectively than before

- Professional societies have an opportunity to serve their members and their industry by publishing high-quality peer-reviewed papers on their websites as well as in printed form

- Older publications can be scanned, and optical character recognition can be used to provide searchable text

Number of journals and papers

- Science has grown exponentially since the late 1600s, both in terms of number of researchers and publications
- The size of science has increased by an estimated five orders of magnitude in three centuries. “We can say that 80 to 90 percent of all the scientists that have ever lived are alive now".
- An estimated 1.35 million scientific journal papers were published in 23 750 refereed journals in 2006. (Perhaps there are about two million papers per year now.) The total number of active scholarly journals, refereed plus non-refereed, was 60 911. There were 2690 open access scholarly journals, including 1735 that were also refereed
- *PLoS ONE* (paid open access) published 6749 papers in 2010, making it the world's largest journal
- In 1983, China produced just 0.6 percent of articles surveyed by Thomson Reuters in the Science Citation Index (Web of Science). By 2013, China produced some 13 percent of the literature, second only to the United States at 29 percent.

A short abstract

IOP FTC

IOP PUBLISHING — JOURNAL OF PHYSICS A: MATHEMATICAL AND THEORETICAL

J. Phys. A: Math. Theor. **44** (2011) 492001 (5pp) — doi:10.1088/1751-8113/44/49/492001

FAST TRACK COMMUNICATION

Can apparent superluminal neutrino speeds be explained as a quantum weak measurement?

M V Berry[1], **N Brunner**[1], **S Popescu**[1] **and P Shukla**[2]

[1] H H Wills Physics Laboratory, Tyndall Avenue, Bristol BS8 1TL, UK
[2] Department of Physics, Indian Institute of Technology, Kharagpur, India

Received 12 October 2011, in final form 27 October 2011
Published 11 November 2011
Online at stacks.iop.org/JPhysA/44/492001

Abstract
Probably not.

PACS numbers: 03.65.Ta, 03.65.Xp, 14.60.Pq

Citations are rare

- A study of over 1.3 million papers across all disciplines and years, found that 61% of papers had zero citations, and 12% had only one citation, with 4% having 16 or more citations
- Another study found that 55% of papers are not cited at all
- Citations of journal papers are most common in biological and physical sciences, less so in engineering where conferences are more important and where implementation is more important than publication. Social sciences tend to place greater reliance on books than on journals.

Open access

- The Universal Declaration of Human Rights that says "everyone has the right freely to … share in scientific advancement and its benefits"

Open access

- There is an important distinction to be drawn between work that scholars want to give away (journal articles, conference papers) and work from which they expect to derive income (books), and this present discussion focuses only on the former

- In academic publishing, authors are not paid by the journal for their writing; peer reviewers give their time voluntarily to evaluate papers and suggest improvements; and, journal editorial boards (and, sometimes, even journal editors) are unpaid. However, scientific papers can be very costly for readers to obtain (typically tens of dollars per paper).

- Low-cost publishing on websites has changed the cost structure of information publication and dissemination, making it possible to extend the reach of information to a much greater worldwide audience. Online publishing also allows articles to become available sooner than printed versions. There is considerable support for the idea of a more open research literature.

Open access

- Open access publishing aims to provide universal, unrestricted free access to full-text scholarly materials via the Internet
- This is different from the dissemination of research articles that has traditionally been controlled by publishers that regulate access by means of subscriptions and licences fees levied on users and academic libraries
- Open access publishing is based on the principle of maximizing the ease with which publications can be obtained. Open publications are visible to search engines, such as Google, thereby making them easily found. This serves the interests of authors, in that they want their writing to have maximum reach and impact, and want their work to be discussed and cited often. It also serves the interests of readers who desire low-cost or free access to a wide range of information. Peer reviewers are usually also authors and readers, so their interests are looked after too. However, it is also very important to safeguard the interests of publishers or scientific societies, as these bodies have high reputations and ensure the trustworthiness of what is published.

Open access

- Sources of funding for scientific societies include: membership fees, subscriptions to journals, sale of publications, sponsorship, limited advertising, conference fees, article-processing charges, and subsidies from funding agencies

- Through proper management of open-access publication, it is possible to maintain the same standards of high-quality production of peer-reviewed papers, with the potential for greatly increased dissemination and citation

Open access

- The high cost of obtaining published papers can be seen as a barrier for scientists and engineers working in the poorer countries of the developing world, as well as for unemployed scientists who would like to keep abreast of their fields while seeking work. Open access provides unrestricted availability to publications to anyone with a connection to the Internet.

- Another benefit of having open access to scientific papers is that search engines can be used to find associations between items that would be missed by human readers of individual papers. ‘Re-inventing the wheel’ can sometimes be avoided by ensuring that older papers are readily available to researchers.

Open access

- Today's world faces many policy choices, on issues such as climate change, energy generation, and recycling. These issues cannot be properly addressed without widespread access (even by the general public) to the results of scientific research in each of these areas.
- The British government has recently announced an initiative to make all taxpayer-funded research available online to anyone who wants to read or use it
- From 2014, the results of all research funded by the European Union must be open access
- In February 2013, the US White House announced that government-funded research should be made free to read within 12 months of publication
- Estimates of the proportion of papers available free online range from 30% to 50%
- It is hard to argue that keeping scientific information behind expensive pay walls is in the best interest of society. There is an increasing degree of dissatisfaction in the scientific community with the degree of control exerted by the big three publishing houses (namely Elsevier, Springer, and Wiley) who own most of the world's more than 20 000 academic journals, and account for about 42% of all journal articles published.

Open access

- Commercial publishers are concerned about losing their revenue stream to open-access competition. There is some scepticism about the economic viability of the open-access model of publishing, but these objections are often easily overcome.

- It has also been argued that there is a moral imperative to remove restrictions of access to the scientific literature.

- There is a fundamental question that needs to be addressed by learned societies – are they there to make a lot of money from publishing to underwrite their other activities, or are they there to advance science?

Open access

- It is in the best interests of authors and readers, and society at large, to ensure that scientific and technical papers are searchable and freely available to as wide an audience as possible
- It is argued that this can be done effectively by scientific and technical societies who are willing to publish their papers online using one or another variant of the open-access model

SAIMM

- SAIMM Journal was first published in 1894, and is currently published monthly
- Electronic copies of papers in PDF format are produced
- Old papers (initially from 1969) were scanned and subjected to optical character recognition (OCR) to make them searchable
- All of the previously unscanned historical copies of the journal (for the period 1894 to 1969) were digitised with the support of the Carnegie Foundation (as part of their effort to get a wide range of African journals online) and Sabinet
- Journal papers were initially made available on the SAIMM website at a cost, but were free to members
- In 2007, SAIMM's journal and conference papers were made free of charge to everyone
- A main strategic objective of SAIMM is to disseminate scientific and technical information to the benefit of the mining and metallurgical industries
- It is intended to provide the widest possible readership for authors who publish papers in the SAIMM Journal

SAIMM

- SAIMM also makes its papers available to other online databases, in order to reach as wide an audience as possible. This includes the international OneMine database and the SciELO system that is managed by the Academy of Science of South Africa.
- It is in the best interests of authors to ensure that their papers are freely available to as wide an audience as possible. It is also in the best interests of readers to have free access to papers. For these reasons, SAIMM neither charges authors nor readers for papers published in their journal or conference proceedings.

SAIMM

- SAIMM also publishes proceedings of many mining and metallurgical conferences that it organizes. The cost of producing conference proceedings is covered by conference fees. Conference papers are made available online via open access, a short period after the conference has taken place.
- SAIMM is a not-for-profit society where much of the work is done by volunteers, supported by a small secretariat with paid full-time staff. Authors and peer reviewers give their time free of charge (or are paid by their employers to do the work), but other aspects of the publishing chain (such as editing, type-setting, proof-reading, printing, packaging, posting, and website management) clearly come at a cost. Advertising and sponsorship cover many of these costs, but the principal sources of income of the institute are membership subscriptions and conference fees.

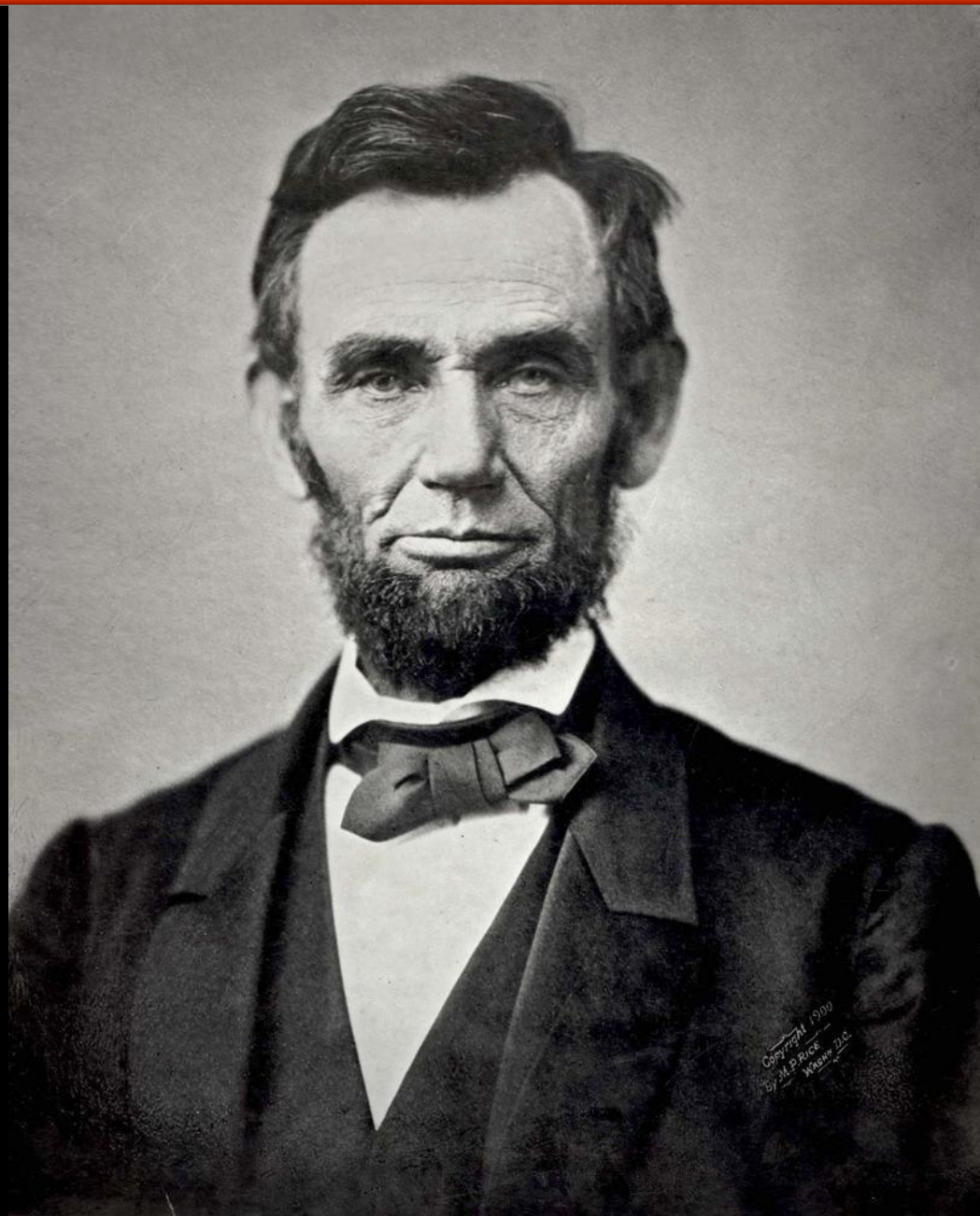
CAPACI OCCASIO
SAIMM

Mintek

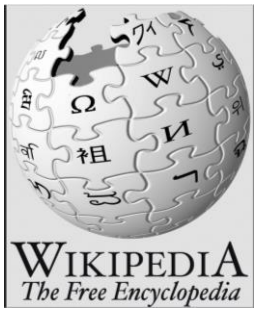
- The Pyrometallurgy Division at Mintek started in about 1996 to make available all of its open publications (papers and patents) on its website
- In order to make the initial selection of older papers available, the question of existing copyright restrictions had to be addressed
- In some cases, this was easily managed, by getting the permission of societies such as SAIMM to make copies of the papers available on the Mintek website
- Many scholarly publishers, including Elsevier, Springer, and John Wiley & Sons, now include concessions in their copyright policies to allow some form of self-archiving of full-text articles by authors in their institutional repositories or on personal websites

Don’t believe
everything
you read
on the Internet

Abraham Lincoln



Wikipedia and Encyclopedia Britannica



- In 2005, *Nature* published a study it undertook of the accuracy of articles from *Wikipedia* and *Encyclopedia Britannica*. This showed the difference in accuracy was not particularly great (averages out to 2.92 mistakes per article for Britannica, and 3.86 for Wikipedia).
- Britannica is based on strong scholarship, sound judgment, and disciplined editorial review
- Wikipedia is very current, comprehensive in coverage, and is based on continuous correction by many people. The core Wikipedia values include: neutral point of view, no original research, verifiable information only, and citing sources.

Stigler's Law of Eponymy

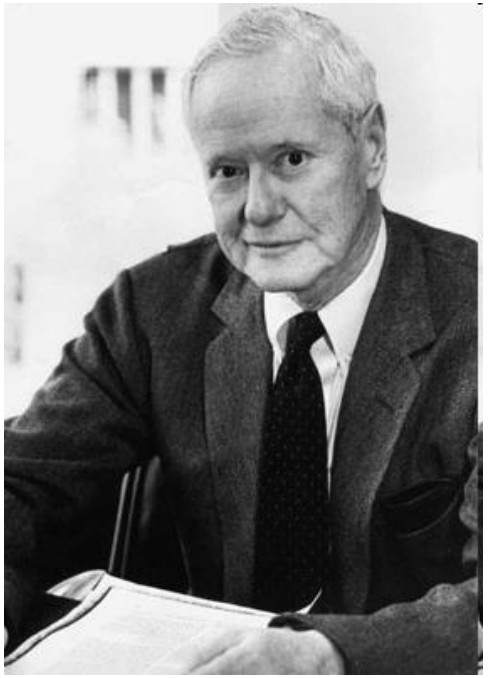
- "No scientific discovery is named after its original discoverer"
- Named after University of Chicago statistics professor Stephen Stigler for his 1980 publication
- Historical acclaim for discoveries is often assigned to persons of note who bring attention to an idea that is not yet widely known, whether or not that person was its original inventor

Stigler's Law

- Alzheimer's disease had been previously described by at least half a dozen others before Alois Alzheimer's 1906 report which is often (wrongly) regarded as the first description of the disorder
- The Bessemer process was discovered by William Kelly in 1851. Henry Bessemer was the first to obtain a patent in 1855
- Fibonacci was not the first to discover the famous sequence of Fibonacci numbers. They had existed in Indian mathematics since 200 BC. Fibonacci provided the series in 1202 AD.
- The normal or Gaussian distribution was introduced by Abraham de Moivre in 1733, but was named after Carl Friedrich Gauss who began using it in 1794
- Newton's first and second laws of mechanics were known and proposed in separate ways by Galileo, Hooke, and Huygens before Newton described these in his *Philosophiæ Naturalis Principia Mathematica*. Newton owns the discovery of only the third one.
- The Reynolds number in fluid mechanics was introduced by George Stokes, but is named after Osborne Reynolds, who popularised its use
- Stokes's theorem was discovered by Lord Kelvin

Stigler’s Law

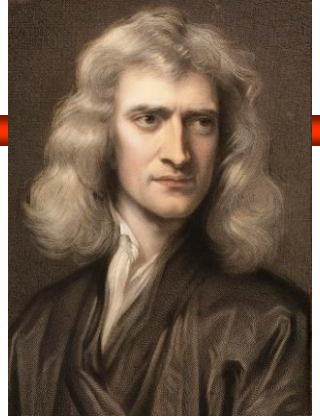
- Stigler named the sociologist Robert K. Merton as the discoverer of "Stigler's law", so making his ‘law’ an example of itself



Robert K. Merton

- Robert Merton described the principle in his 1957 Presidential Address to the American Sociological Society

Newton and Leibniz



- Robert Merton talks about the pressure of priority in scientific discovery.

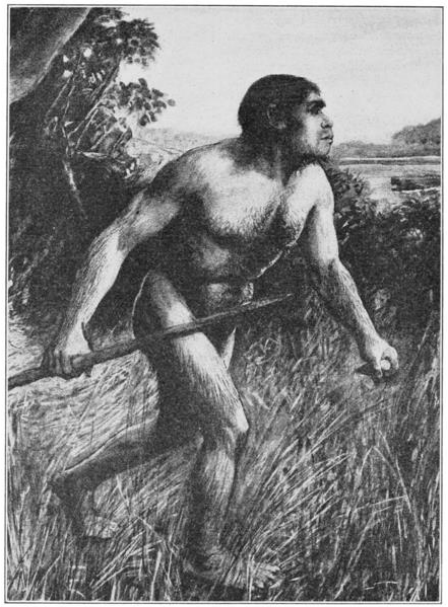
- There was a great deal of controversy between Newton and Leibniz over the invention of calculus. When the Royal Society finally established a committee to adjudicate the rival claims, Newton, who was then president of the Royal Society, packed the committee, helped direct its activities, anonymously wrote the preface for the second published report – the draft is in his handwriting – and included in that preface a disarming reference to the old legal maxim that “no one is a proper witness for himself”. Newton must have felt intense pressure for self-vindication that made him adopt such offensive means for the defence of his valid claims. It was not because Newton was so weak but because the institutionalized values were so strong that he was driven to such lengths.”

Fraud by Vrain-Lucas

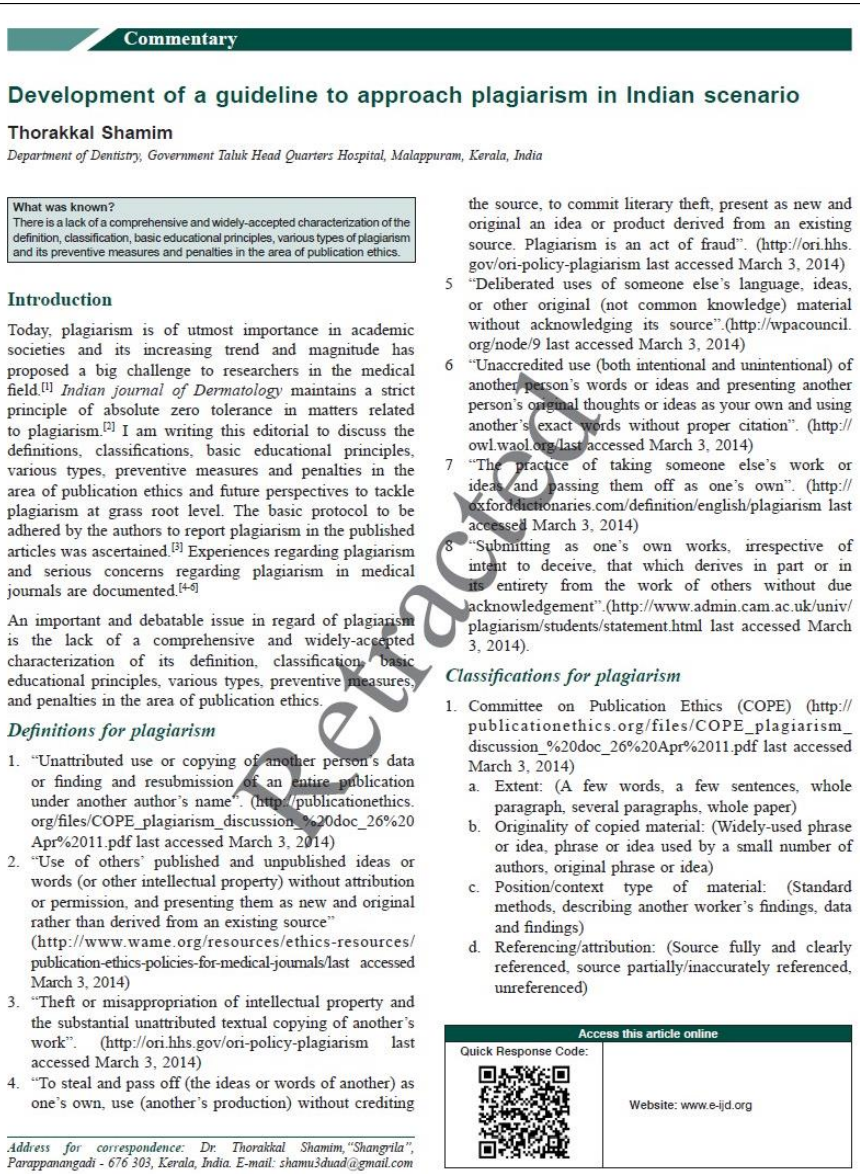
- Merton (1957) mentions the remarkably prolific Vrain-Lucas who, in the mid-1800s, created and sold more than 27 000 pieces of manuscript that included letters by Pontius Pilate, Mary Magdalene, the resurrected Lazarus, Ovid, Luther, Dante, Shakespeare, Galileo, Pascal, and Newton, all written on paper and in modern French. "Most provocative among these documents was the correspondence between Pascal and the then eleven-year-old Newton (all in French, of course, although even at the advanced age of thirty-one Newton could struggle through French only with the aid of a dictionary), for these letters made it plain that Pascal, not Newton, had, to the greater glory of France, first discovered the law of gravitation, a momentous correction of history, which for several years excited the interest of the *Académie des Sciences* and usurped many pages of the *Comptes Rendus* until, in 1869, Vrain-Lucas was finally brought to book and sentenced to two years in prison."

Piltdown hoax

- The Piltdown Man was an infamous paleoanthropological hoax in 1912 in which bone fragments (parts of a skull and jawbone) were presented as the fossilised remains of a previously unknown early human
- These fragments were said to have been collected in 1912 from a gravel pit at Piltdown, East Sussex, England by Charles Dawson
- The significance of the specimen remained controversial until it was exposed in 1953 as a forgery, consisting of the lower jawbone of an orangutan deliberately combined with the cranium of a fully developed modern human



Plagiarism



Commentary

Development of a guideline to approach plagiarism in Indian scenario

Thorakkal Shamim

Department of Dentistry, Government Taluk Head Quarters Hospital, Malappuram, Kerala, India

What was known?
There is a lack of a comprehensive and widely-accepted characterization of the definition, classification, basic educational principles, various types of plagiarism and its preventive measures and penalties in the area of publication ethics.

Introduction

Today, plagiarism is of utmost importance in academic societies and its increasing trend and magnitude has proposed a big challenge to researchers in the medical field.[1] *Indian journal of Dermatology* maintains a strict principle of absolute zero tolerance in matters related to plagiarism.[2] I am writing this editorial to discuss the definitions, classifications, basic educational principles, various types, preventive measures and penalties in the area of publication ethics and future perspectives to tackle plagiarism at grass root level. The basic protocol to be adhered by the authors to report plagiarism in the published articles was ascertained.[3] Experiences regarding plagiarism and serious concerns regarding plagiarism in medical journals are documented.[4-6]

An important and debatable issue in regard of plagiarism is the lack of a comprehensive and widely-accepted characterization of its definition, classification, basic educational principles, various types, preventive measures, and penalties in the area of publication ethics.

Definitions for plagiarism

1. "Unattributed use or copying of another person's data or finding and resubmission of an entire publication under another author's name". (http://publicationethics.org/files/COPE_plagiarism_discussion_%20doc_26%20Apr%2011.pdf last accessed March 3, 2014)
2. "Use of others' published and unpublished ideas or words (or other intellectual property) without attribution or permission, and presenting them as new and original rather than derived from an existing source" (http://www.wame.org/resources/ethics-resources/publication-ethics-policies-for-medical-journals/last accessed March 3, 2014)
3. "Theft or misappropriation of intellectual property and the substantial unattributed textual copying of another's work". (http://ori.hhs.gov/ori-policy-plagiarism last accessed March 3, 2014)
4. "To steal and pass off (the ideas or words of another) as one's own, use (another's production) without crediting the source, to commit literary theft, present as new and original an idea or product derived from an existing source. Plagiarism is an act of fraud". (http://ori.hhs.gov/ori-policy-plagiarism last accessed March 3, 2014)
5. "Deliberated uses of someone else's language, ideas, or other original (not common knowledge) material without acknowledging its source".(http://wpacouncil.org/node/9 last accessed March 3, 2014)
6. "Unaccredited use (both intentional and unintentional) of another person's words or ideas and presenting another person's original thoughts or ideas as your own and using another's exact words without proper citation". (http://owl.waol.org/last accessed March 3, 2014)
7. "The practice of taking someone else's work or ideas and passing them off as one's own". (http://oxforddictionaries.com/definition/english/plagiarism last accessed March 3, 2014)
8. "Submitting as one's own works, irrespective of intent to deceive, that which derives in part or in its entirety from the work of others without due acknowledgement".(http://www.admin.cam.ac.uk/univ/plagiarism/students/statement.html last accessed March 3, 2014).

Classifications for plagiarism

1. Committee on Publication Ethics (COPE) (http://publicationethics.org/files/COPE_plagiarism_discussion_%20doc_26%20Apr%2011.pdf last accessed March 3, 2014)
   a. Extent: (A few words, a few sentences, whole paragraph, several paragraphs, whole paper)
   b. Originality of copied material: (Widely-used phrase or idea, phrase or idea used by a small number of authors, original phrase or idea)
   c. Position/context type of material: (Standard methods, describing another worker's findings, data and findings)
   d. Referencing/attribution: (Source fully and clearly referenced, source partially/inaccurately referenced, unreferenced)

Address for correspondence: Dr. Thorakkal Shamim, "Shangrila", Parappanangadi - 676 303, Kerala, India. E-mail: shamu3duad@gmail.com

Access this article online
Quick Response Code:
Website: www.e-ijd.org

- An article from 2014 entitled "Development of a guideline to approach plagiarism in Indian scenario" was retracted by the Editors of the Indian Journal of Dermatology (2015), as large portions of the manuscript were copied from the first round questionnaire of a dissertation entitled 'Developing a comprehensive guideline for overcoming and preventing plagiarism at the international level based on expert opinion with the Delphi method' by another author

Peer review

- The system of peer review is used to maintain standards and to improve the quality of papers. This vital system is, however, significantly flawed. There is little incentive for reviewers to invest sufficient time in picking up all errors in publications, and any ineptitude on their part is usually protected by anonymity. It has reached the point where some reviewers have mistakenly permitted the publication of hoax papers deliberately presented with a complicated scientific facade. In light of such astounding inadequacies, perhaps a more open review process would be an improvement.

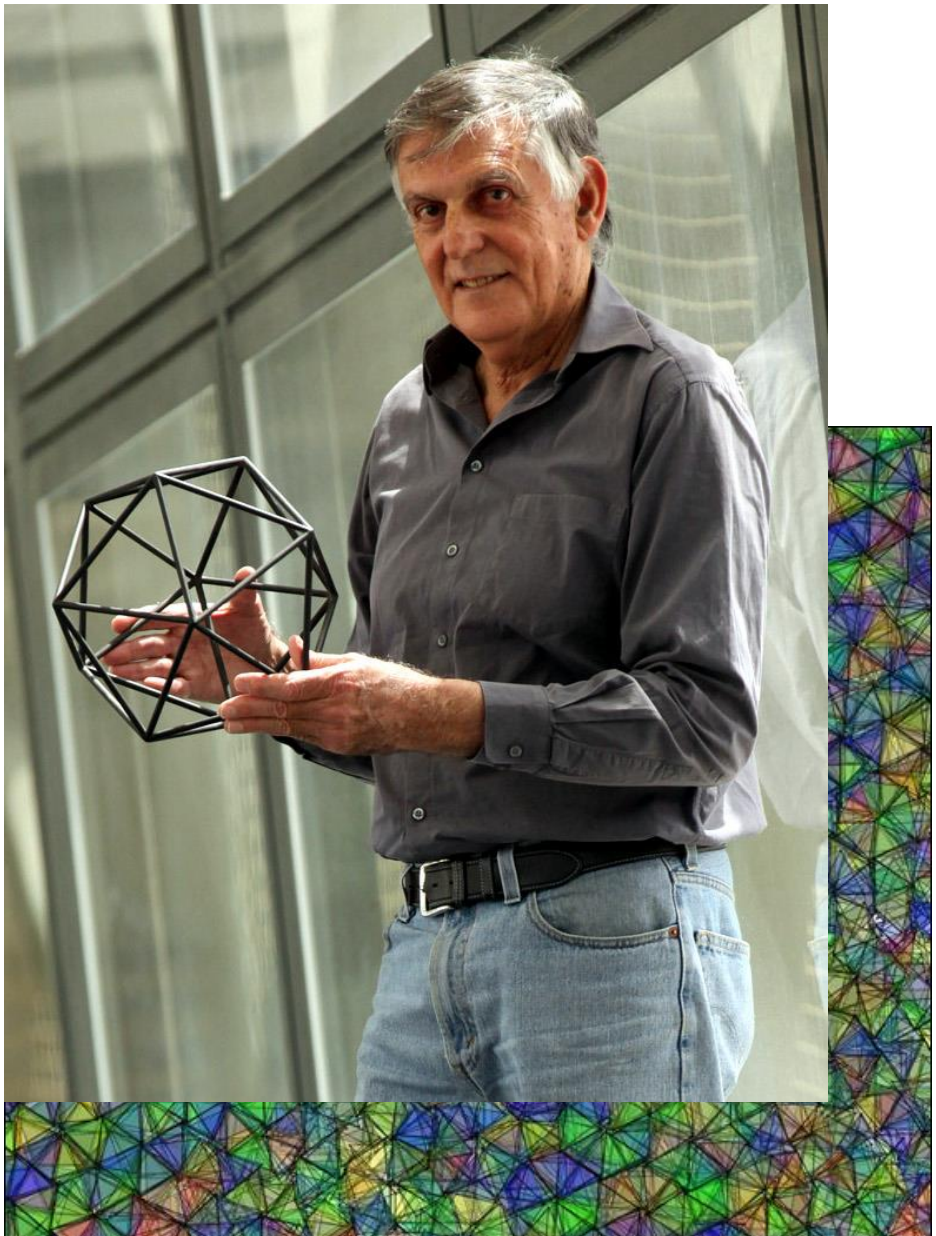
Peer review

- Peer reviewers are not paid, nor held accountable
- Costly and time-consuming
- Peer review is resistant to new or controversial ideas
- Is it true and is it new?
- Agreement between referees is often little higher than by chance
- Vulnerable to misconduct, plagiarism and breach of confidentiality. Some of the most cited papers of all times were rejected by referees

Rejected / delayed recognition – Bad peer review

- Discovery of blood typing
- Jenner's 1796 paper describing vaccinations against smallpox
- Krebs's paper describing the citric acid cycles
- Murray Gell-Mann's work on quarks
- At least 35 articles that would eventually earn the Nobel Prize and fame for their authors were rejected outright during the initial inspection by reviewers

Harsh reviews - Quasicrystals



- Paper rejected by Journal of Applied Physics in 1984
- Later published in Metallurgical Transactions A and Physical Review Letters
- Linus Pauling said: “There is no such thing as quasicrystals, only quasi-scientists”
- Dan Schechtman won the Nobel Prize for Chemistry in 2011

More damaging reviews

- The nineteenth-century physicist Waterston, his classic paper on molecular velocity having been rejected by the Royal Society as “nothing but nonsense”, became hopelessly discouraged and left science altogether

- Deeply disappointed by the lack of response to his historic papers on heredity, Mendel refused to publish the now-permanently lost results of his further research and, after becoming abbot of his monastery, gave up his research on heredity

- Robert Mayer, tormented by refusals to grant him priority for the principle of conservation of energy, tried a suicide leap from a third-story window and succeeded only in breaking his legs and being straitjacketed, for a time, in an asylum

Bowman's equation for arc shape

- Very important equation for modelling plasma arcs appears twice in the literature – wrongly!

$$\frac{r_a}{r_k} = 1 + 2.2\left[1 - \exp\left(-\frac{z}{5r_k}\right)\right]$$

the separation of the graphite tip from the metal bath. As clearly demonstrated by the measurements of Ao et al. and Miyashita et al. (3,5), jet penetration is considerable, especially at short gaps. Thus the interpretation of voltage/gap measurements in terms of voltage/arc length characteristics requires considerable care.

Arc voltage fluctuations

An experimental observation which is both relevant and requires explanation is the large fluctuation of voltage at any fixed gap and current. The envelopes of these fluctuations are shown in figure 1. At very short gaps the voltage range can be even larger than shown in the figure. Miyashita et al. have recorded swings of +/- 200 V and more around the mean. A frequency analysis by Ahlers and Timm did not show any particular frequency in these swings; the spectrum demonstrated a general decrease of amplitude with frequency.

ARC SHAPES

Although several of the authors quoted above recorded high-speed photographs of arcs, itself a major achievement, they have not reported any detailed analysis of arc dimensions. Accordingly the older data of Montgomery and Sharp (7), Bowman et al. (8) and Jordan et al. (9)havebeen reanalysed.

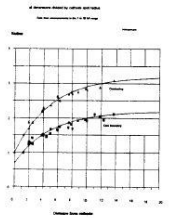


Figure 2 Arc shapes

Figure 2 shows arc shapes in dimensionless co-ordinates; distance from the cathode is divided by cathode spot radius, while arc radii are also referred to the cathode spot radius. Shown are both core radius, whose boundary is close to the 10,000 degK isotherm, and the visible radius, which is probably very close to the electrically conducting boundary. The core shows itself in photographs as distinctly brighter than the surrounding plasma. This is due mainly to the strong continuum radiation from the free electrons which dominate radiation above 10,000 degK. Data for core shapes comes from both AC and DC arcs, while the outer boundary data comes from DC arcs only.

When photographs are taken with liquid metal present the black-body background radiation, some 70 W/cm2 at 1600 degC, tends to overwhelm the radiation from the arc outside of the core. For this reason it is difficult to locate the outer boundary of the arc from such photographs. This was the case for the AC data included in figure 2. The DC arc photos were, however, taken with a room temperature background so that outer arc boundaries were much clearer.

Although Strachan (10) has made accurate measurements on arcs up to 20 kA they cannot be included in the data of figure 2. His arcs lasted only a few msec and therefore ran with a much smaller cathode spot than the long-duration industrial arcs. As a result his arc shapes show a greater range of radii than those in figure 2.

The fact that arc shapes seem to be the same over the range 1 to 10 kA encourages an extrapolation to higher currents. For this it is only necessary to decide on a value for the cathode spot radius. This is readily available if the cathode spot current density is known. In fact the sources for figure 2 suggest that a constant density in the range 3.0 to 3.5 kA/cm2 would seem appropriate for higher current DC arcs.

For quantitative analysis the equation for the outer boundary in figure 2 is :-

$$\frac{r_a}{r_k} = 1 + 2.2\exp\left(-\left(\frac{z}{5r_k}\right)\right) \quad (1)$$

where: z is the axial distance from the cathode spot.
r_a is the radius of the arc conducting boundary
r_k is the radius of the cathode spot.

When all radii are divided by 1.45 the same equation also applies to the core dimensions.

It should be noted that this data applies to an open, stable arc. The effect of slag and instabilities will be discussed later.

$$\frac{r_c}{r_k} = 1 + 2.2\left(1 - e^{\left(\frac{z}{5\,r_k}\right)^2}\right)$$

2009 book

$$\frac{r_a}{r_k} = 1 + 2.2\exp\left(-\left(\frac{z}{5r_k}\right)\right)$$

1994 paper

period. The voltage gradients are in same range as recorded above; in addition **figure 28** shows that the gradient increases with decreasing arc voltage, respectively length.

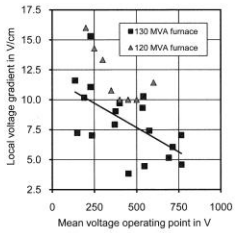


Fig. 28: Voltage gradient during liquid bath period [Den02], [Tre08]

3.2.2 Arc shape

A useful empirical finding, derived from photographs of arcs in the 1 to 10 kA range (figure 1), was that DC arcs from graphite electrodes tend to have a similar shape, expressed in dimensionless form by [Bow94]

$$\frac{r_c}{r_k} = 1 + 2.2\left(1 - e^{\left(\frac{z}{5\,r_k}\right)^2}\right) \quad (3.3)$$

In this equation r_c and r_k are the arc conducting radius and cathode spot radius respectively, and z is the axial distance along the column from the cathode. As the axial distance increases beyond ~10 r_k the conducting radius approaches an asymptotic limit of 3.2 r_k. Assuming that the same relation holds at higher currents the equation allows the arc radius to be calculated from the dimensions of the cathode spot.

3.2.3 Arc voltage/length relationship

An arc voltage/length relation can be calculated from the shape, using equation (3.4), fixing the dimension of the cathode spot from an assumed constant current density. Empirical values of voltage gradient, as indicated in figures 27 and 28, are assumed to apply to the fully expanded arc conducting radius, that is, for an axial distance of more than 20 times the cathode spot radius. The voltage gradient along the column is assumed to be inversely proportional to the cross-sectional area of the conducting channel. **Figure 29** shows the results of such calculations for four different asymptotic voltage gradients for which the cathode spot current density was taken as 3.5 kA/cm^2.

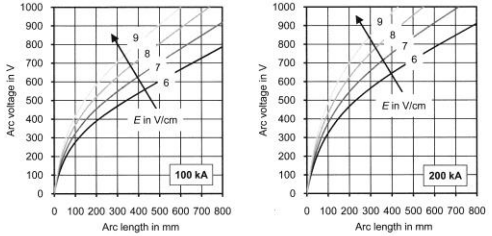


Fig. 29: Arc voltage/length relations for DC arcs for different asymptotic voltage gradients

CoO activity coefficient

$$\ln \gamma_{CoO} = 8.3283 + \frac{14.9763 \times 10^3}{T} + 0.9552 \times B \qquad (13)$$

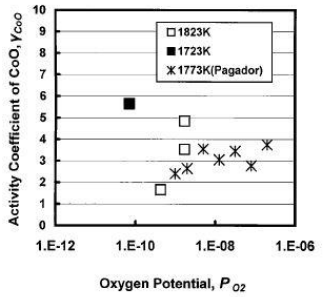


Fig. 5 Calculated activity coefficient of CoO in slag in comparison with the reported data.

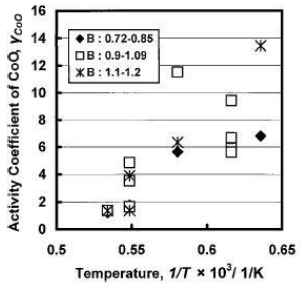


Fig. 6 Influence of temperature and slag composition on the activity coefficient of CoO in slag.

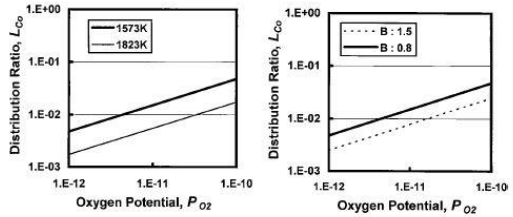


Fig. 7 Estimation of cobalt distribution ratio in steelmaking process.

Low-basicity slag has a smaller activity coefficient of CoO as compared to that of high-basicity slag. On the basis of this result, eq. (13) is obtained for the activity coefficient of CoO in slag.

$$\ln \gamma_{CoO} = 8.3283 + \frac{14.9763 \times 10^3}{T} + 0.9552 \times B \qquad (13)$$

4. Discussion

The behavior of cobalt in the steelmaking process can be estimated from these results. The activity coefficient of CoO in slag phase was estimated from eq. (13), and the Fe-C-Co alloy was considered as the metal phase. In this case, the effect of carbon on the activity coefficient of cobalt was not taken into account. The cobalt content and the slag volume before the oxidation treatment were assumed to be 50 ppm and 300 kg/t, respectively. Figure 7 shows the calculation results. The distribution ratio increases with the increase in temperature as the strong temperature dependence of equilibrium constant, although the activity coefficient of CoO increases. Therefore, the treatment with acid slag under low temperature is considered to be favorable for increasing the distribution ratio. Nevertheless, as the maximum oxygen potential in the steelmaking process is about 10^{-10}, the distribution ratio would be less than 0.1, and the decrease in the cobalt content would be negligibly small.

5. Conclusions

For the production of low-activation steel by the steelmaking process, the cobalt distribution between MgO-saturated FeO_x-MgO-CaO-SiO_2 slag and Fe-Cu-Co molten alloy has been investigated. The results are summarized as follows: (1) the distribution ratio increases with decreasing temperature and (2) the activity coefficient of CoO decreases with increasing temperature and decreasing slag basicity. Even though low temperature and low basicity are suitable conditions for refining, it would be very difficult to reduce the cobalt content by the oxidation reaction, owing to the small distribution ratio of cobalt.

- Calculated activity coefficient of CoO(l) in slag as a function of oxygen partial pressure
- At 1400°C (1673 K), graph shows γCoO=10, but equation has the wrong form and shows 96 000 000.
- Author said: “You are right. This equation is not correct. I made a serious mistake. Please never use this equation.”

Errata

- Electronic publishing allows errata to be linked to the original papers. This might improve the current situation where errors tend to be propagated from one paper to the next.

Sokal hoax

- Alan Sokal (Professor of Physics at New York University) published a hoax article in *Social Text*, a journal of postmodern cultural studies, in 1996 entitled “Transgressing the Boundaries: Towards a Transformative Hermeneutics of Quantum Gravity”
- He wanted to see if the journal would publish an article that liberally salted with nonsense, but sounded good, and flattered the editors’ ideological preconceptions
- The paper proposed that “quantum gravity is a social and linguistic construct”
- The hoax sparked a debate about the scholarly merit of humanistic commentary about the physical sciences
- Sokal maintains that it's almost impossible to function in the world without some functional sense of truth

The 'two worlds' of CP Snow

- C.P. Snow (1964) once suggested the second law of thermodynamics as a test of scientific literacy for the humanist, and said it was ‘about the scientific equivalent of: *Have you read a work of Shakespeare’s?*’.

Bohannon's test of peer review

- Between January and August 2013, John Bohannon (a journalist for *Science*) submitted almost identical fake papers (obviously and seriously flawed) to 304 supposedly peer-reviewed journals
- The paper's template is "Molecule X from lichen species Y inhibits the growth of cancer cell Z". Authors and affiliations were unique and fake.
- 157 journals accepted, and 98 rejected the paper
- Some accepting journals were published by major publishers (e.g., Elsevier), but predatory paid open access journals fared the worst
- An article entitled "Who's afraid of peer review?" was published in Science on 4 October 2013
- This study was seriously flawed, but made quite an impact

Measurements of scientific output

- Citation indexing was pioneered by Eugene Garfield in 1955
- Number of publications
  - ‘least publishable units’
- Journal impact factor
- Number of citations
- Number of citations per paper
- h-index

Journal Impact Factor

- Eugene Garfield first introduced the concept of an impact factor in 1955, when he was director of the Institute for Scientific Information (ISI) - now part of Thomson Reuters

- About 70% of refereed scientific journal papers are indexed, even though less than 39% of refereed journals are indexed (more than 15 000 refereed journals are not indexed)

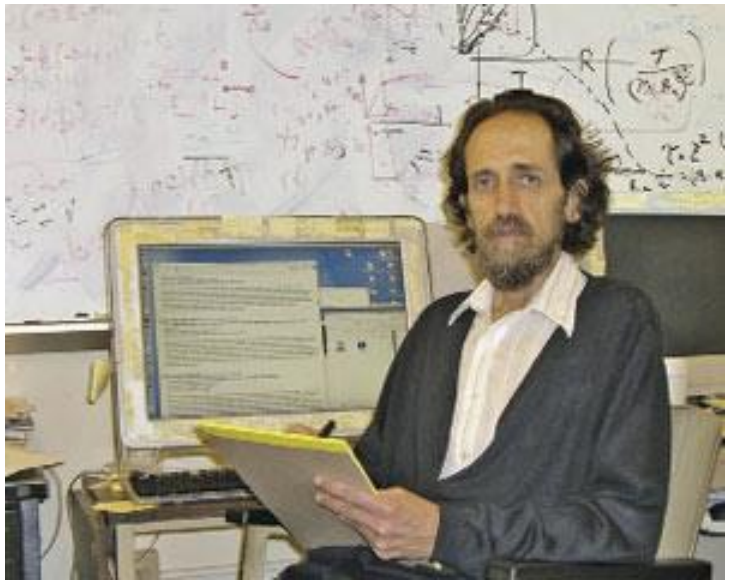
- The ratio of (the number of citations in the current year to the items published in the previous two years) to (the total number of articles published in the same two years)

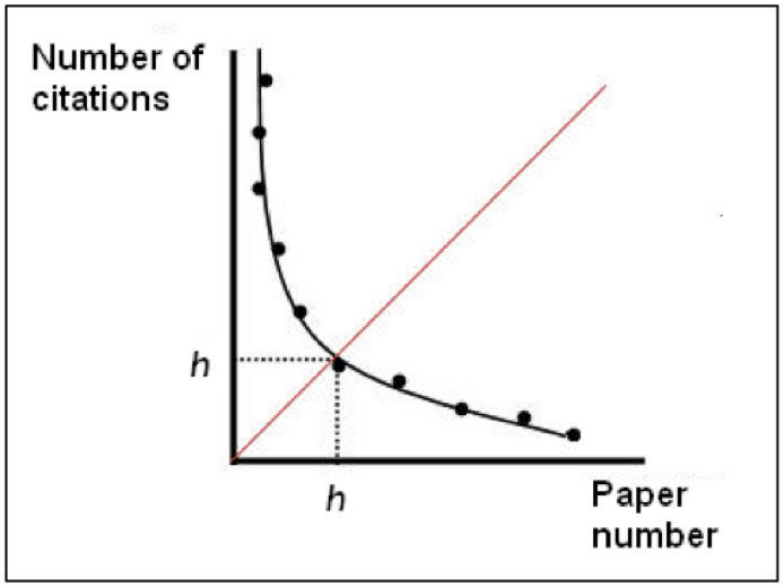
- JIF (2014) = (number of citations in 2014 to papers articles published in 2012 and 2013) / (total number of articles published in 2012 and 2013)

Journal Impact Factor

- Cannot be used to meaningfully compare journals in different fields
- Widely criticised for use as a lazy proxy for quality of a particular paper
- The intrinsic merit of a paper is much more important than where it is published
- Publications in high ranking journals are not only more likely to be fraudulent than articles in lower ranking journals, but also more likely to present discoveries which are less reliable (i.e., are inflated, or cannot subsequently be replicated). Some of the sociological mechanisms behind these correlations have been documented, such as pressure to publish (preferably positive results in high-ranking journals), leading to the potential for decreased ethical standards.

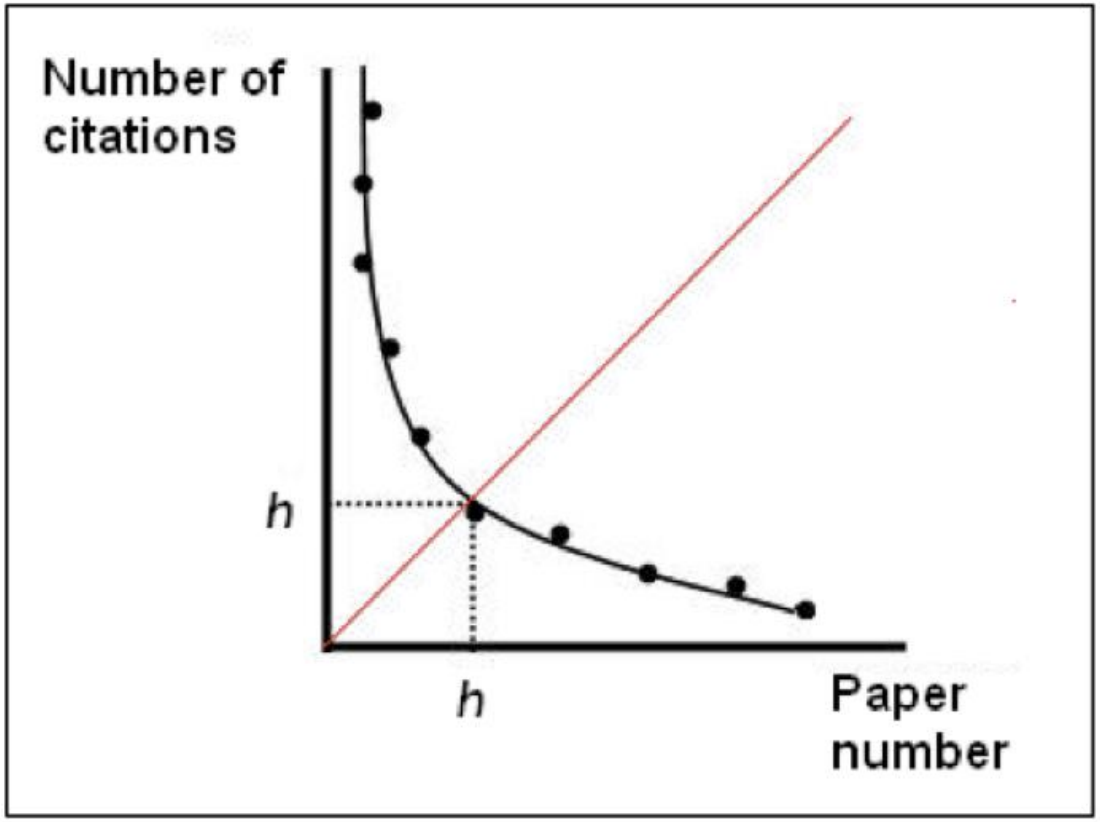
h-index





- Proposed in 2005 by Jorge E. Hirsch, a physics professor at the University of California, San Diego
- Measure of impact and quantity of scientific output as a single number
- h stands for Hirsch or highly cited
- He defines the h index as the maximum number of an author's papers that have been cited at least h times
- Having an *h* index of 15 means that a scientist has published 15 papers that each had at least 15 citations
- Avoids problems where the total number of papers does not indicate the quality of scientific publications, whereas citation counts can be disproportionately affected by a single publication of major influence
- Useful comparison between authors within a specific subject area

h-index



- Curve for number of citations versus paper number, with papers numbered in order of decreasing citations (Hirsch, 2005). The intersection of the 45° line with the curve indicates h

Google Scholar *h*-index

Albert Einstein

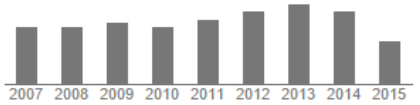
Follow

Institute of Advanced Studies, Princeton
Physics
No verified email

Title 1–20	Cited by	Year
Can quantum-mechanical description of physical reality be considered complete? A Einstein, B Podolsky, N Rosen Physical review 47 (10), 777	13652	1935
Uber einen die Erzeugung und Verwandlung des Lichtes betreffenden heurischen Gesichtpunkt A Einstein Ann. Phys. 17, 132-148	7769 *	1905
On the movement of small particles suspended in stationary liquids required by the molecular-kinetic theory of heat A Einstein Annalen der Physik 17, 549-560	~~5891~~ *	1905
Zur Elektrodynamik bewegter Körper A Einstein	4161 *	
Investigations on the Theory of the Brownian Movement A Einstein Dover publications	3717	1956
Eine neue bestimmung der moleküldimensionen A Einstein Annalen der Physik 324 (2), 289-306	3148	1906
The meaning of relativity A Einstein, EP Adams, EG Straus, S Bargmann Princeton University Press	2948 *	1950
The meaning of relativity A Einstein Routledge	~~2784~~ *	2003

Google Scholar

Citation indices	All	Since 2010
Citations	93075	30023
h-index	105	64
i10-index	362	213

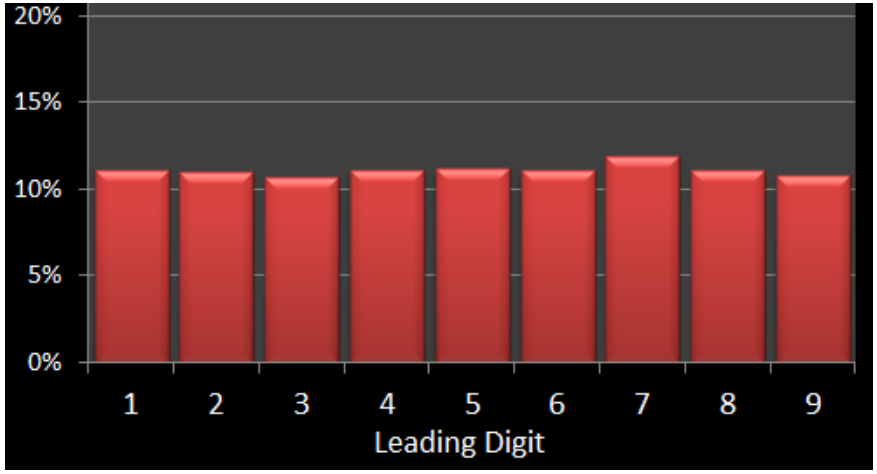


Google and page rank

- Google ranks the importance of a document in a search according to the networked importance of the pages that link to it
- Algorithm developed by Larry Page and Sergey Brin

Benford's Law

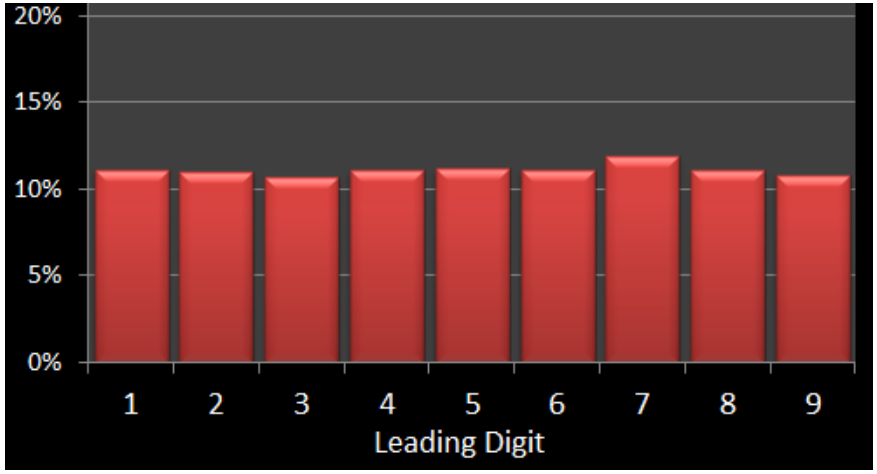
- Imagine having a list of randomly occurring measurements, for example, the altitude (in metres) of the 122 000 most highly populated towns in the world
- What would the frequency distribution of the leftmost digit (1-9) look like? Would it be evenly distributed at about 11% each?



Benford's Law

- Imagine having a list of randomly occurring measurements, for example, the altitude (in metres) of the 122 000 most highly populated towns in the world
- What would the frequency distribution of the leftmost digit (1-9) look like? Would it be evenly distributed at about 11% each?

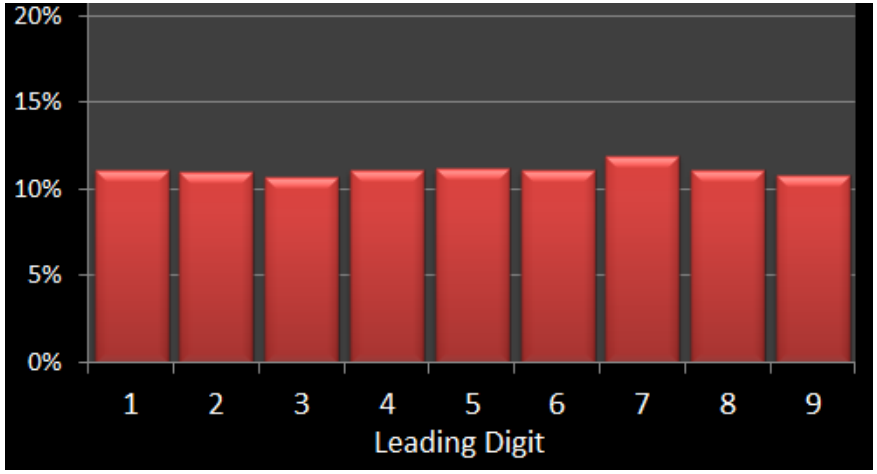
Wrong!

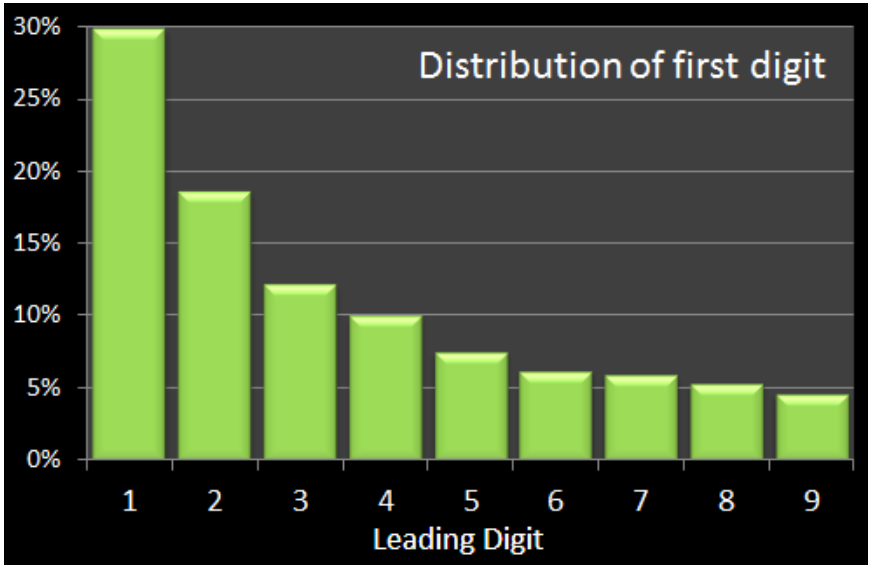


Benford's Law

- Imagine having a list of randomly occurring measurements, for example, the altitude (in metres) of the 122 000 most highly populated towns in the world
- What would the frequency distribution of the leftmost digit (1-9) look like? Would it be evenly distributed at about 11% each?

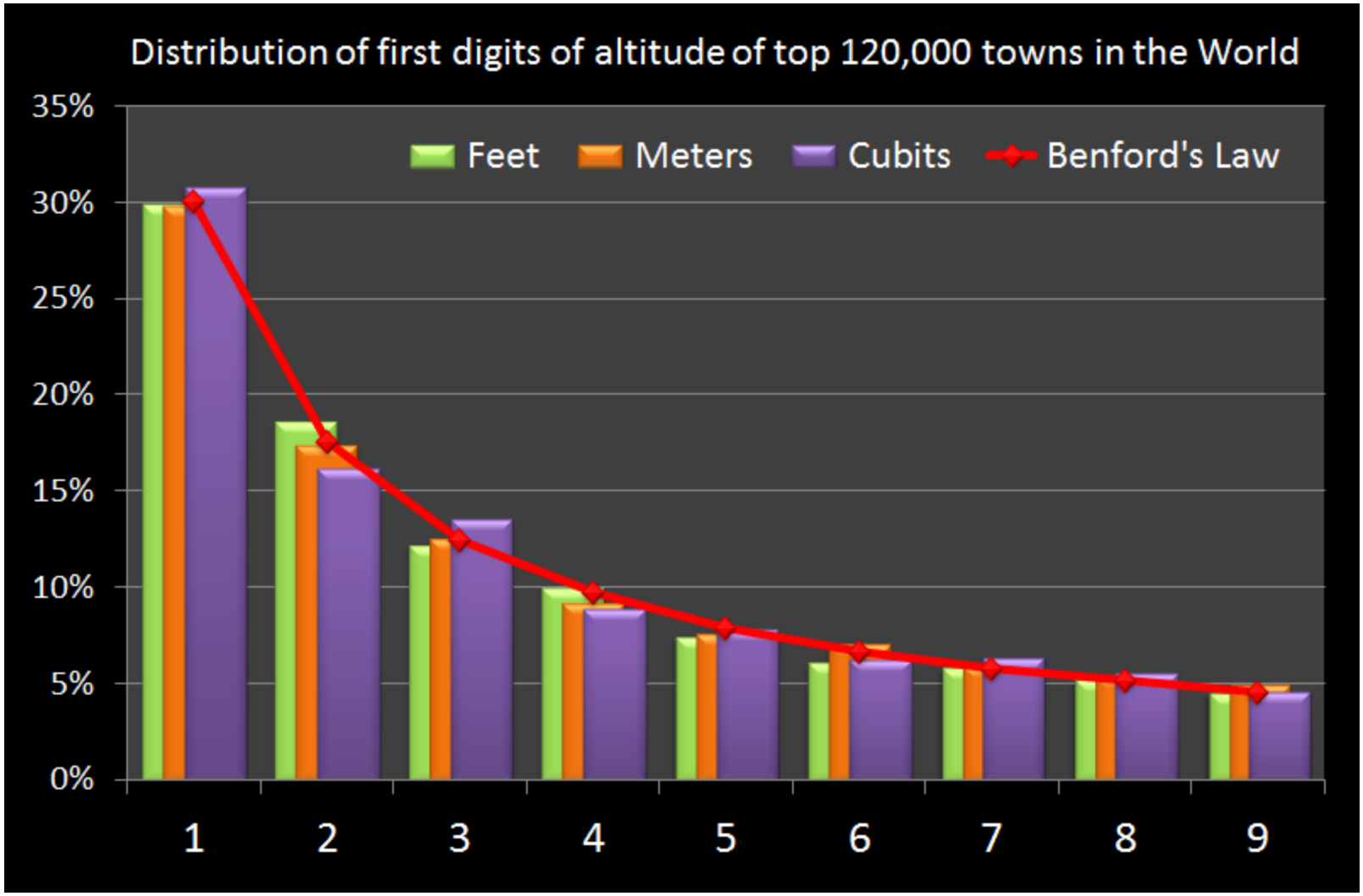
Wrong!





Benford's Law

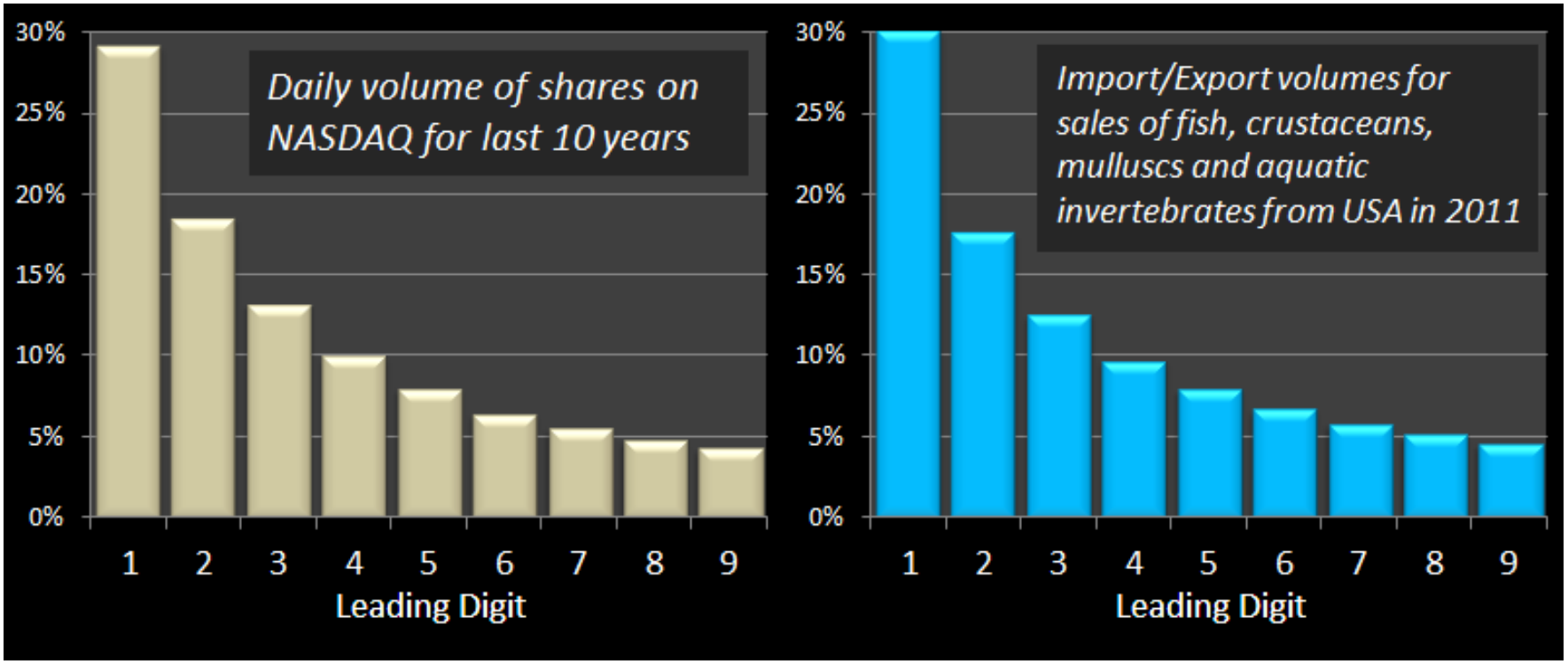
- The distribution remains the same even if the units are changed from metres, to feet, or to cubits (523 mm)



http://datagenetics.com/blog/march52012/index.html

Benford's Law

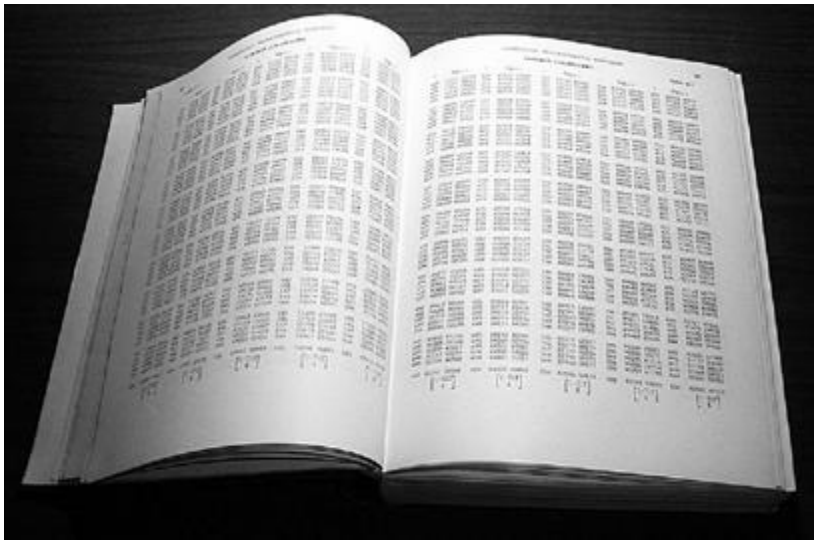
- The same pattern occurs with many other data sources, such as stock market volume, distances to stars, etc.



Benford’s Law

- Simon Newcomb, a Canadian mathematician, noticed that, when using his book of log tables, the earlier pages (which contained numbers that start with 1) were much more worn than the other pages. He documented this discovery in 1881.

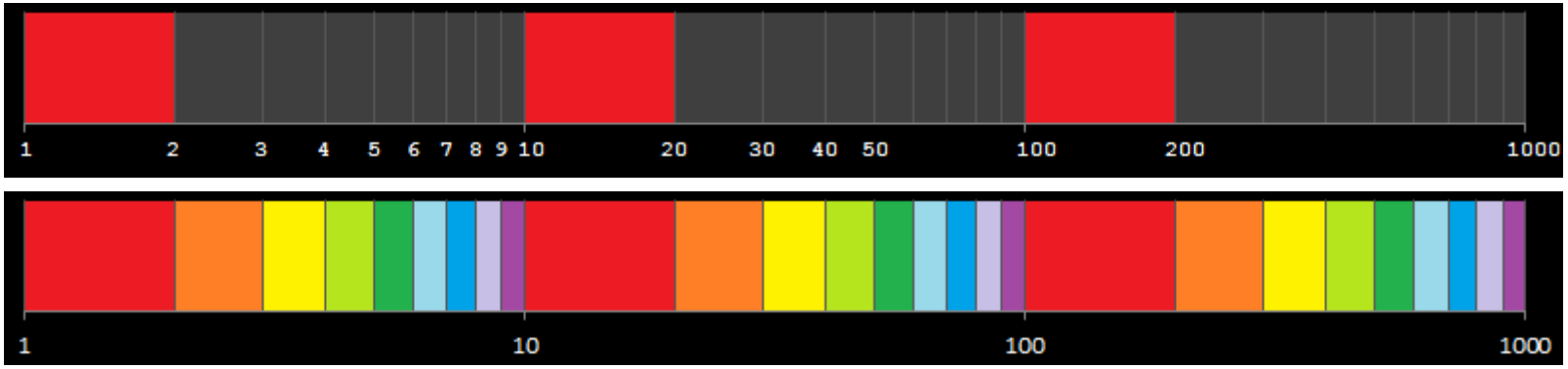
Frank Benford



- The physicist Frank Benford rediscovered this observation in 1938 and, as a great example of Stigler’s Law, it is named after him as Benford's Law.
- Benford’s Law is very useful in the detection of fraudulent data.

Benford's Law

- Imagine a plant growing. To go from length 1.0 to 2.0 it would have to increase by 100%, but to go from 2.0 to 3.0 it would need to grow only 50%, and so on.



- Benford's formula states that the probability of the leading digit being of a certain value can be described by the following function:

d	*Pr (d)*	Relative size of *Pr (d)*
1	30.1%	
2	17.6%	
3	12.5%	
4	9.7%	
5	7.9%	
6	6.7%	
7	5.8%	
8	5.1%	
9	4.6%	

$$\Pr(d) = \log_{10}(d+1) - \log_{10}(d)$$

This simplifies to:

$$\Pr(d) = \log_{10}\left(1 + (1/d)\right)$$

Conclusions

- Scientific publishing remains alive and well, despite some problems and challenges
- Electronic technology provides some wonderful opportunities to improve the way we communicate scientific results
- The approaches outlined here should help us to get closer to the truth