Being Ethical in Large-Team Science:
A Quantitative Historical Perspective

History of Science Society 2013 Annual Meeting
Boston, Massachusetts
21-24 November 2013

Alexander M. Petersen
IMT Institute for Advanced Studies, Lucca Italy

Institutions
Markets
Technologies
IMT
INSTITUTE
FOR ADVANCED
STUDIES
LUCCA

Outline

Growth of Science

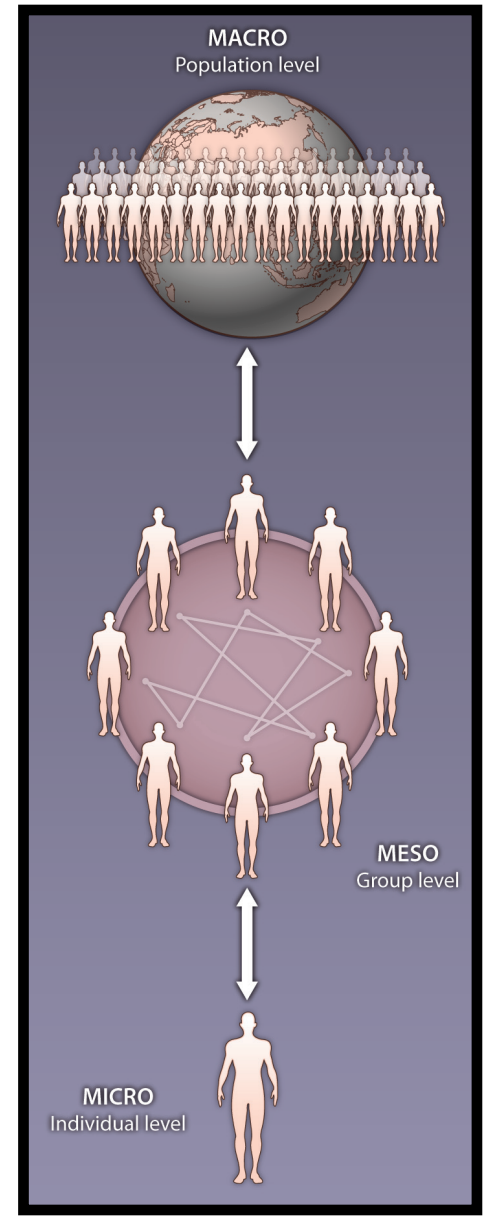
- Institutional perspective
- Computational historical perspective on

the increasing prevalence of teams in science

Discourse on the ethics of Team Science

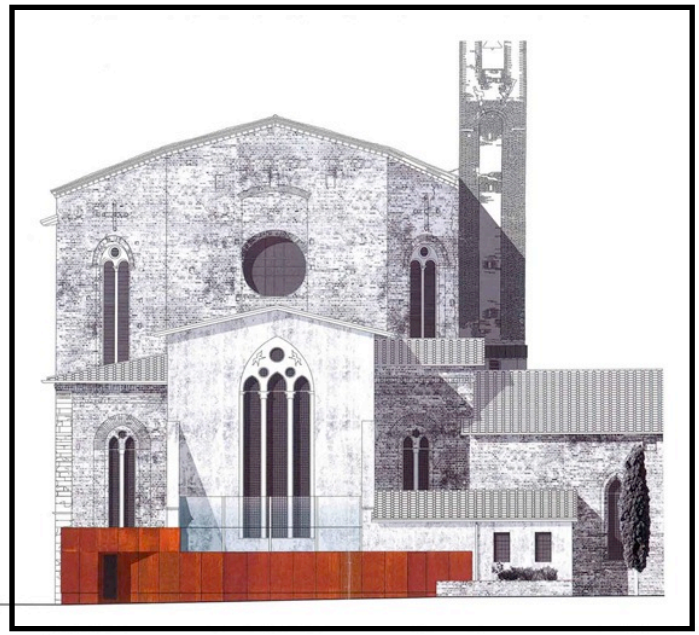
5 ethics issues in team settings:

(i) Ethics of credit
(ii) Parasitic authorship
(iii) Conflicts of interest
(iv) International variations in ethics codes
(v) One-size-fits-all ethics

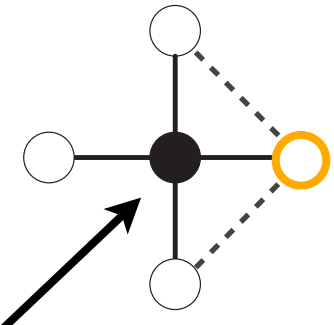


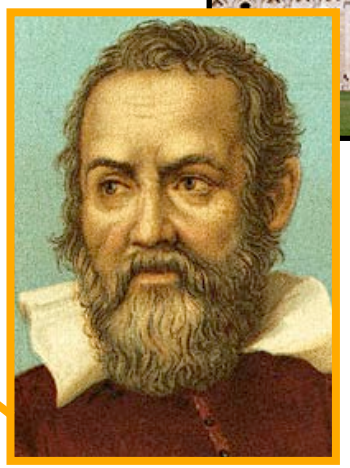
K. Börner, et al. A multi-level systems perspective for the science of team science. Sci. Transl. Med. 2, 49cm24 (2010).

Evolution of scientific networks
In the beginning…



Early academic network





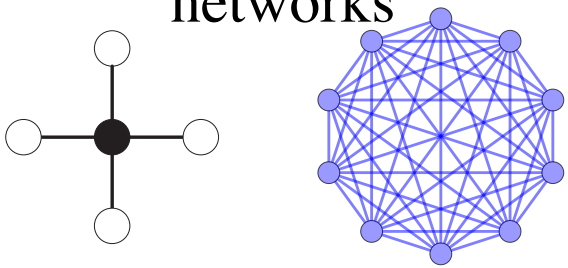
Noble patron (king, wealthy aristocrat, Pope)

Galileo Galilei

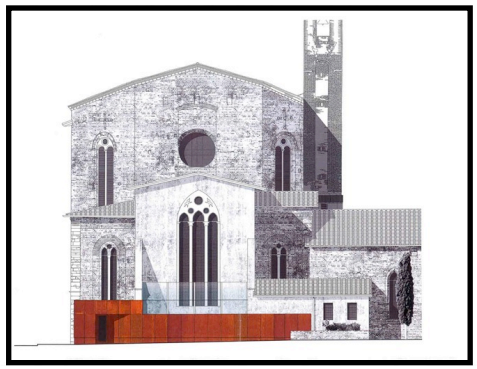
Consider that the norms of scientific precedence/credit emerged during a period of "singleton" science, and not (big) team science

Paul A. David. *The Historical Origins of 'Open Science': An essay on patronage, reputation, and common agency contracting in the scientific revolution.* Capitalism and Society 3(2): Article 5 (2008).

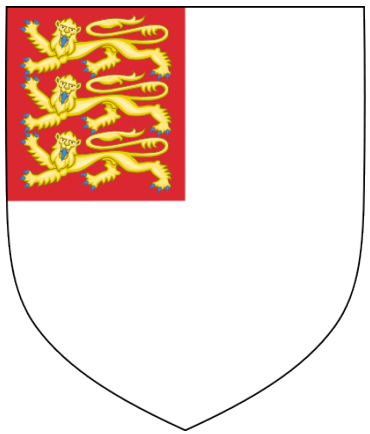
Limited complexity in small collaboration networks



Early scholarly societies, e.g. national societies, scholastic monasteries, noble courts



The Royal Society of London for Improving Natural Knowledge, Established 1660

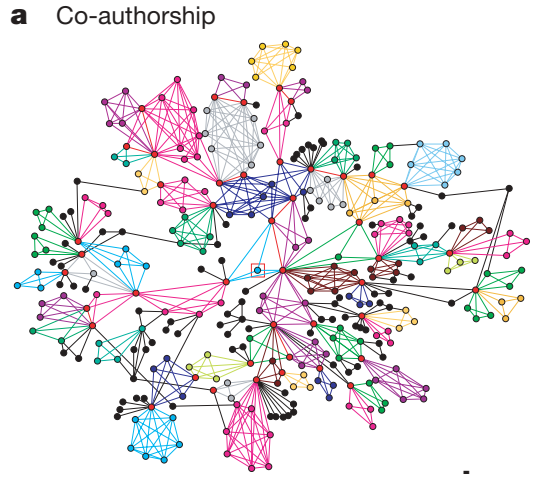


Growth, increasing complexity

⇒

New issues in team science ethics

Vast complexity in large collaboration networks



G. Palla, A.-L. Barabasi, T. Vicsek. Quantifying social group evolution. Nature 446, 664-667 (2007)

S. Wuchty, B. F. Jones, B. Uzzi. The increasing dominance of teams in production of knowledge. Science 316, 1036-9 (2007)

Urban property

210 acres (85 ha) (Main campus)
21 acres (8.5 ha) (Medical campus)
360 acres (150 ha) (Allston campus)
4,500 acres (1,800 ha) (other holdings)

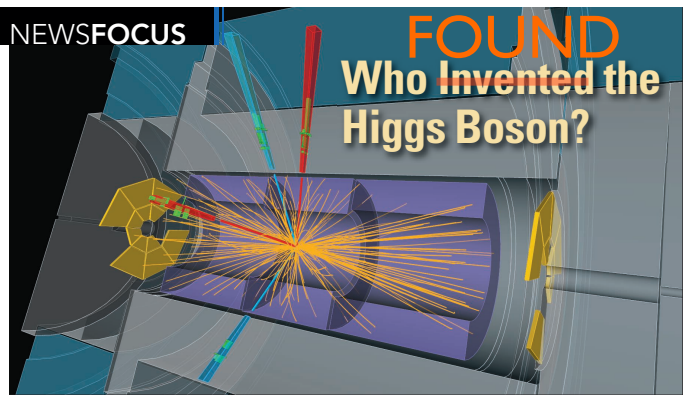
Academic staff	**Admin. staff**
2,100	2,500 non-medical
	11,000 medical

Endowment

US$30 billion (2012) (Large-cap company, e.g. same market capitalization as Enel and Mitsubishi)

Institutions Markets Technologies
IMT
INSTITUTE FOR ADVANCED STUDIES LUCCA

Institutional context: *Increasing team size, cross border collaboration, & changing incentive system*



Five living theorists have claims to having dreamed up the most famous subatomic particle in physics. But what did they really do?

Kingdom. Others question whether the advance was a big enough step beyond previous work to merit science's biggest prize.

14 SEPTEMBER 2012 VOL 337 **SCIENCE** www.sciencemag.org

"50-way tie for the Nobel Prize"

www.sciencemag.org **SCIENCE** VOL 336 6 APRIL 2012

Published by AAAS

CITATION IMPACT 9 DECEMBER 2011 VOL 334 **SCIENCE**

Saudi Universities Offer Cash In Exchange for Academic Prestige

Two Saudi institutions are aggressively acquiring the affiliations of overseas scientists with an eye to gaining visibility in research journals

SCIENCE POLICY 5 AUGUST 2011 VOL 333 **SCIENCE**

Changing Incentives to Publish

Chiara Franzoni,[1] Giuseppe Scellato,[2,3] Paula Stephan[4,5,6]*

genome.gov
National Human Genome Research Institute
National Institutes of Health

ATLAS EXPERIMENT

Compact Muon Solenoid experiment at CERN's LHC

ENCODE

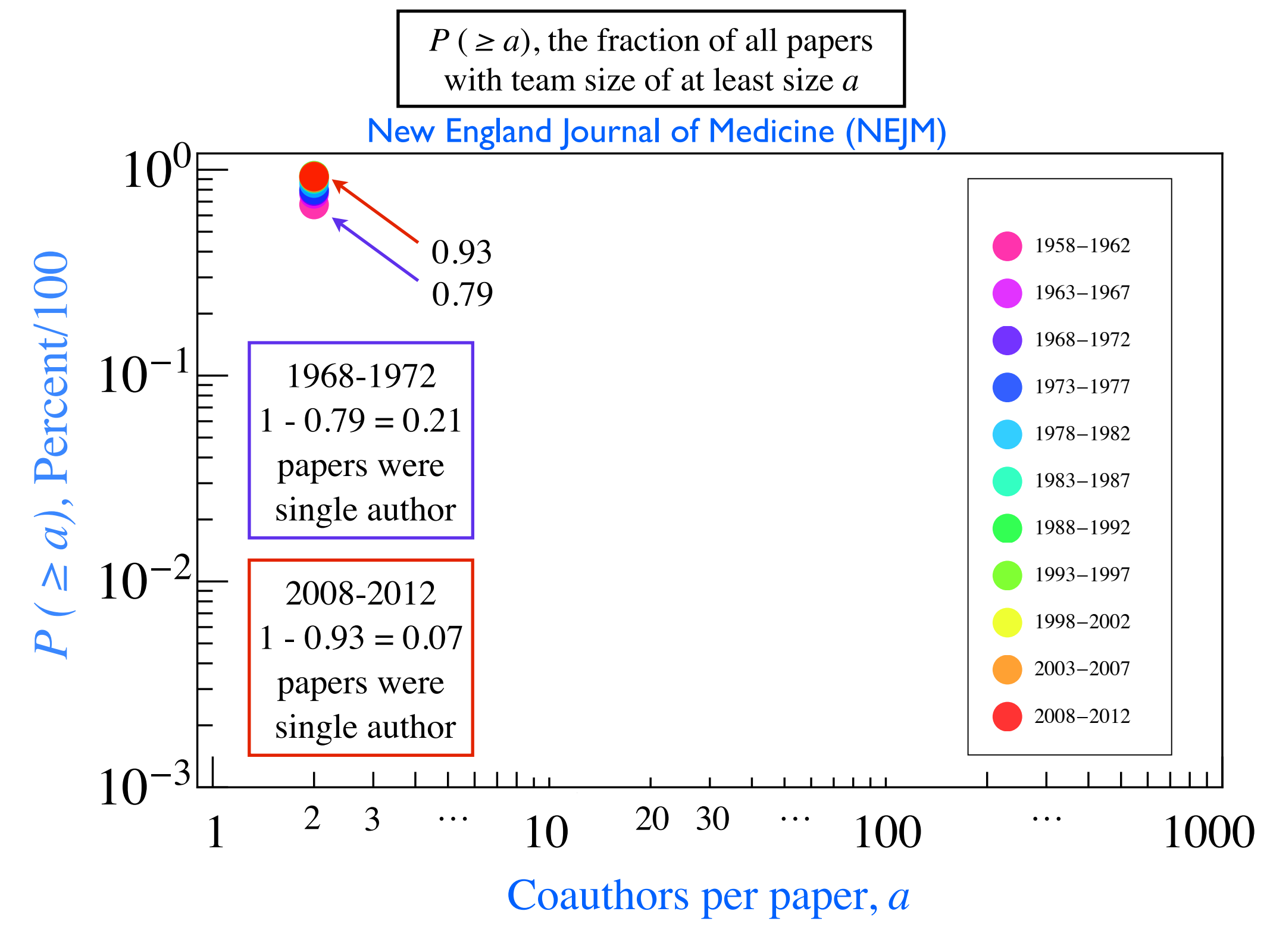
"The CMS experiment is one of the largest international scientific collaborations in history, involving more than 3000 scientists, engineers, and students from 172 institutes in 40 countries. Now, the data collected by CMS are being delivered to institutes around the world to be analysed." - CMS website

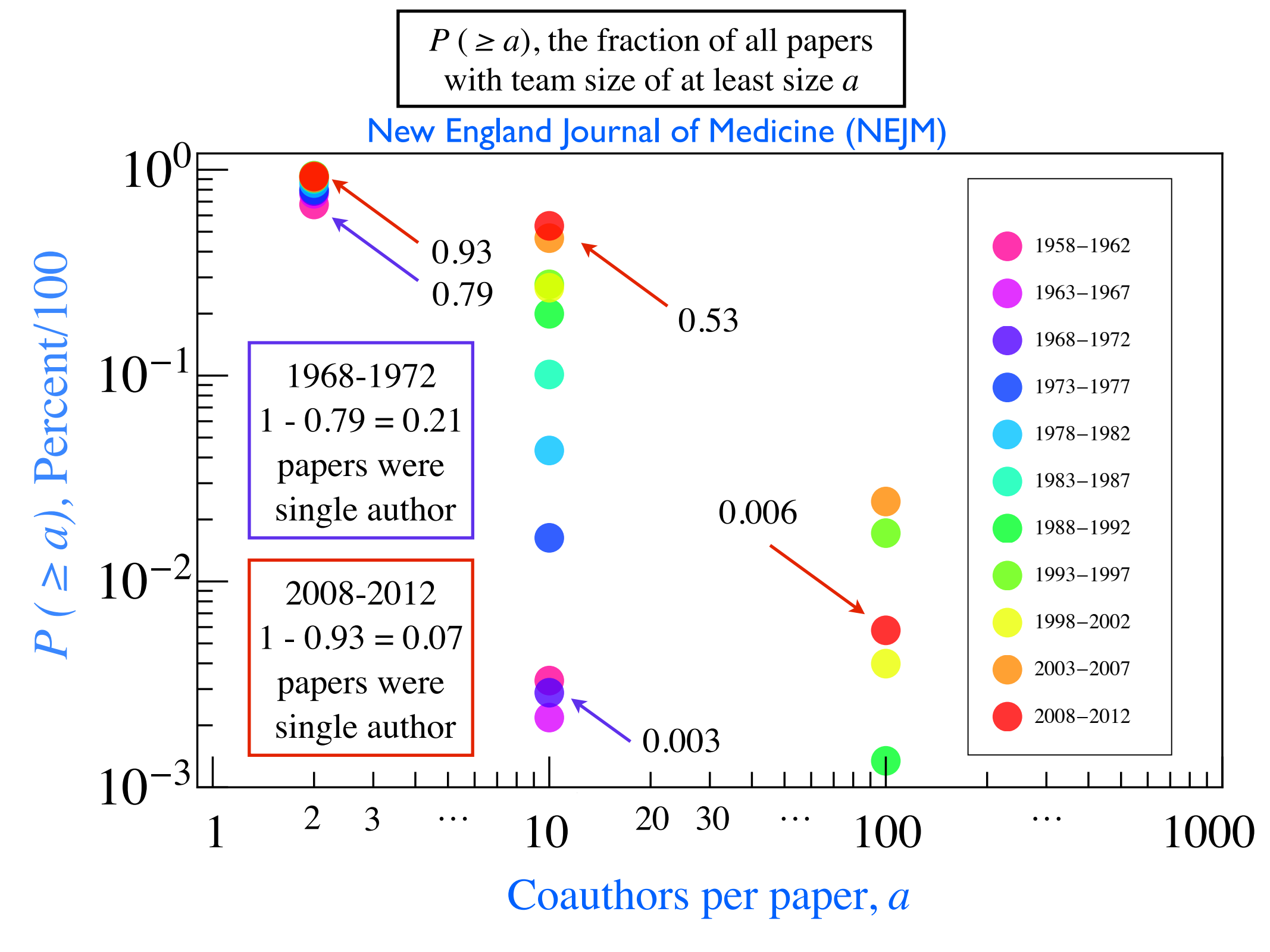
Institutions Markets Technologies
IMT
INSTITUTE FOR ADVANCED STUDIES LUCCA

Computational Perspective

What can we learn
from a "big data" perspective
on the increasing role of large teams in Science ?

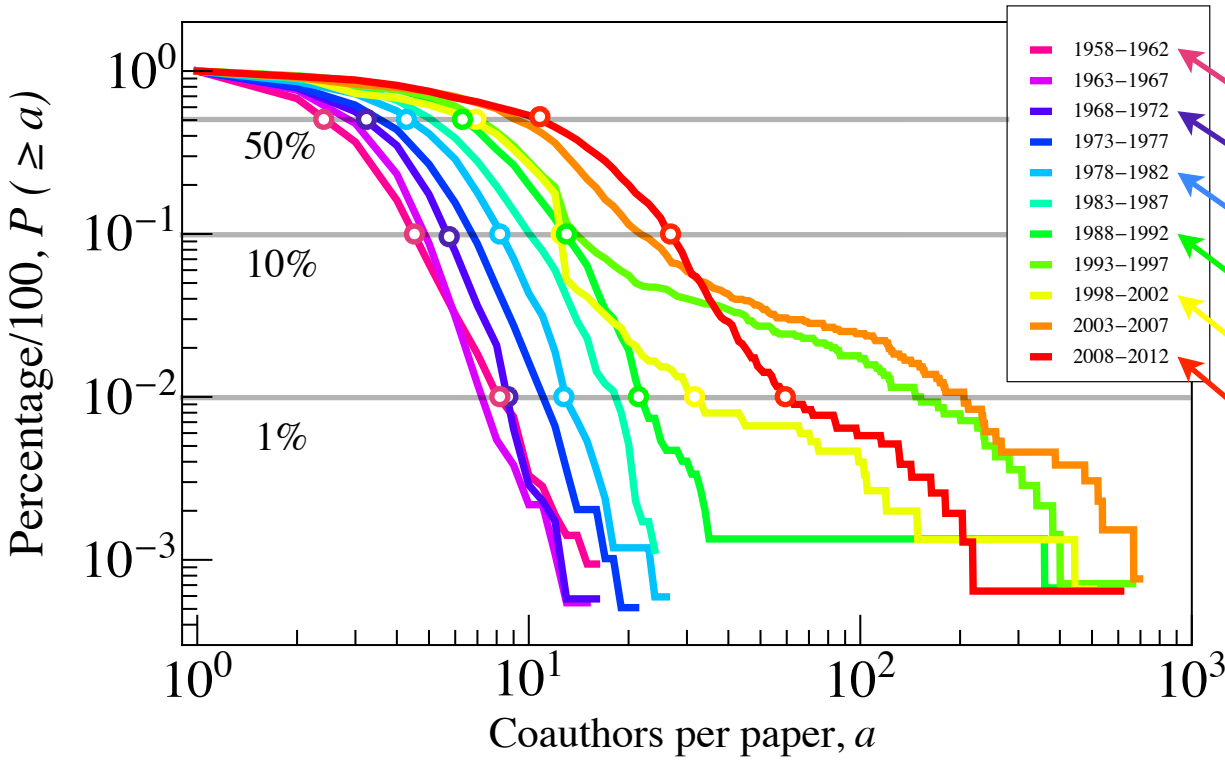
Given the growth patterns, what are the
implications on role of ethics in large team settings?


$P\,(\geq a)$, the fraction of all papers with team size of at least size a
New England Journal of Medicine (NEJM)
$P\,(\geq a)$, Percent/100
10^{0}
10^{-1}
10^{-2}
10^{-3}
0.93
0.79
1968-1972
1 - 0.79 = 0.21
papers were
single author
2008-2012
1 - 0.93 = 0.07
papers were
single author
1958–1962
1963–1967
1968–1972
1973–1977
1978–1982
1983–1987
1988–1992
1993–1997
1998–2002
2003–2007
2008–2012
1
2
3
⋯
10
20
30
⋯
100
⋯
1000
Coauthors per paper, a


$P\,(\geq a)$, the fraction of all papers with team size of at least size a
New England Journal of Medicine (NEJM)
$P\,(\geq a)$, Percent/100
10^{0}
10^{-1}
10^{-2}
10^{-3}
0.93
0.79
0.53
0.006
0.003
1968-1972
1 - 0.79 = 0.21
papers were
single author
2008-2012
1 - 0.93 = 0.07
papers were
single author
1958–1962
1963–1967
1968–1972
1973–1977
1978–1982
1983–1987
1988–1992
1993–1997
1998–2002
2003–2007
2008–2012
1
2
3
…
10
20
30
…
100
…
1000
Coauthors per paper, a

Connecting the dots reveals the persistent growth of team size in R&D

New England Journal of Medicine (*NEJM*)



50%, 10%, and 1% of team sizes are greater than:

NEJM

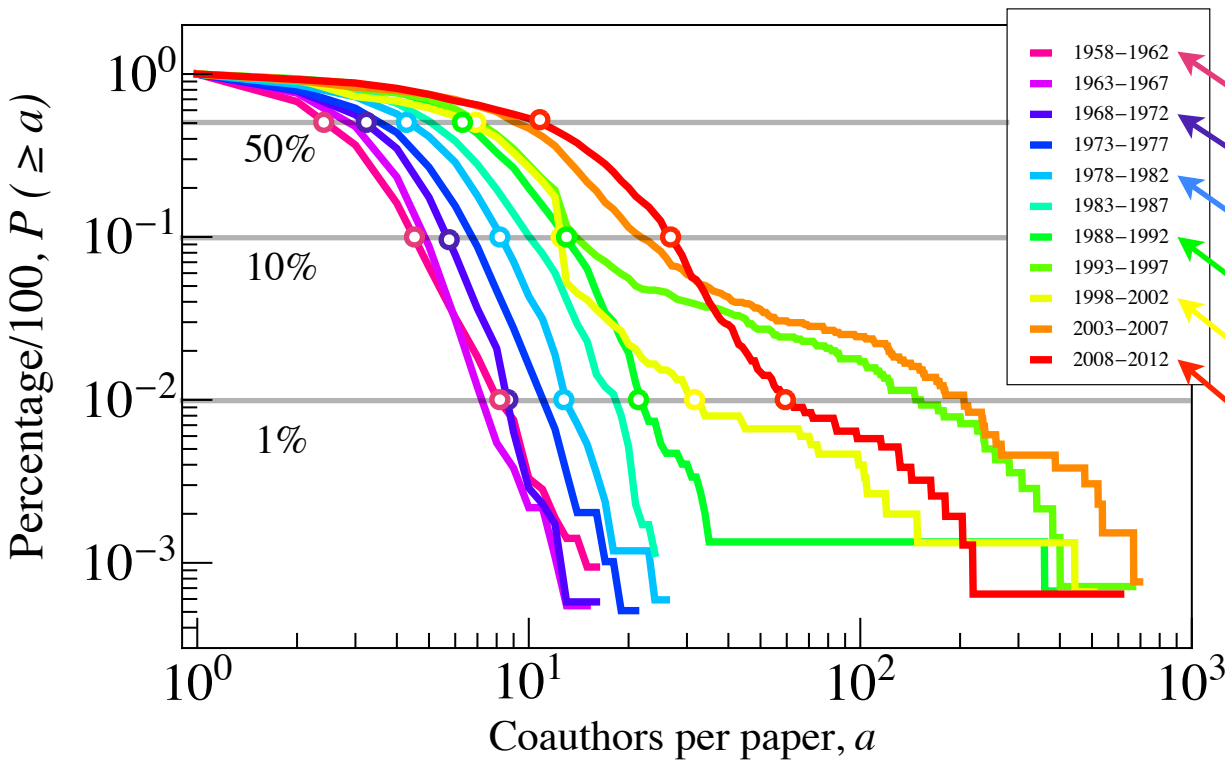
	50%	10%	1%
a =	2	4	8
a =	3	5	8
a =	4	8	12
a =	6	12	21
a =	6	12	30
a =	10	26	60

Physical Review Letters

	50%	10%	1%
a =	2	4	10
a =	2	5	13
a =	3	8	38
a =	3	7	69
a =	3	8	226
a =	4	11	550

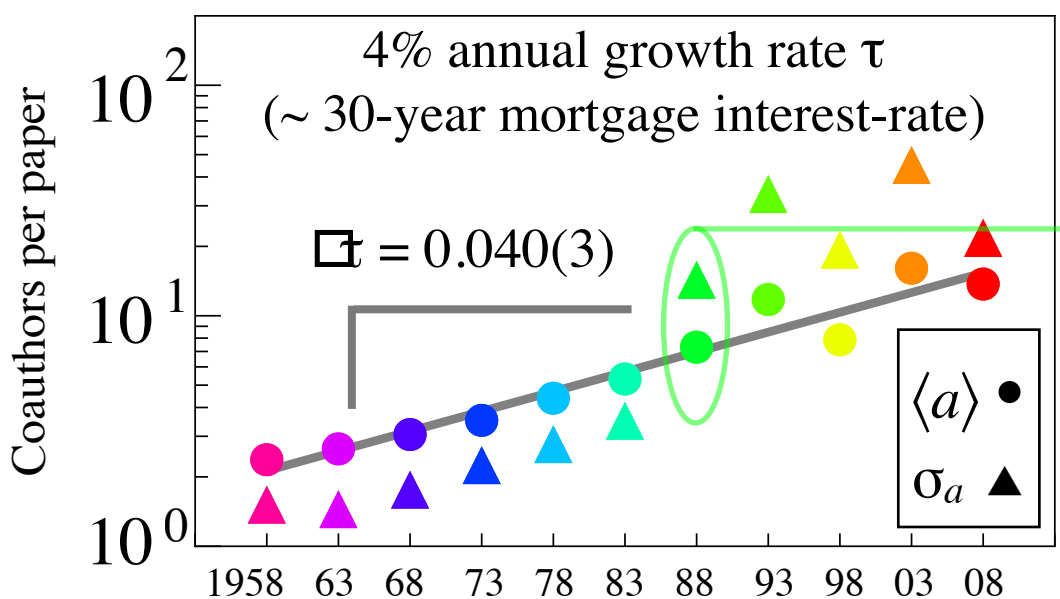
Connecting the dots reveals the persistent growth of team size in R&D

New England Journal of Medicine (*NEJM*)

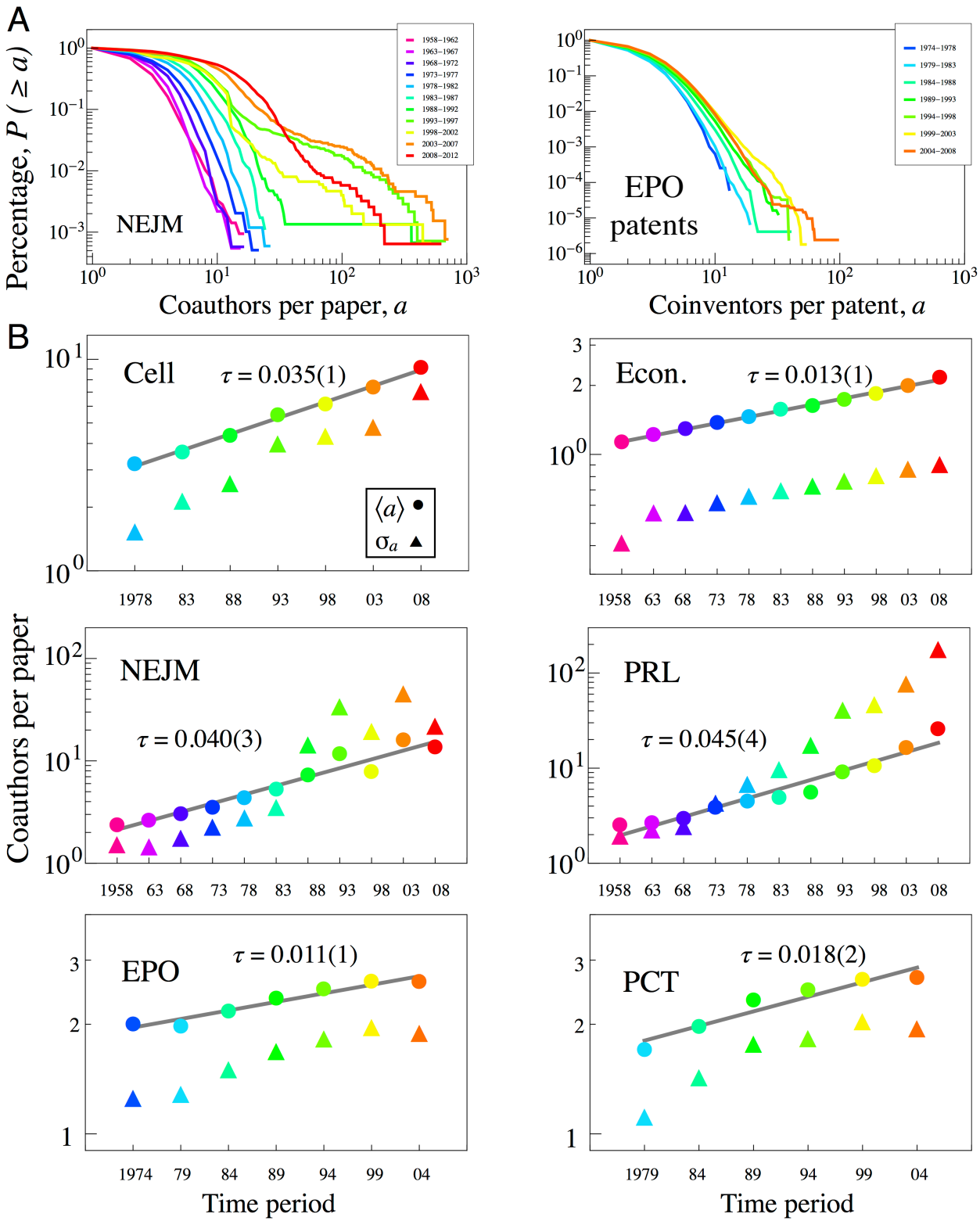


50%, 10%, and 1% of team sizes are greater than:

	50%	10%	1%
$a =$	2	4	8
$a =$	3	5	8
$a =$	4	8	12
$a =$	6	12	21
$a =$	6	12	30
$a =$	10	26	60



Arrival of big-team science in ~ 1998-1992 when the standard deviation first exceeded the average, $\sigma_a > \langle a \rangle$



Big data approach

Dataset	Years analyzed	Articles / Patents	Team size growth rate τ
Cell	1978 – 2012	11,637	0.035(1)
Economics journals	1958 – 2012	36,466	0.013(1)
New England J. Med. (NEJM)	1958 – 2012	18,347	0.040(3)
Phys Rev. Lett. (PRL)	1958 – 2012	98,739	0.045(4)
OECD EPO (patent)	1974 – 2008	2,207,204	0.011(1)
OECD PCT (patent)	1979 – 2008	1,695,339	0.018(2)

What does the future possibly hold?

For example, if we extend the growth trend observed for the journal *Cell* over the last 35 years, extrapolating to the year 2050, the average team size is roughly 40 coauthors per paper.

For comparison, repeating the same extrapolation for the European Patent Office (EPO) growth trend, suggests that by 2050 the average patent will have roughly 4.4 coinventors, the same average team size for Cell publications in 1988.

Ethics in science from a team perspective

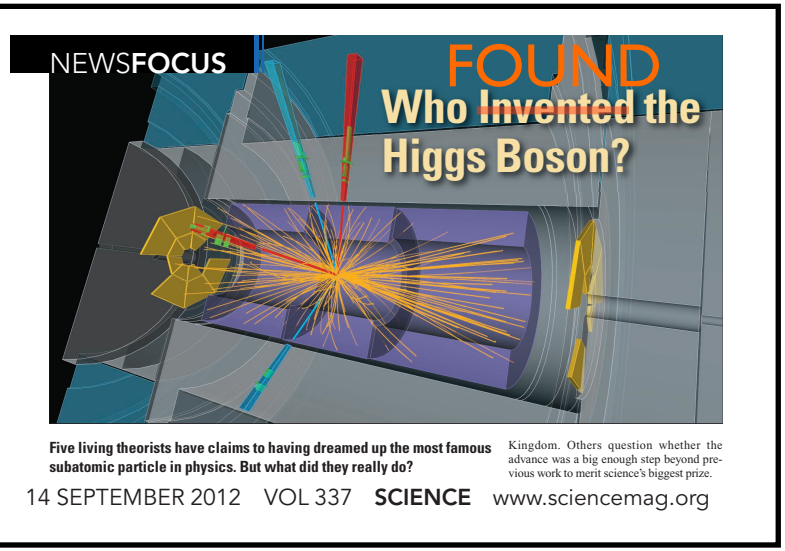
1. *Ethics of credit* and the problem of how to attribute / reward achievement, especially in the case of extremely large team projects
2. *Parasitic coauthorship* and free-riding in team science
3. *Conflicts of interest* increase non-linearly with team size
4. *International variations* in ethics codes create dilemmas in multi-national teams
5. *One-size-fits-all* and the universality of team ethics

Many of these issues stem from the competitive features of an growing scientific system:

strategizing and extreme behaviors emerge (e.g. scientific fraud, Cognitive Enhancing Drugs (CEDs), parasitic coauthorship) which reveal the price of success

I: Ethics of credit distribution in large team science

Case example: the hunt for the Higgs Boson



Consider, for example, the recent dilemma of crediting the six scientists, who in a series of 3 papers in 1964, predicted the existence of a Higgs Boson particle, a fundamental but elusive component of the "standard model" of elementary particles.

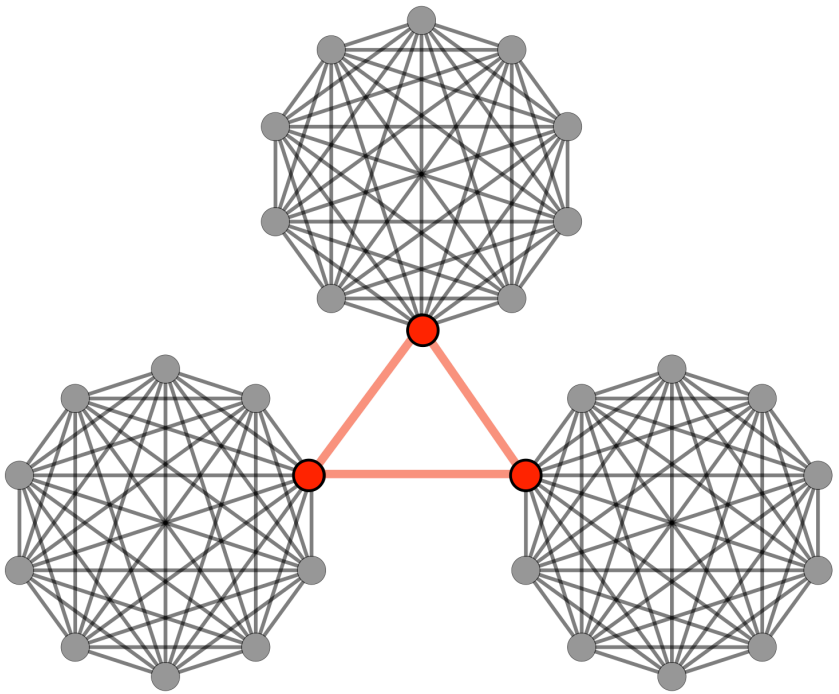
The three seminal papers published in the journal *Physical Review Letters* were submitted only 104 days apart and were published only 77 days apart, a relatively short time period given the pace of publication 60 years ago. Nevertheless, only two of the theoreticians, Francois Englert and Peter Higgs, were awarded the Nobel prize in Physics in 2013.

"The CMS experiment is one of the largest international scientific collaborations in history, involving more than 3000 scientists, engineers, and students from 172 institutes in 40 countries. Now, the data collected by CMS are being delivered to institutes around the world to be analysed." - CMS website

The award to only two individuals is further complicated when juxtaposed to the two large detector teams, the ATLAS detector and the CMS detector, each comprising 1000s of scientists, which orchestrated in tandem the monumental discovery of the Higgs particle using the LHC collider (10 years to construct, ~ 9 Billion USD$)

I: Ethics of credit distribution in large team science

The reward system in science developed during a period when teams were relatively small. Hence, there is an inherent difficulty in distributing fairly sliced credits in large modular teams comprised of heterogenous members



$a = 30, N = 138$

2008-2012

NEJM (Medicine), $P(\geq 30) = 0.065$

PRL (Physics), $P(\geq 30) = 0.040$

Cell (Biology), $P(\geq 30) = 0.017$

Cutting the "credit pie" fairly: Who gets credit? "Who's on first"?

Coauthorship: disciplinary norms

Distinguished credit:
- first author(s)
- corresponding (principal) author(s)
- specific credits: designed, performed research, contributed new reagents/analytic tools, analyzed data, wrote the paper...

Citation (impact) credit:
- shared equally amongst *a* coauthors

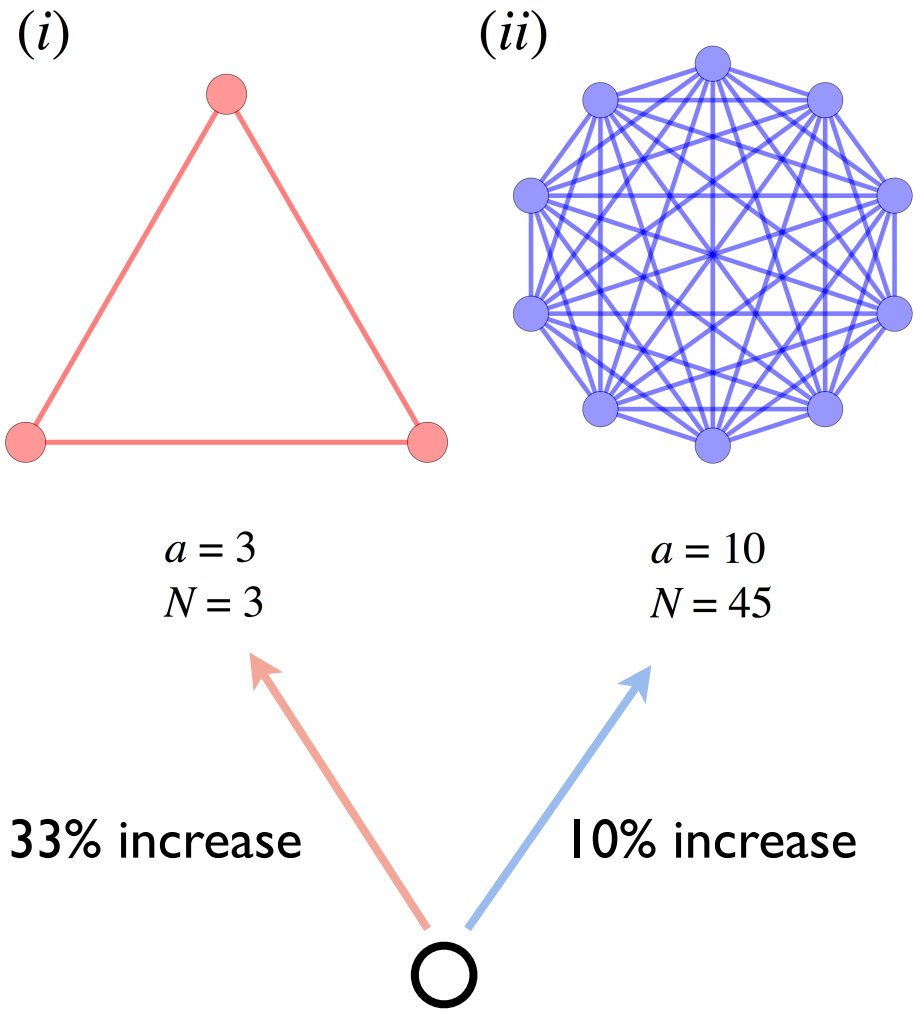
Fraud/Retraction anti-credit:
- can impact all *a* coauthors

~ factor of 20 increase in retractions from 2000 - 2010

The retraction penalty: Evidence from the web of science. Lu SF, Jin GZ, Uzzi B, Jones B.
Scientific Reports 3, 3146 (2013).

II: Parasitic coauthorship: free-riding in large team science

In a large team setting, it is increasingly difficult to discriminate coauthor list order, and furthermore, to determine what merits inclusion on the coauthor list.



The addition of a single coauthor, from $a \rightarrow a + 1$, appears to be only a marginal modification when a is large.

Parasitic coauthorship, "the White Bull effect"

Some senior researchers take advantage of team coauthorship culture by directly manipulating "*experience and deviousness to exploit uncertainties or ambiguities in research guidelines and prosper in poorly regulated, grey areas*."

> The White Bull effect: abusive coauthorship and publication parasitism. L. S. Kwok. *J. Med. Ethics* 31, 554–556 (2005).

What constitutes authorship?

- disciplinary dependent
- grey areas: common for international teams to employ "english helpers", does this deserve credit as help or coauthor?

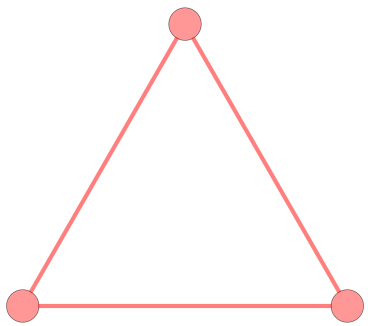
III: Conflicts of interest with ethics implications

- Incentives to (not) publish results may differ between mentor and trainee

- Breakdown of virtue ethics in large teams: the generational transmission of academic character and ethics, is diluted when the ratio of trainees to mentors is too large

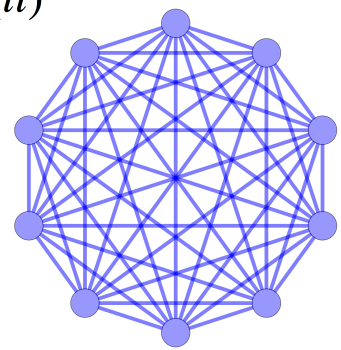
- "Keeping the machine running": Large team projects require efficient management of research scientists. If scientist turnover is high, then there is an inefficient burden to recruit/train replacements. This can lead to a "post-doc" trap, whereby young scientists are lured and kept within big projects by (dishonest) overstated career prospects. Are the next crop of scientists being trained to be leaders or to just fit into a large machine?

(*i*)



$a = 3$
$N = 3$

(*ii*)



$a = 10$
$N = 45$

The maximum number of dependencies N in a team of size a is given by $N \approx a^2/2$. These dependencies (links) represent the multitude of associations, and potential conflicts of interest, between the team members.

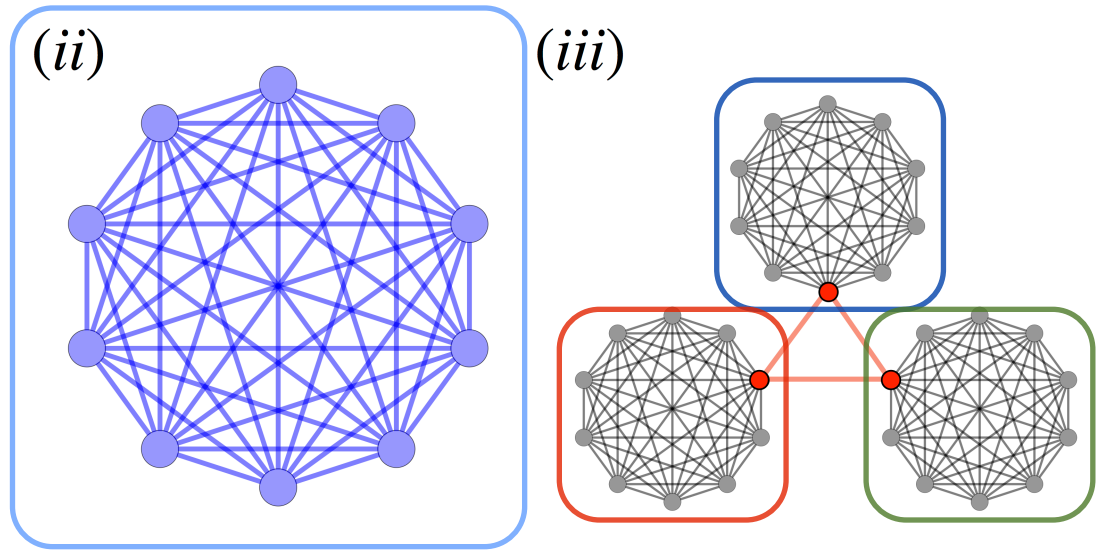
IV: International variations in ethical norms

The norms of leadership, management, and promotion can be largely country dependent. Moreover, the standards for ethical conduct in science and the laws reflecting bioethical standards on research topics involving stem cells, experiments with animals, and human clinical trials, can vary significantly across counties. With collaborations becoming increasingly global and interdisciplinary, the standardization of ethical norms is crucial.

- intense localization of proprietary biomedical R&D in countries with less restrictive bio-ethics legislature

- outsourcing of clinical trials is a prime example of side-stepping potential ethical impediments by going abroad

Single set of ethical norms / laws



Multiple sets of ethical norms / laws

V: Universality of team ethics across varying team size?

Should we expect the ethics of small team science to map across scale and apply unflinchingly to large team science?

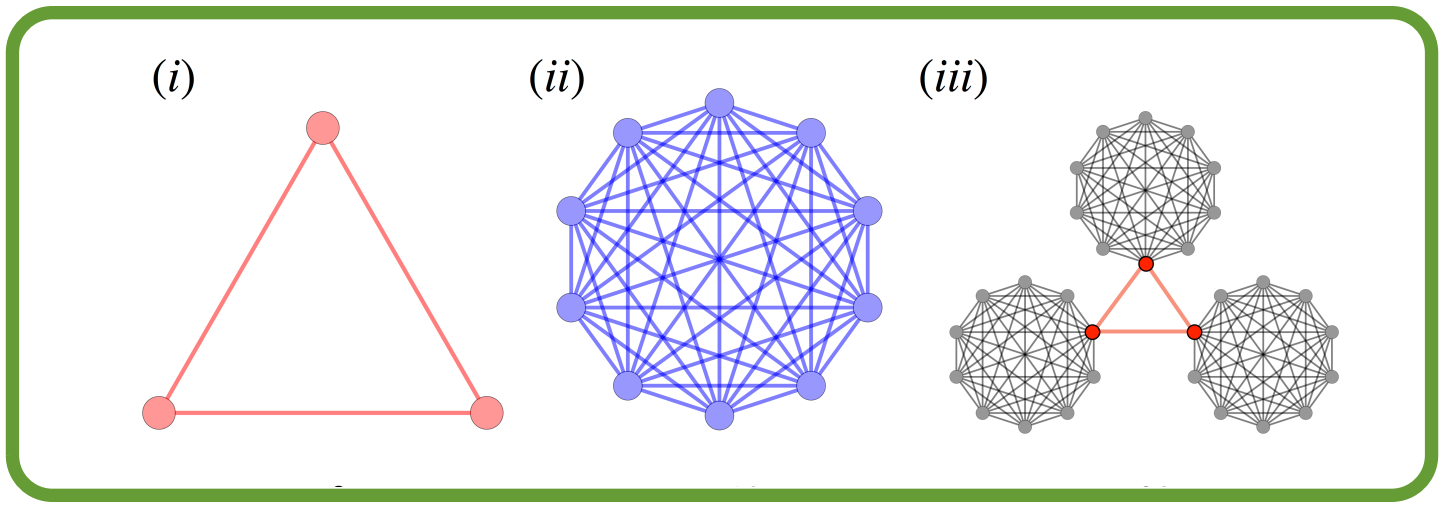
Several features of large team science challenge the institutions constructed for small team science:

- the reproducibility of such large projects (inherently requiring complementary large teams committed to verification)

- the distribution of (equal) credits to all participants.

- the growth of interdisciplinarity science is a factor contributing to larger team sizes. Is it possible to expect that social norms of ethical conduct in publication to be shared by different disciplines?

?



?

Institutions Markets Technologies
IMT
INSTITUTE FOR ADVANCED STUDIES LUCCA

Food for thought

- Steady growth of team size across private and public R&D challenges longstanding system of credit and ethics in science
- Large teams are composed of heterogenous structure, leading to difficulties in allocating credit, in sanctioning unethical behavior, and in discouraging parasitic coauthorship. Also, the potential number of conflicts of interest increase dramatically with team size.
- Does a universal set of ethics apply across teams of all sizes? Also, will large teams contribute to the erosion of ethical standards in science? Nevertheless, despite these problems, the burden of consciousness and morality lies in the individual, independent of team size.
- A body of ethical scientists is indeed an invaluable community resource since the adoption of norms is a self reinforcing process, gaining strength with size. Team ethics policies should emphasize a bottom-up approach consisting of producing ethical scientists through directed education, starting with a student's first introduction to science in secondary school, in a very general sense, the training of team science ethics should become a corollary of the longstanding scientific method.

A computational perspective on the ethics of big team science,
A. M. Petersen, I. Pavlidis, I. Semendeferi.
In preparation (2013) *http://physics.bu.edu/~amp17/*

Thank You!
A special thanks to my collaborators:
I. Pavlidis, I. Semendeferi

Being Ethical in Large-Team Science:
A Quantitative Historical Perspective

Alexander M. Petersen,[1] Ioannis Pavlidis,[2] and Ioanna Semendeferi[3]

[1]*Laboratory for the Analysis of Complex Economic Systems,*
IMT Lucca Institute for Advanced Studies, Lucca 55100, Italy
[2]*Department of Computer Science,University of Houston, Houston 77204, USA*
[3]*College of Natural Sciences and Mathematics,*
University of Houston, Houston 77204, USA

(Dated: November 7, 2013)

Abstract

Since the middle of the twentieth century, science has increasingly become a large-team endeavor. The gradual crowding out of singleton and small-team science challenges key features of the scientific system, such as assignment of recognition, incentives to collaborate and share, and the moral responsibility of the scientist. In a system dominated by large teams, it becomes increasingly difficult to distribute credit, to efficiently allocate funding, and to fairly award career achievement prizes. It becomes equally difficult to assign blame: Who is to be held responsible in ethical breaches, and who is responsible for monitoring ethical standards and sanctioning in a system composed of large groups? And, since science is becoming increasingly international and interdisciplinary, is it conceivable to apply a universal set of ethical standards across endeavors of all sizes and all disciplines in all countries? Further complicating things, the hierarchical structure of large teams strengthens the stratification mechanisms that distinguish the "haves" and "have-nots" in science, leaving a system which is not necessarily fair. Using expository analysis of historical publication and patent datasets to quantify the growth of team size and the distribution of scientific credit, we discuss various ethics issues in the context of large-team science. Our analysis begs the question of whether or not the ethics in singleton and small-team science is relevant to the understanding of ethics in large-team science. Our analysis renders quantitative support and complements qualitative analysis that has been ongoing in the history of science.

Cognizant Enhancement Drugs (CED)

Professor's little helper

The use of cognitive-enhancing drugs by both ill and healthy individuals raises ethical questions that should not be ignored, argue **Barbara Sahakian** and **Sharon Morein-Zamir**.

NATURE|Vol 450|20/27 December 2007

" Is it cheating to use cognitive-enhancing drugs?.... How would you react if you knew your colleagues — or your students — were taking cognitive enhancers?... we know that a number of our scientific colleagues ... already use modafinil [Modiodal, Provigil] to counteract the effects of jetlag, to enhance productivity or mental energy, or to deal with demanding and important intellectual challenges..."

NATURE|Vol 452|10 April 2008

Poll results: look who's doping

In January, *Nature* launched an informal survey into readers' use of cognition-enhancing drugs. **Brendan Maher** has waded through the results and found large-scale use and a mix of attitudes towards the drugs.

"One in five respondents said they had used drugs for non-medical reasons to stimulate their focus, concentration or memory. Use did not differ greatly across age-groups..., which will surprise some. "

Towards responsible use of cognitive-enhancing drugs by the healthy

Society must respond to the growing demand for cognitive enhancement. That response must start by rejecting the idea that 'enhancement' is a dirty word, argue **Henry Greely and colleagues**.

"...one survey estimated that almost 7% of students in US universities have used prescription stimulants [Adderall and Ritalin] in this way, and that on some campuses, up to 25% of students had used them in the past year. These students are early adopters of a trend that is likely to grow, and indications suggest that they're not alone."

NATURE|Vol 456|11 December 2008

I: Ethics of anti-credit and the price of success

Achievement-oriented systems: incentives for cut-throat "zero-sum" behavior, possibly leading to blatant cheating/falsification, i.e. use of performance / cognitive enhancing drugs

Jan Hendrik Schön Scandal (2001)

On October 31, 2002, *Science* withdrew **eight** papers written by Schön

On December 20, 2002, *Physical Review* withdrew **six** papers

On March 5, 2003, *Nature* withdrew **seven** papers

Diederik Alexander Stapel Scandal (2011)

Social psychologist made up data for at least 30 publications according to preliminary investigation, which is still ongoing.

Hisashi Moriguchi Scandal (2012)

"Transplant of induced pluripotent stem cells to treat heart failure probably never happened.... He is affiliated with University of Tokyo but not with Massachusetts General Hospital nor with Harvard Medical School. The study did not receive Institutional Review Board approval." nature.com

Q: Are large teams more susceptible to ethics breaches?

Q: Whose responsibility is it to sanction unethical behavior in large team science?

Q: And who takes the blame in a large team scandal?