

DISTRIBUTION OF SCIENTIFIC EXPERTS AS RECOGNIZED BY PEER CONSENSUS

M. KOCHEN, R. CRICKMAN, A. BLAIVAS

*Mental Health Research Institute, University of Michigan,
205 Washtenaw Place, Ann Arbor, Michigan 48109 (USA)*

(Received February 14, 1981 in revised form July 10, 1981)

Peer review plays an important role in maintaining the quality of science. Selection of peers is at the heart of the process by which science advances. Editors and others responsible for selecting a group of peers often rely on their position in a network by which experts in a field are linked to one another by bonds of common interest and recognized expertise. In this paper, we report one aspect of a study aimed at characterizing the structure of this network: the asymmetry of the fraction of experts receiving varying numbers of nominations as experts by peers. The distribution of such nominations is very skew, and we have found that a law of cumulative advantage provides the best theoretical approximation for the distribution of nominations, especially when the overall pool of data is broken down into well-defined specialties.

Introduction

Peer groups are often charged with evaluating a candidate for tenure, with evaluating a research proposal for funding, or with evaluating a manuscript for publication in a scientific journal. The composition of the group of peers is an important determinant of the decision made.¹ The process of selecting referees or reviewers has not been as widely studied as might be expected, considering its importance. The status and the theoretical stance of reviewers can influence reviewers' recommendations about publishability of a manuscript.² The degree of consensus in peer ratings of proposals submitted to the National Science Foundation is very low.³

Evaluations of the quality of graduate departments based on peer ratings do not correlate well with evaluations based on quality and quantity of published output. This may be so because those rating graduate departments take account of the wider university context.⁴ Individual recognition may also be influenced by people with whom the subject communicates and with whom he socially interacts.⁵ Generalizations are, however, difficult to support.

Consensus among investigators about quality has been suggested⁶ to be greater in fields with more comprehensive and more precise theoretical structures. The im-

portance of new ideas may be more difficult to identify in more descriptive and loosely organized fields. Even within a field, e.g. within high energy physics, phenomenologists familiar with both central theory and experiments may have greater professional recognition than those who are further from the hub of scientific communication.

The practice of selecting referees for peer evaluation tasks has been an art that varies greatly with the selector, as well as with the field. The following quote, attributed to Dr. Benton *Cooley*, illustrates one extreme attitude: "A successful surgeon should be a man who, when asked to name the three best surgeons in the world, would have difficulty deciding the other two." In other fields, self-selection is very rare. A combined theoretical and practical basis for a more systematic and scientific approach to referee-selection has been under development in recent years.⁷ When an editor chooses potential referees, he uses his connections with authors based on his recognition of their interests and expertise. These choices could be improved by utilizing these connections more fully and by improving these connections. In order to explain how this network of connections attains and maintains the structure it has, we have begun to build models to explain the dynamics by which specialties grow⁸ and develop into the present structure of the network.

That structure is not homogeneous. Some nodes in the network have many more links than others. The nodes cluster. Some clusters are denser than others. The clusters form hierarchies. Nodes with many links to start with tend to grow even more links, accentuating the asymmetry or skewness.

In this paper, we examine the skewness in which scientists are recognized as experts by their peers. We asked experts in various specialties to name other experts in their field whose work and competence they esteem. We found that most experts so named were named by just one nominator. Only a few experts were named by more than five others. Our aim in this paper is to characterize the distribution of experts according to the number of times they are nominated and to explain the nature of that distribution in the context of the peer network. This work is very similar to that done by *Cole, Cole* and *Dietrich*⁹ and can be regarded as an independent verification of their findings. They were, however, not concerned with the distribution of nominations.

Theoretical Considerations

A person develops expertise and attains recognition for it over time. If this first important contribution attracts the attention of people occupying key positions in the communication network, word may spread and his visibility may increase. Peers

may begin to cite his publication. A bandwagon effect may begin. If experts, whose judgments are respected, speak highly of the person's contribution, others may become favorably disposed to speak highly of him as well. The resulting reinforcement helps him to make further contributions of the kind that earn him the critics' favors.

This phenomenon has been called the "Matthew effect" or a "law" of cumulative advantage. It has been studied by a number of people.¹⁰ It can be modeled by an application of *Polya's* urn scheme¹¹ for contagion, which is a special case of the general effect. This likens the process of being nominated as an expert to drawing balls from an urn of black and white balls so that whenever a white ball is drawn, k white balls are added to the urn so that the probability of drawing a white ball increases; the white balls correspond to advantaged persons, and being drawn corresponds to an act of recognition. Peer recognition, such as citation, is such an act of recognition; it may be reflected in being nominated in our survey.

Yule's Distribution

Logical considerations used by *Price* to derive a (Yule) distribution for the Cumulative Advantage effect for the description of the results of the citation analysis have some similarity with the distribution of nominations. Citation of a paper may be considered as an act of recognition of a scientist somewhat similar to being nominated. The Yule distribution is described by the expression:¹²

$$\text{Prob. } (x = n) = (m + 1)B(n, m + 2), n = 1, 2, \dots$$

where $B(a,b)$ is the beta function,

$$B(a,b) = \frac{\Gamma(a)\Gamma(b)}{\Gamma(a+b)}, \quad \Gamma(a+1) = a\Gamma(a), \Gamma(0) = 1;$$

and m may be a noninteger number, x is the number of nominations. m may be estimated by the relationship: $m = 1/(\bar{x} - 1)$, where \bar{x} is the mean of the experimental distribution we would like to approximate.

We could evoke a Matthew effect by asking some people to nominate experts, publicize who nominated whom and then ask more people to nominate experts in second and further rounds. We did not do this. We asked editors to name research-

ers from whom they would most like to receive manuscripts or whom they would most like to use as referees for manuscripts they received. Then, we asked similar questions of the persons so named. If a bandwagon or Matthew effect were at work, it must precede the stage at which we are observing the choices. If only a few people are named by many others, it may be the consequence of these few having been advantaged and reached a position of advantage, presumably because their superior contributions and qualifications earned them more recognition by key people. That past recognition increased the number of others who recognized them as well.

Negative Binomial

Previous studies of the distribution of nominations in a large sociogram showed that the distribution is a negative binomial. An investigation into the social structure of nomination by friends in high schools^{1,3} led us to ask if our data might fit the theoretical negative binomial distribution. An application of the truncated negative binomial might be especially suitable, because we do not know how many persons received no nominations.

This distribution is expressed as follows:^{1,3}

$$\text{Prob. } (x = n) = (1 - Q^{-N})^{-1} \binom{N + n - 1}{N - 1} (P/Q)^n (1 - P/Q)^N$$

$$n = 1, 2, 3, \dots$$

where x = the number of times a randomly chosen person is nominated,

$$P + Q = 1, P > 0$$

The parameters P and Q may be estimated from an appropriate experimental distribution by:

$$\hat{Q} = s^2(\bar{x})^{-1} (1 - f_1)^{-1}; \hat{N} = (\bar{x} - \hat{Q}_{f_1}) (\hat{Q} - 1)^{-1}$$

where s^2 is the sample variance, f_1 is the frequency of the first observation (in our case, the frequency of only one nomination), and \bar{x} is the sample average.

Design of the Study

We selected six specialties: Topology, Differential Geometry, Human Systems Management, Information Science, General Systems Theory, and Future Studies.

Low-dimensional topology and differential geometry are well-defined, related mathematical specialties. They are highly codified and not capital-intensive.

The next four areas are not traditional scientific disciplines. They were chosen, in part, because of the first author's involvement in them. He was one of the first to use "information science" to describe this new field.¹⁴ *Human Systems Management* is the title of a new journal (North-Holland, 1979), which he helped to start and for which he serves as managing editor. He and Michael *Marien* are co-ordinators of the "Futures Information Network," an attempt to provide for the exchange of recommendations and opinions about newsworthy future studies that have been published or are in progress. General systems theory is the area of a group of researchers, many of whom are interconnected by a computer conferencing system led by Stuart Umpleby, in the Electronic Information Exchange System, in which *Kochen* participated.

Fifty-four editors of journals in these six fields of study were selected to receive identical questionnaires. We tried to identify all the key journals in each of the fields. 54 responded and made a total of 417 nominations to 336 different people. The names of the 40 respondents and the names of the 336 people they nominated were stored in two computerized data bases. These files constituted the first round of our study.

We then sent identical requests to 336 different people named by the 40 editors and called the responses received from this group the second round. The new persons named by the second round respondents (some second round respondents name one another or one of the forty editors) received questionnaires in turn. Their responses became the third round data. The data presented and analyzed is based on these first three rounds. A similar method was used by *Kadushin*¹⁵ for the American intellectual elite. The difficulty with the method is precisely the skewness or long tail of the peer recognition distribution curve. Because so many more experts are nominated once rather than several times, we could expect even the 10th round to add some singly nominated individuals.

Our first round requests for nominations of experts varied somewhat by field. We asked journal editors of *Low Dimensional Topology* and *Classical Differential Geometry* to nominate people in both those fields. We asked all the members of the editorial board of the new journal, *Human Systems Management*, for nominations in that field. In information science, we started with requests to the principal editors of 11 different journals in that field. Initial requests in future studies were

given to members of the Futures Information Network. Requests for nominations were sent over the EIES computer conferencing network to all members of the general systems theory group.

Results and Their Interpretation

The actual number of requests, the response rate, and the number of people nominated broken down into fields is shown in Table 1.

As shown in Table 1, 40 respondents made 417 nominations in the first round of the study, and 104 of these were respondents in the second round. They made 1049 nominations in the second round. Note that six of these responded and identified their specialty as polymer chemistry, even though they might have been named as experts in future studies or general systems theory. 156 respondents made 1044 nominations in the third round. In total, 300 respondents made 2564 nominations in which they listed 1460 different scientists as experts in their field.

The number of people receiving exactly n nominations broken down in fields is shown in Table 2. At the bottom of the table we provide means and standard deviations for corresponding fields and for all 7 fields combined.

Some of the data are plotted in Fig. 1 as bar graphs. We first checked if these data fit a truncated negative binomial distribution. The estimate of the parameters Q and N gives

$$\hat{Q} = 1.387 \text{ and } \hat{N} = 5.15$$

It can be seen that the truncated negative binomial distribution is not a good candidate for the approximation of our histogram.

One of the reasons that the distribution does not work in our case might be that we asked only a limited number of people to nominate others each round. Consequently, our data do not meet the eligibility for approximation by the Poisson distribution. The second reason might be that there is an obvious interdependence between choices of nominators in sequential rounds. A new statistical technique that is more appropriate for this kind of chain-letter technique to explore a network than either the negative binomial or the Yule–Greenwood distribution is being developed by Peter *Lenk* of our research team. The resulting distribution is likely to exist as a computer program rather than in closed form. When this work is completed, it will be tested against our data and reported in a separate paper.

Second, we tested the *Yule–Greenwood* distribution,¹⁶ based on the (Matthew) effect of cumulative advantage. An attempt to approximate the experimental statistical histogram for all 7 fields of nominations by the Yule distribution is shown in

Table 1
 Number of nominators, response rate, number of nominations, and number of nominees in the first three rounds of the study

Field	First Round				Second Round				Third Round				Total		
	Nominators	Response Rate	Nominations	Nominees	Nominators	Response Rate	Nominations	Nominees	Nominators	Response Rate	Nominations	Nominees	Nominators	Nominations	Nominees
Future Studies	6	55%	60	57	11	28%	111	84	11	39%	85	73	28	256	171
General Systems Theory	8	32%	70	45	9	18%	86	83	23	50%	168	130	40	324	211
Human Systems Management	15	61%	155	141	36	34%	376	293	62	31%	368	296	113	899	589
Information Science	8	73%	79	59	32	60%	316	196	49	35%	330	251	89	725	381
Differential Geometry	(2)*	30%	20	14	5	36%	51	33	2	20%	39	39	7	110	63
Topology	3	30%	33	20	5	29%	69	54	2	13%	9	9	10	111	67
Polymers	-	-	-	-	6	-	40	28	7	27%	45	39	13	85	60
Total	40	74%	417	336	104	28%	1049	771	156	24%	1044	819	300	2564**	1460

*The nominators in this subject field were themselves in subject area of topology and supplied nominees in both subject area.

**An additional 54 nominations arise from the section of journal editors who formed the seed of our nominating process.

Table 2
Peer recognition data

n	Fields:		Future Studies		General Systems Theory		Human Systems Mgmt		Inform. Science		Differ. Geometry		Topology		Polymer Chemistry		All 7 fields Combined	
	#	%	#	%	#	%	#	%	#	%	#	%	#	%	#	%	#	%
1	130	(76)	151	(72)	462	(78)	258	(68)	42	(67)	43	(64)	44	(73)	1029	(70)		
2	21	(12)	37	(18)	62	(11)	62	(16)	10	(16)	13	(19)	10	(17)	204	(14)		
3	9	(5)	10	(5)	23	(4)	18	(5)	3	(5)	4	(6)	4	(7)	82	(6)		
4	5	(3)	7	(3.3)	16	(3)	6	(1.6)	3	(5)	5	(7)	1	(1.7)	46	(3)		
5	4	(2)	2	(0.9)	9	(1.5)	12	(3)	4	(6)	2	(3)	1	(1.7)	38	(3)		
6	0		1	(0.5)	5	(0.8)	7	(1.8)	0		0		1	(1.7)	16	(1.1)		
7	0		1	(0.5)	6	(1.0)	5	(1.3)	1	(1.6)	1	(1.6)			14	(0.9)		
8	1	(0.6)	1	(0.5)	0		6	(1.6)	6	(1.6)					10	(0.7)		
9	1	(0.6)	0		0		3	(0.8)	3	(0.8)					8	(0.5)		
10			1	(0.5)	4	(0.7)	1	(0.3)	1	(0.3)					4	(0.3)		
11					1	(0.2)	0		0						3	(0.2)		
12					1	(0.2)	1	(0.3)	0						0			
13							0		0						3	(0.2)		
14							0		0						0			
15							1	(0.3)	1	(0.3)					0			
16							1	(0.3)	1	(0.3)					2	(0.13)		
17															1	(0.06)		
Totals	171		211		589		381		63		67		60		1460*			
Sample mean; \bar{x}	1.18		1.54		1.53		1.90		1.75		1.66		1.43		1.76			
Standard Deviation of x:	1.42		1.20		1.41		2.01		1.36		1.08		0.83		1.73			
Yule distribution parameter m	5.56		1.85		1.89		1.11		1.33		1.52		2.33		0.56			
Chi-square	43.91		3.27		15.46		13.74		2.22		2.33		0.43		20.59			
Degrees of freedom	2		3		4		4		2		2		1					
Strategy	Rejected		Accepted		Rejected		Rejected		Accepted		Accepted		Accepted		Rejected			
	p < 0.01		p > 0.3		p < 0.01		p < 0.01		p > 0.3		p > 0.3		p > 0.5		p < 0.01			

*Fields total more than 1460 because some experts received nominations in more than one field.

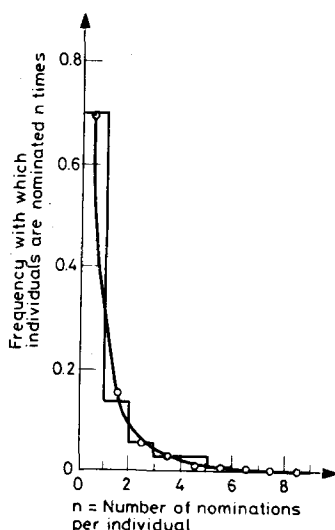


Fig. 1. The distribution of nominations

Fig. 1. The experimental distribution is represented by bars. The theoretical distribution is represented by a smooth curve drawn through the dots corresponding to sequential nominations. The curve was drawn for clarity's sake, so that the reader can better visualize the result. Visually, it fits the data on the bar graph quite well.

To test the goodness of fit, we computed Chi-square. The bins greater than 8 were aggregated into one observation. The result is a large value, $\chi^2 = 20.95$, which with 7 degrees of freedom shows the fit to be statistically insignificant.

Numerous observations may contribute to the diversity of the points. A similar effect has been observed in fitting radioactive decay by the Poisson distribution. For smaller groups of data the Poisson distribution fits well. But if the observations exceed a certain level, then a paradoxical effect occurs: the Chi-square test begins to force a rejection of the null-hypothesis.

The nomination data differ for the 7 different fields. Information scientists do not know polymer chemists and seldom nominate them. To test this hypothesis, we fit all seven fields separately by the Yule distribution and the results are shown in the same Table 2 at the bottom. The hypothesis that the Yule distribution is a fair approximation for the corresponding data sets proves true for the following specialties: Polymer Chemistry, Differential Geometry, and Topology with a statistically significant fit for these "hard" sciences at the 0.3 level. For General Sys-

tems Theory, the fit is statistically significant at the 0.5 level. Corresponding values for Chi-square and the p-values are found in the Table. The hypothesis has been rejected at the 0.01 level for the Human System Management, Information Science and Future Studies fields, demonstrating that these fields are nonhomogeneous. It is plausible that they comprise several groups or clusters, each of which has a potential to develop into a separate specialty in the future.

In order to further substantiate the hypothesis we conducted a statistical experiment: we combined three hard sciences (Diff. Geometry, Topology and Polymer Chemistry) together and, taking the distribution thus obtained, tried to approximate it using the Yule distribution. Although each of these three fields separately displays a satisfactory fit, together they give a worse result: the hypothesis has been rejected:

$$\chi = 5.295, \text{ with } m = 1.69 \text{ and three degrees of freedom;}$$

$$p\text{-value} > 0.10$$

We cannot be sure that each group analyzed is homogeneous. It is possible that each group includes representatives of several different fields. The problem of uniformity of the field bears a strong relationship to the degree of consensus among respondents concerning whom to nominate.

The degree of consensus about who had expertise was lower than expected. Only 30% of the sample received at least two nominations, and the proportion of the scientist receiving at least three nominations was only 16%.

Only in the two math fields were there scientists who were acknowledged as experts by 50% or more of the respondents. Future studies and general systems theory only had two scientists (1.0% of the total) who were nominated as experts by a quarter of the respondents. In information science, only two scientists (0.5% of the total) who were nominated by 18% of the respondents, and only one of the 589 nominees in human systems management was nominated by 11% of their group. Clearly then, the degree of consensus about the selection of scientific expertise in unstructured fields is quite low.

The data on consensual agreement in each of the six fields supports the idea that they form two groups. In human systems management, general systems theory and future studies, the proportion of scientists nominated by at least two others was 22, 29 and 24% correspondingly, and the fraction of scientists recognized by three or more peers ranged from 10–12%. In classical differential geometry, low dimensional topology and information science, the proportion of scientists receiving at least two nominations was between 32% and 36%, and the proportion with three or more nominations was between 16% and 17%.

Similarly, the proportion of "stars" with 5 or more nominations ranged from 3–5% in human systems theory management, general systems and future studies, while in the two math fields and information science, it was 5–10%. There was, however, some overlap among the field so that nominators from both information science and general systems theory might nominate the same experts. In short, the more codified fields seemed to have a higher degree of consensus in the recognition of scientific expertise.

The inclusion of information science with the mathematical field merits some comment. Information Science has a number of theoretical formulations, such as the class of skew distributions, cumulative advantage processes, Lotka's law, which are well know and mathematically expressed. They provide a coherent intellectual structure for much of the field.

In this regard, it is interesting to note that the two most codified fields in this study, classical differential geometry and low dimensional topology, seemed to have the highest degree of consensus about scientific expertise. Both fields had nine scientists (7% of the sample) who were nominated by more than 49% of their peers. In none of the other fields was there even one scientist recognized as widely.

The degree of codification in a scientific discipline may be an important component in developing consensus about scientific expertise. The nature of the peer recognition curve is a very skew, irregular function that seems to reflect an underlying law of cumulative advantage for hard sciences, such as Polymer Chemistry. Its study is of interest for the social aspects of science as much as for its possible application to science policy analysis and the improvement of peer review.

This research was made possible by support from National Science Foundation grant 1ST-78-16629. The authors were assisted by J. Hersey while he was on a post-doctoral fellowship. Cecilia Hallock, Charles Barr and L. Leamon helped greatly in preparing the data base.

References

1. M. KOCHEN, Quality Control in the Publishing Process and Theoretical Foundations for Information Retrieval, in: J. TOU (Ed.), *Software Engineering*, Vol. 2 Academic Press, New York, 1971; M. KOCHEN, B. PERKEL, Improving Referee Selection and Manuscript Evaluation, in: M. BALABAN (Ed.), *Scientific Information Transfer: The Editor's Role* Reidel, Dordrecht, Holland, 1978.
2. D. LINDSEY, *The Scientific Publication System in Social Science*. Jossey-Bass, San Francisco, 1978.
3. J. COLE, S. COLE, *Social Stratification in Science*. University of Chicago Press, Chicago, 1973; and Which Researcher Will Get the Grant? *Nature*, 279 (1979) June 14.

4. R. C. ANDERSON, F. NARIN, P. McALLISTER, Publication Ratings vs. Peer Ratings of Universities, *Journal of the American Society for Information Science*, 29 (1978) 91–103.
5. S. CRAWFORD, Informal Communication Among Scientists in Sleep Research, *Journal of the American Society for Information Science*, 22 (1971) 335–352; D. CRANE, *Invisible Colleges*. University of Chicago Press, Chicago, 1972.
6. R. K. MERTON, H. ZUCKERMAN, Age, Aging, and Age Structure in Science, in: R. K. MERTON (Ed.), *The Sociology of Science*. University of Chicago Press, Chicago, 1973, 497–560.
7. KOCHEN, op. cit., note 1.
8. M. KOCHEN, A. BLAIVAS, A Model for the Growth of Mathematical Specialties, *Scientometrics*, 3 (1981) 265.
9. COLE, COLE, DIETRICH, Measuring the Cognitive State of Scientific Disciplines, in: Yehuda ELKANA (Ed.), *Toward a Metric of Science*. John Wiley & Sons, New York, 1978, 209–252.
10. R. K. MERTON, The Matthew Effect in Science, *Science*, 159 (1968) 56–63; also *The Sociology of Science* University of Chicago, Chicago, 1973; D. de SOLLA PRICE, Cumulative Advantage Games Explained: A Reply to Kantor, *Journal of the American Society for Information Science*, 29 (1978) 204–206.
11. F. EGGENBERGER, G. POLYA, Ueber die Statistik Verketteter Vorgaenge, *Zeitschrift fuer Angewandte Mathematik und Mechanik*, 3 (1923) 279–289.
12. M. GREENWOOD, G. V. YULE, An Inquiry into the Nature and Frequency Distributions Representative of Multiple Happenings, *Journal Royal. Stat. Soc.*, 83 (1920) 255–279.
13. N. L. JOHNSON, S. I. KOTZ, Negative Binomial Distribution, Ch. 5, *Discrete Distributions* Wiley-Interscience 1970.
14. M. KOCHEN, *Some Problems in Information Science*, Scarecrow, Metuchen, 1965.
15. C. KADUSHIN, *The American Intellectual Elite*, Little and Brown, Boston, 1974.