Size matters, except perhaps for pure mathematicians

Ralph Kenna [1] *and Bertrand Berche* [2]

[1] Coventry University, England
[2] Nancy Université, France

CompPhys10, Leipzig, November 2010

RK & B. Berche, EPL 90 (2010) 58002; Scientometrics, published online 24 September 2010.

Concentration or competition ?

“In the current economic conditions we must concentrate investment where it will have the greatest impact – within our world-class research-intensive universities. ”

[Press release from the Russell Group, 09 July 2010]

“Concentrating research funding and doctoral students in just a handful of universities will damage both the UK university sector and the economy as a whole. ... There is no economic or academic case for the concentration of funding for research or doctoral students.”

[Press release from Million+, 02 March 2010]

So which is best : concentration or competition ?

Prelude

Note 1 : Throughout the talk we will refer to the RAE.

This is the Research Assessment Exercise, a peer-review process which is carried out in the UK every 5-7 years. The last RAE was in 2008.

Academia was divided into 15 broad areas consisting of 67 units of assessment (UOA's).

E.g., *applied maths* was one UOA. *Physics* was another.

RAE measures the quality of research carried out by groups at universities.

Subsequent government funding is determined by the results of RAE.

Prelude

Note 2 : We will use word "group" in RAE sense.

E.g., if the Uni Leipzig were in UK, their "Applied Mathematics Group" may be drawn from research-active, permanent staff in

- Institut für Theoretische Physik (Computer-oriented QFT, Molecular Dynamics, QFT & Gravity, Statistical Physics, Particle Theory, Condensed Matter)
- Mathematisches Institut (Numerik, Optimierung/Finanzmathematik, stochastik, . . .)
- plus possibly some postdocs.

Leipzig itself would decide on who to submit to RAE.

Group Quantity & Quality

One may expect that, *on average*, a group of size $N = 10$ is twice as strong as a group of size $N = 5$.

I.e., strength S of a group activity is proportional to size N,

$$S \propto N.$$

Define the quality s of a group activity as the strength per head : s = S/N.

Then we might expect that *on average* that

$$s = S/N \sim \text{constant},$$

i.e., that quality is independent of quantity.

The purpose of this talk is to show that this is not the case !

Group Quantity & Quality

Lets check quality of group activity from RAE 2008.

Denote the research strength of the i^{th} member of the g^{th} research group in a given discipline by a_{g_i}.

This includes the added strenth gained or lost by by

- journal & library facilities, computer & equipment access,
- teaching &administration loads,
- managerial support,
- prestige & confidence inspired by history of institute and individual previous successes,
- extramural collaborations,
- etc.

Group Quantity & Quality

Naively, we expect the strength S_g of a group of size N to be

$$S_g = \sum_{i=1}^{N} a_{g_i} = N\bar{a}_g,$$

where $\bar{a}_g$ is the mean strength of the N individuals.

Define quality as the strength per head : $s_g = \frac{S_g}{N}$.

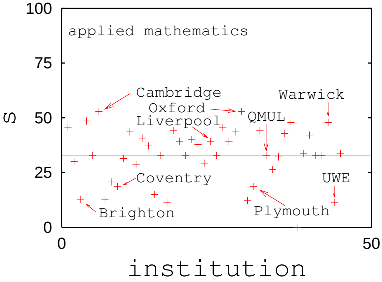
Then

$$s_g = \bar{a}_g$$

so that the quality of the group is the average quality of its individuals.

Let us check this in applied mathematics at RAE 2008.

Group Quantity & Quality



Indeed, $s_g = \text{constant} \pm \text{noise}$.

Quality indeed appears randomly distributed about a mean, with older, prestigious universities tending to be above and newer ones below.

But this is not the correct picture....

Note : If you were at or collaborated with a UK uni between 2001-2008, you may be in this plot (or the physics one) !

Group Quantity & Quality

The previous picture is the basis on which institute & universities are ranked after RAE.

We will see that interpretations based on this are wrong.

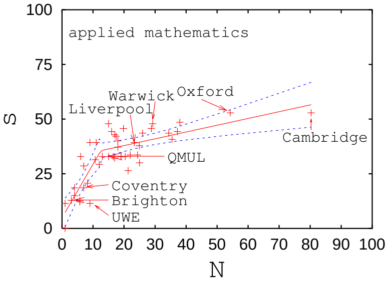
They are also dangerous :
For example Oppenheim & Summers (2008) made the approximation

$$a_{g_i} \approx \bar{a}_g = s_g$$

to estimate individual quality through RAE group quality.

We next plot the data in a different way : quality against quantity...

Group Quantity & Quality



Quality s is clearly correlated with groups size N : $s = s(N)$.

What is the nature of this correlation ?

Group Quantity & Quality

We have to treat research groups as complex systems. With

$$b_{g\,\langle i,j\rangle} = \text{strength of interaction between individuals} \quad i \quad \text{and} \quad j$$

the strength of the group is

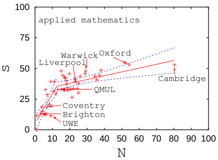
$$S_g = \sum_{i=1}^{N} a_{g_i} + \sum_{\langle i,j\rangle=1}^{N(N-1)/2} b_{g\,\langle i,j\rangle} = N\bar{a}_g + \frac{1}{2}N(N-1)\bar{b}_g,$$

where $\bar{b}_g$ is the mean 2-way interaction strength in group g.

So the quality is $s_g = \bar{a}_g + \frac{1}{2}(N-1)\bar{b}_g$.

Then the expected quality over all groups takes the form

$$s = a_1 + b_1 N.$$



Group Quantity & Quality

However, two-way communication cannot be carried out between every pair of nodes if N is too large, say above N_c.

If $N > N_c$, *subgroups* or *cliques* or *blobs* form.

If there are $N/(\alpha N_c)$ subgroups of mean size αN_c, and if the subgroups interact with strength β_g, then

$$S_g = \underbrace{N\bar{a}_g + \frac{1}{2}N(\alpha N_c - 1)\bar{b}_g}_{\text{intra-subgroup}} + \underbrace{\frac{1}{2}\frac{N}{\alpha N_c}\left(\frac{N}{\alpha N_c} - 1\right)\beta_g}_{\text{inter-subgroup interaction}} .$$

I.e., we expect a different linear dependency of quality s on size N (like a phase transition !).

Note that the quadratic term in S or linear term in s is proportional to $1/N_c^2$

Group Quantity & Quality

We therefore expect

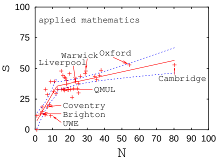
$$s = \begin{cases} a_1 + b_1 N & \text{if } N \leq N_c \\ a_2 + b_2 N & \text{if } N \geq N_c . \end{cases}$$

If $N > N_c$, we call the group large.

If $N < N_c$, we call the group small or medium.

Since $b_2 \sim 1/N_c^2$, the slope to the right is small for large N_c.

We have now explained



Critical Mass

Critical mass is traditionally viewed as a threshold beyond which research quality “takes off”.

But no evidence for such a threshold has ever been presented [Thompson Reuters (July 2010)].

Mark Harrison [Warwick, 2009] : “Is there a critical mass in research ? How big is it ? ... I do not claim to know the answer. But I am thoroughly familiar with the question.”

We will now answer the question in two ways.

First we consider the group as a grand canonical ensemble and then as a canonical one.

Critical Mass

The *total* strength of a group is

$$S(N) = a_i N + b_i N^2 \quad \text{where } i = 1 \text{ or } 2 \text{ if } N < N_c \text{ or } N > N_c.$$

If M new nodes become available where should they be allocated?

$$S(N+M) = a_i(N+M) + b_i(N+M)^2.$$

The average increase in societal strength is

$$\frac{\Delta S}{M} = \frac{S(N+M) - S(N)}{M} = a_i + b_i(2N+M)$$

and depends on whether $N < N_c$ or $N > N_c$.

Next we take the limit of this as $M \to 0$.

(I.e., we simply maximize the gradient of $S(N)$.)

Critical Mass

It turns out it is best to support the small/medium group if $N > N_k$ where

$$N_k = \frac{N_c}{2}.$$

If $N < N_k$, we call the group small.

If $N_k < N < N_c$, we call the group medium.

It is best to support the medium group ($N > N_k$). Then to support the large group ($N > N_c$) and then the small one.

Critical Mass

So there are two critical masses in research (for each discipline).

$$\begin{aligned} N_k &= \text{lower critical mass,} \\ N_c &= \text{upper critical mass.} \end{aligned}$$

There is no threshold (as Thompson Reuters have said). But N_k corresponds more closely to the traditional notion of critical mass.

I.e.,

(Lower) critical mass is the minimum size a research team must achieve for it to be viable in the longer term.

Mass Transfer

We next consider the transfer of nodes within an academic discipline with total numbers fixed.

Group I (e.g., Coventry) is small/medium and has $N_I < N_c$ staff.

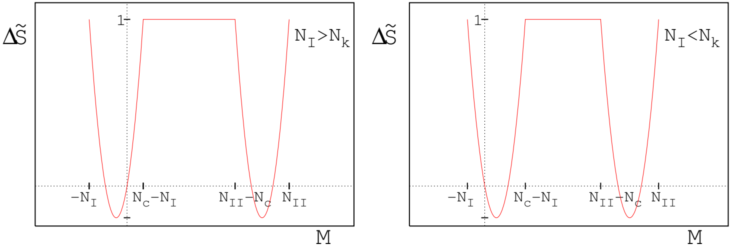
Group II (e.g., Cambridge) is large and has $N_{II} > N_c$ staff.

We transfer M nodes from Group II to Group I.

What is ΔS, the change in strength of the *society*?

Mass Transfer

We assume $b_2 = 0$ (large N_c).



$N_k < N_l < N_c$ $\quad$ $N_l < N_k$

It is sensible to transfer nodes from the large group to the medium one.

How does RAE measures quality ?

At RAE 2008, subject areas were scrutinised to assess the proportion of research work submitted which fell into categories

- 4* : Quality that is world-leading in terms of originality, significance and rigour
- 3* : Quality that is internationally excellent in terms of originality, significance and rigour but which nonetheless falls short of the highest standards of excellence
- 2* : Quality that is recognised internationally in terms of originality, significance and rigour
- 1* : Quality that is recognised nationally in terms of originality, significance and rigour
- Unclassified : Quality that falls below the standard of nationally recognised work

How does RAE measures quality ?

An example of a UOA quality profile is as follows :

Quality level	4*	3*	2*	1*	u/c
$p_{n*} =$ % of research activity	$p_{4*} = 15$	$p_{3*} = 25$	$p_{2*} = 30$	$p_{1*} = 20$	$p_{0*} = 10$

Based upon quality profiles, a formula is used to determine how research funding is distributed. In England in 2009/2010 the funding formula is

$$s = p_{4*} + \frac{3}{7}p_{3*} + \frac{1}{7}p_{2*}.$$

The funding allocated to a groups of size N is then proportional to $S = sN$.

Summary so far

RAE has measured research quality in many academic disciplines in the UK.
We fit the resulting data to the ansatz

$$s = \begin{cases} a_1 + b_1 N & \text{if } N \leq N_c \\ a_2 + b_2 N & \text{if } N \geq N_c \,. \end{cases}$$

This gives us the upper and lower critical masses N_c and N_k for that subject area.
We then divide research groups into

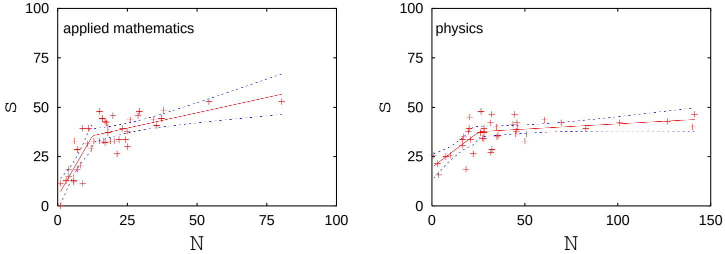
———————————- N_k ———————————- N_c ———————————-

small medium large

It is best to support medium groups, then large ones and then small.

Natural Sciences

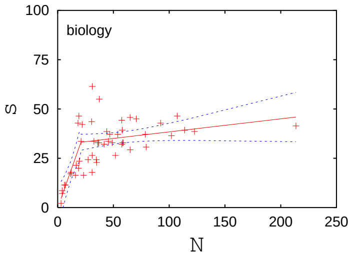
We start the analysis with theoretical/experimental physics

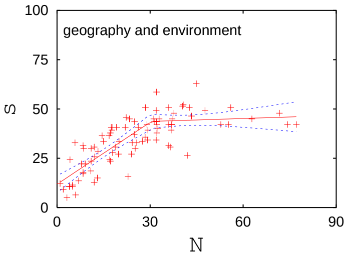


We measure $N_k = 6 \pm 1$
(with $R^2 = 74\%$)

We measure $N_k = 13 \pm 3$
(with $R^2 = 53\%$)

Natural Sciences

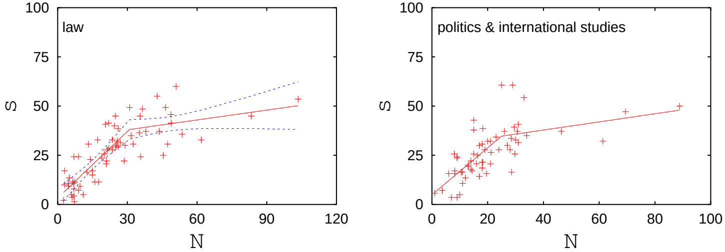




We measure $N_k = 10 \pm 2$ (with $R^2 = 54\%$)

and $N_k = 15 \pm 2$ (with $R^2 = 66\%$).

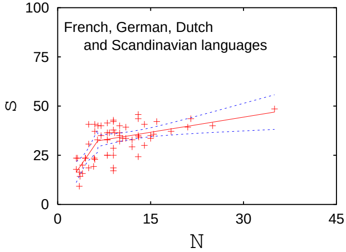
Other subject areas

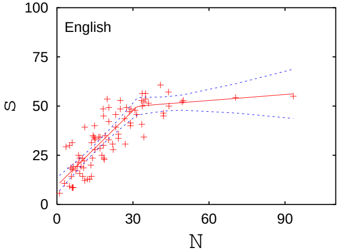


$N_k = 15 \pm 2$

$N_k = 13 \pm 3$

Other subject areas

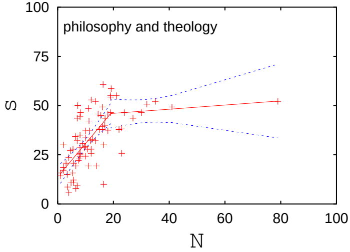


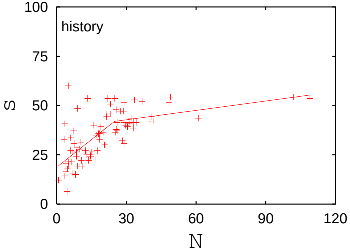


$N_k = 3 \pm 1$

$N_k = 16 \pm 2$

Other subject areas



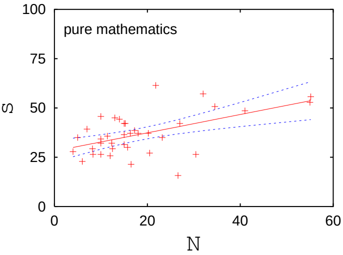


$N_k = 9 \pm 2$

$N_k = 12 \pm 2$

Pure mathematics

In pure mathematics, no breakpoint was found :



A linear fit has relatively small slope and high intercept implying all groups are large.

The smallest group submitted had $N = 4$. So we may estimate $N_c \lesssim 4$ or

$$N_k \lesssim 2.$$

I.e., pure mathematicians work alone or in pairs - size doesn't matter.

Critical masses

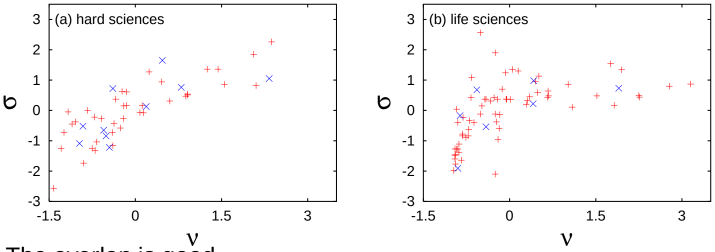
The estimates for the lower critical masses of 24 (combined) subjects are

Subject	$N_k = N_c/2$	Subject	$N_k = N_c/2$
Applied mathematics	6 ± 1	Nursing, etc.	9 ± 3
Physics	13 ± 3	Computer science$^\sharp$	25 ± 5
Earth sciences	15 ± 2	Archaeology$^\sharp$	8 ± 2
Biology‡	10 ± 2	Economics	5 ± 2
Chemistry	18 ± 7	Business	24 ± 4
Agriculture, vet. etc‡	5 ± 2	Politics	13 ± 3
Law†	15 ± 2	Sociology	7 ± 2
Architecture & planning	7 ± 2	Education	15 ± 3
Non-English languages	3 ± 1	History‡	12 ± 3
English	16 ± 2	Philosophy & theology	9 ± 2
Pure mathematics*	≤ 2	Art & design‡	12 ± 4
Medical sciences	20 ± 4	Other arts	4 ± 1

The French System (AERES)

The data presented above is from the UK's RAE.
Is the phenomenon more general ?
We plot the standardised data for the hard and life sciences in the UK (red) and France (blue), where

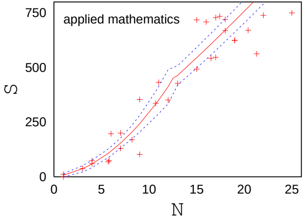
$$\sigma = \frac{s - \bar{s}}{\sigma_s} \qquad \nu = \frac{N - \bar{N}}{\sigma_N}$$



The overlap is good.

Absolute Strength

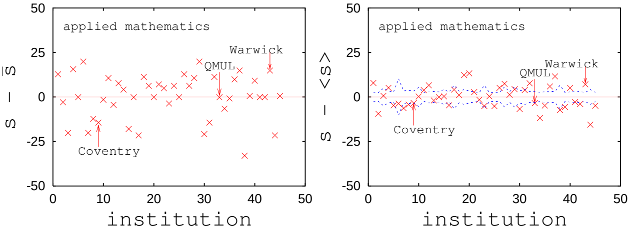
To reiterate the main point concerning the correlation between quality and quantity (group size), we plot the *absolute* strength S for applied mathematics (in the UK) as a function of N to show the breakpoint :



Without the breakpoint, the steepest gradient condition would lead to supporting only the biggest among the large groups so that only one group remains in the end.

Intra-Disciplinary Renormalization

For applied mathematics, we plot $s - \bar{s}$ and $s - \langle s \rangle$ against N, highlighting the various university groupings

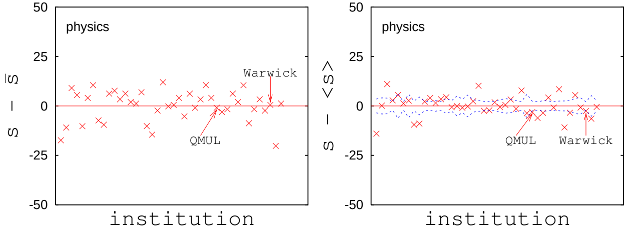


Left plot represents deviation of data from *global average*. St. dev. = 12.6, range = 52.9

Right plot represents deviation of data from *local average*. St. dev. = 6.4, range = 28.8

Intra-Disciplinary Renormalization

Same plots for physics :



Left plot (*global average*) : St. dev. = 7.8, range = 32.1

Right plot (*local average*) : St. dev. = 5.3, range = 25.1

Left line separates groupings ; right line goes through groupings

Conclusions

We explain why the average performance of larger groups appears to exceed those in smaller groups. We also determined critical masses.

- Community is greater than sum of it parts ...
- ...an effect which saturates beyond N_c
- It is unwise to judge a group soley on quality profile : small/medium groups should not be expected to yield same strength as large ones
- It is unwise to judge individuals in a group on the basis of the group strength as this neglects interactions
- Need to take size into account to determine which groups are punching above or below their weight

Conclusions

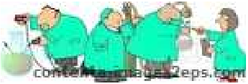
Strength is primarily ascribed to two-way communication links.

Universities should facilitate this

- Hotdesking bad !
- Distance working is bad !
- Collaboration is good !
- Medium-sized groups should be supported
- Small groups must endeavour to achieve (lower) critical mass

Final Message

A medium sized group



is better than a small one

but the effect saturates beyond a certain size.

Thanks

Thanks to Neville Hunt and to David Arundel, Christian von Ferber, Arnaldo Donoso and Andrew Snowdon.