The Long Tail of Idea Generation
by

Robin W. Spencer and Timothy J. Woods

reprinted from

International Journal of Innovation Science

The Long Tail of Idea Generation

Robin W. Spencer* and Timothy J. Woods
*Director of Research, Imaginatik PLC
207 South Street, Boston MA 02111
spencerrw@alum.mit.edu
www.scaledinnovation.com

Abstract

Engaging hundreds or thousands of employees in challenges that address specific business needs is a well established innovation model. Large-scale idea generation is a key element of such challenges, and we have observed that it has remarkably consistent statistical properties across every challenge, company, and industry. It bears a strong analogy to the "long tail" that Chris Anderson recognized in electronic retail of books and music and shares its enabling economics. We propose that idea generation is driven by universal positive feedback properties which account for its observed power law form, scale independence, and dynamics of growth. Simulation reproduces these properties and suggests further details and consequences. These robust observations have significant impact on strategies in collaborative innovation.

1. THE EVENT MODEL FOR COLLABORATIVE INNOVATION

In this paper we show that voluntary large-scale collaboration systems have a consistent statistical signature which has significant consequences that innovation leaders should appreciate. But before getting to the numeric observations it is useful to review the event model, also referred to as the challenge or diverge-converge model, which underpins most of our data. The model is centuries old, proven effective, and well illustrated by the classic Longitude Problem. As recounted by Dava Sobel [1], determining accurate longitude at sea was an unsolved and important problem in the early 18th century, and its consequences became all too apparent in October 1707 when four English warships struck rocks in a thick fog and 2000 men drowned. In response, Parliament's Longitude Act of 1714 set the terms of a great challenge: the problem was the determination of longitude at sea, to win required a specified accuracy demonstrated on a real voyage, anyone could submit an answer (you didn't even have to be English, let alone connected to the navy or scientific establishment), Parliament and the Admiralty would fund and develop the solution, there was a formal review committee and process (initially led by Sir Isaac Newton), and the prize was the unimaginable sum of £20,000. John Harrison, an unknown clockmaker with no formal education, solved detail after detail and ultimately won the challenge (and the prize) by inventing what we now recognize as the first marine chronometer.

The longitude story illustrates the elements of a good challenge: state the problem clearly, do not prejudge the solutions, say why it's important and who will implement the best ideas, expose the challenge to a large diverse audience, let everyone know the review process and criteria, and reward success (with involvement, recognition, or money).

One of us (RWS) has been developing and facilitating this process inside Pfizer Inc. for over four years and over 200 challenges with a system internally referred to as the "Idea Farm." The target audience for an Idea Farm challenge may be as small as 20 people or as large as the entire corporation, but the majority are at a scale of 200 to 2000 employees. These are also the most useful because at this scale the business needs are complex enough to benefit from wide and diverse input while the executive sponsorship is local and direct enough to lead to tight decisions and probable implementation. Here are a few examples of the types of questions that these challenges pose to employees:

> A new product has just passed a significant milestone in its expensive development and is now likely to make it to market. The team has been so tightly focused they may not be aware (or responsible) for needs elsewhere in the company. In the broadest sense, what else can we do with this product? What are unmet customer needs in unrelated areas that it might address?

Regulatory authorities require us to prepare certain lengthy complex documents. Despite being so constrained, what prevents us from cutting 30% off the time it takes to complete them?

For a sales force that often works out of homes or automobiles, what are the most time-wasting or distracting aspects of the job and what are feasible solutions (given various constraints of budget or job requirements)?

The IT department can take on a limited number of new projects each year. Targeting their entire internal audience, what are new needs for IT capabilities this year? Rate them by complexity, consequences, scope, and urgency against preexisting commitments.

As a side effect of the research needed to produce our core products (new pharmaceuticals), additional intellectual property may be generated. What are all of these intellectual property estates which may become opportunities for spin-off businesses?

On an outside website we welcome new product ideas and proposals from the public at large.

These examples show that our challenges do not pigeonhole innovation: business needs around new products, saving time or money, allocating resources, and more are all in scope. If the results are new and useful, that is by definition innovative.

2. "NOT EVEN WRONG"

Our investigation into the mathematical structure of innovation challenges begins with a simple question that could easily come from an executive sponsor: "What's the average number of ideas entered per person?" A little exploration of this innocent and apparently reasonable question reveals that any answer we could give – for example by dividing the total number of ideas by total participants – would be "not even wrong." This nice phrase comes from physicists, who say that an incorrect measurement based on a good theory might be wrong, but a measurement based on a bad theory would be "not even wrong". It's an insult and a warning to recognize that results based on bad theory should not be calculated or presented, because doing so propagates misunderstanding and implicitly endorses the bad theory.

Why this is a case of bad theory is illustrated in Figure 1. If we ask "what's the average height of our employees?" we expect a result like that on the left: a bell-shaped curve, an obvious peak (average), and some reasonable width (variance). From the shortest to the tallest person, height varies by less than a factor of two. The right of Figure 1, presented in the same way, is from the Pfizer Idea Farm over thousands of people and thousands of ideas. There is no symmetric peak, no obvious average, and the range is enormous: many people put in just one idea and a few put in hundreds each.
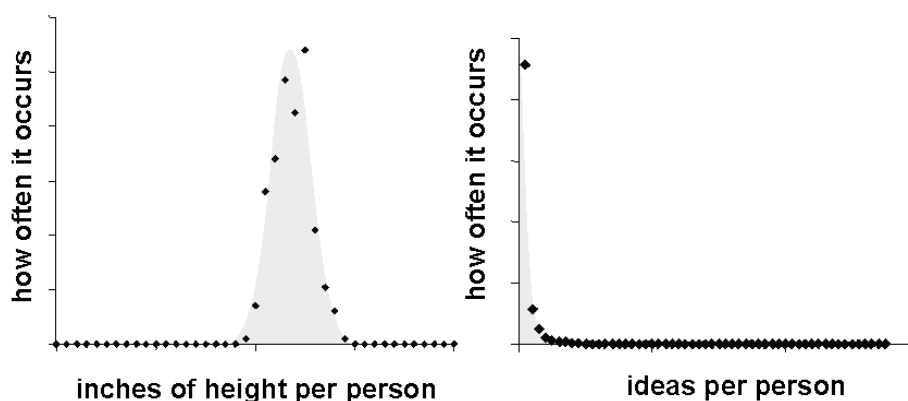

Figure 1. Height vs frequency for several hundred people. (right) Number of ideas per person vs frequency from the Pfizer Idea Farm, 2006-2009. For both graphs the scales are linear with (0,0) at the lower left.

A traditional statistics textbook would say that the height data are "well behaved" with a nice peak, a little scatter, and an easily calculated mean and standard deviation. The data would be well described by a Gaussian or Normal distribution. The idea-generation data would be dismissed as "poorly

behaved" with no peak (except at the y-axis asymptote) and a calculated standard deviation much larger than the calculated mean.

But this is science, and when large, accurately collected data sets disagree with theory, it's the theory that gets discarded, not the data. The theory we're discarding is the tacit assumption that the world should be Gaussian, and we will need to find an alternative way to describe and think about data like this. Nassim Nicolas Taleb and Paul Ormerod, authors of "The Black Swan" and "Why Most Things Fail", would strongly agree [2,3]. We are trained to think in Gaussian terms, and while that's sometimes an appropriate way to describe the world, it can sometimes be very wrong. The consequences can be huge: as Ormerod points out, financial strategies are too often based on an implicit assumption that events are Gaussian, but in the real world that may be very far from Gaussian, this can result in underestimated risk, financial fragility, and economic collapse.
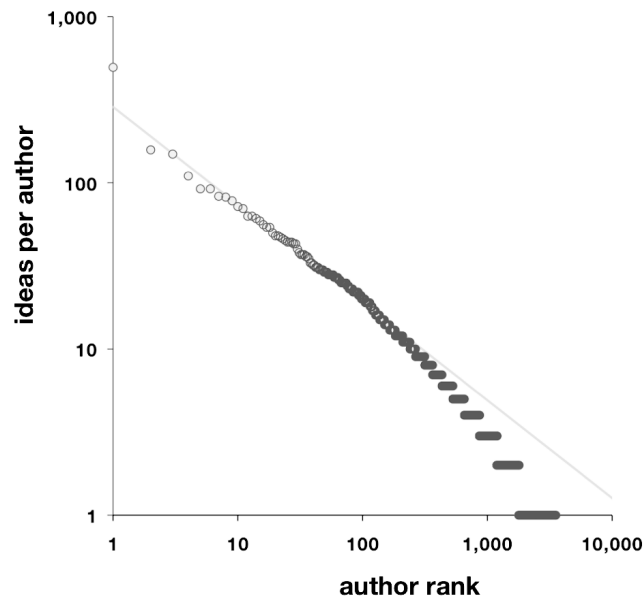

Figure 2. Rank-frequency plot for ideas submitted to the Idea Farm at Pfizer, 2006-2009, from over 200 challenges, 3530 total authors, 13784 total ideas. The line is a power law with α = 2.7. In this paper such data are graphed as frequency (y) vs rank (x) because this is readily interpreted. However the mathematics literature typically graphs such data as probability distributions with reversed axes and exponent α (alpha). Our graphs have exponent $r = 1/(\alpha-1)$

Redrawing the data in Figure 1b with author rank on the horizontal axis (where the rank of the most prolific author is 1, and the least prolific is 3180, the total number of authors in this case) and number of entries per author on the vertical axis, with both axes logarithmic, we get Figure 2. The log scales reveal details that linear scales disguise: now we can see that author #1 entered about 600 ideas and comments, #10 entered about 100, #100 entered about 20, and author #1000 put in about three. Since the data points in Figure 2 lie quite close to a straight line on a log-log plot they would be well described by a power law (i.e. $y=Cx^{-r}$); the mathematics is useful because it suggests reasons why the data take this form, but first a review of the qualitative implications.

3. THE LONG TAIL IN RETAIL

Chris Anderson coined the term "long tail" to describe similar observations in retail sales [4]. Anderson usually draws his diagrams to look like Figure 3, which is simply a graph of y = 1/x plotted with the x-axis logarithmic and the y-axis linear. Anderson shades the graph as shown to make his distinction between the limits of traditional physical retailing (the head) and the vastly longer tail accessible with the reach of the internet. The first lesson from Anderson is why the distinction exists: the head is defined by the items that can profitably be sold in traditional stores, and at some point a book or song is not popular enough to recoup the finite costs of putting it on the shelves. But electronic retailing (think Amazon or iTunes) changes the economics by vastly reducing the cost of making an additional book or song available; this exposes the tail, an enormous latent market for additional content. The second lesson is simply that the tail matters: even in 2005-2006 for Amazon (books), Netflix (movies),

and Rhapsody (music), the long tail (defined as content not available except from these e-retailers) accounted for 20-40% of their sales and was growing rapidly [4, pp 19-26]. The head consists of selling hundreds of books that sell thousands of copies each, the tail consists of selling millions of books that sell a few copies each. Smart retailers do both.
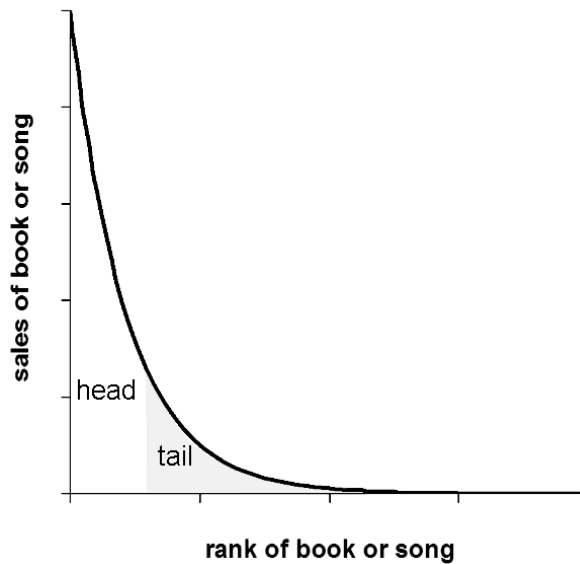

Figure 3. Rank-frequency plot of retail sales as in Anderson [4]

4. THE LONG TAIL IN IDEA GENERATION

Figure 2 shows that a large aggregated data set of participation in a collaborative innovation system appears to follow a power law. In this section we dissect these data to look at individual challenges, data from different companies and industries, data from a different business process, and examine the kinetics of growth of a challenge.
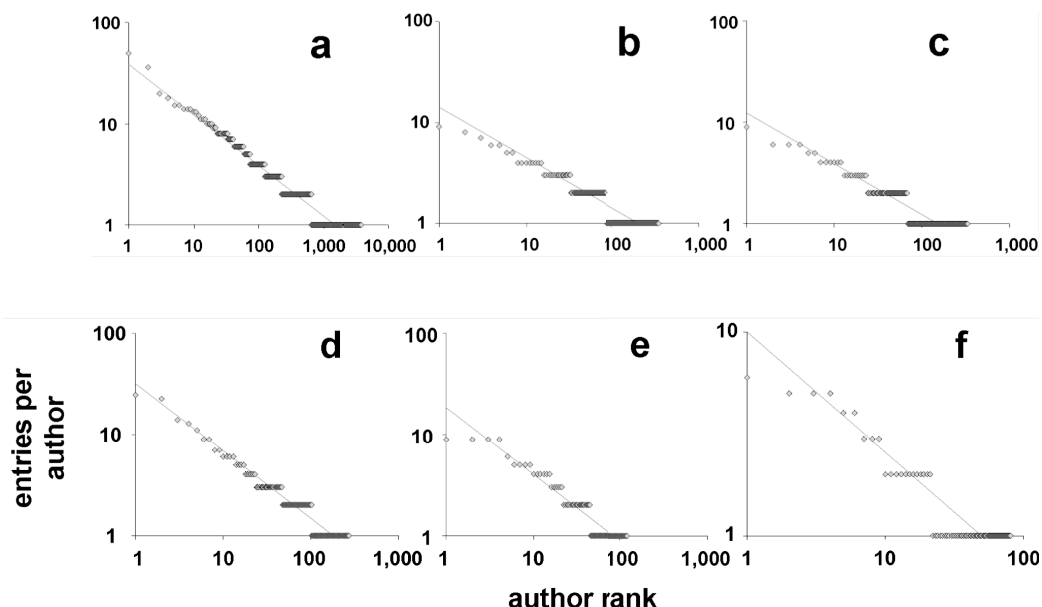

Figure 4. Rank-frequency plots for separate challenges at Pfizer. (a) unsolicited external ideas, (b) cost savings for a sales division, (c) new sources of revenue in the research division, (d) encouraging employee relocation following site closure, (e) support for a Six Sigma time-saving project in Clinical Development, (f) ideas for additional indications for a new drug in late-stage development. In all cases the x-axis is author rank and y-axis is entries per author. The lines are power laws with α = 3.0, 3.0, 3.0, 2.5, 2.5, 2.7 for (a)-(e) respectively.

Figure 4 demonstrates that challenges show the same statistical behavior individually that they do in the aggregate (Figure 2). All the details of these challenges were different: different problem to address, different business sponsor and goal, different audience, different type of knowledge or level of experience needed, different levels of participation among those invited (about 20% for Figure 4b vs 2% for Figure 4f). All have the shape of a power law, all have about the same slope ($\alpha = 2.7 \pm 0.3$). In fact this is the general case: of all the data sets that we have examined only a few fail to have this pattern and they share a common characteristic (see "Exceptions" below). Data that do not change their distribution when aggregated or dissected are "scale free", and since the power law is the only distribution with this property [5], Figures 2 and 4 are strong evidence for a power-law generating principle underlying our innovation challenges. Anderson sees the same thing when he dives into the "tails within tails" of genres of music downloaded from Rhapsody ([4] pp 139-140).

Does this long tail tell us something about Pfizer, or about collaborative problem solving in the pharmaceutical industry? Yes, but it is much broader than that. Several equally large and diverse data sets generously offered by other clients (of TJW) show the familiar pattern (Figure 5).
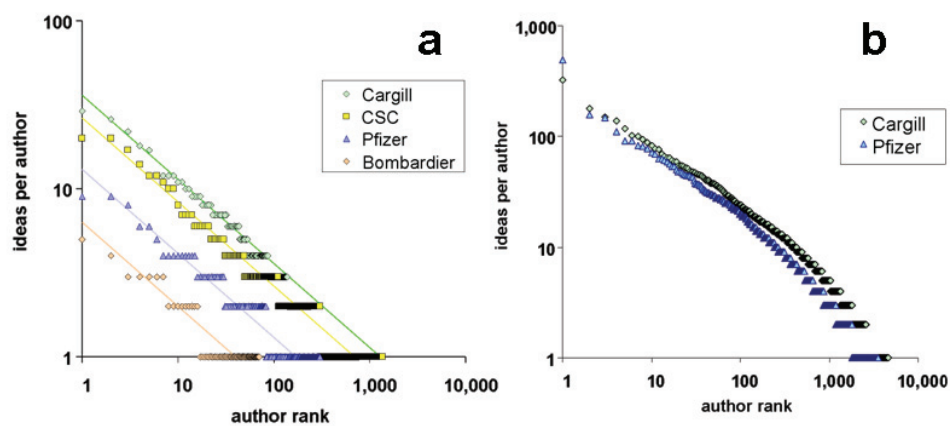

Figure 5. Rank-frequency plots for challenges at different companies. (a) individual challenges at (top to bottom) Cargill, Computer Sciences Corporation, Pfizer, and Bombardier, all lines have $\alpha = 3$, (b) aggregated data sets from Cargill and Pfizer over 3-5 years each in large business units of thousands of employees.

These companies represent agribusiness, pharmaceuticals, computer services, and aerospace manufacturing. The employees are all different and separated by secure corporate firewalls. We see the same patterns from others (who prefer to remain anonymous) in banking, insurance, and heavy manufacturing. There are no exceptions we have seen (barring the special cases noted below).

Nearly all of the data we have seen derive from the general challenge-based process and enabling software offered by Imaginatik (which has Cargill, Pfizer, CSC, Bombardier, etc. as clients). But this is not a special requirement either: the data of Figure 4a are derived from a public email "request for ideas" system run by Pfizer for a number of years, and it shows the identical pattern and slope. Indeed, with differences only in the exponent of the power law, this pattern is seen at enormous scale in open participation forums like Digg and Wikipedia [6].
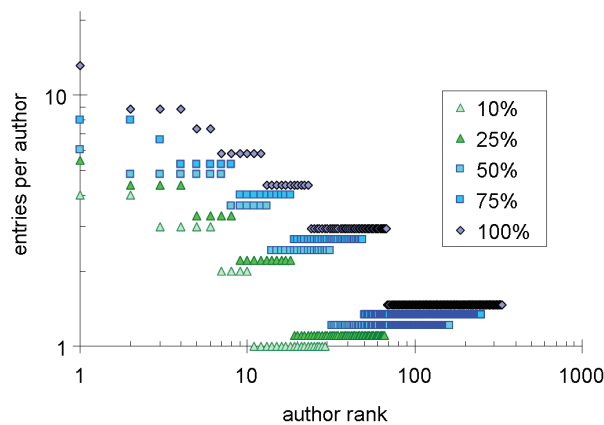
Figure 6. Growth in time of the challenge of Figure 4c. Slices of the data are shown when 10%, 25%, 50%, 75%, and 100% of the total entries has been received. Each subset is vertically offset by a constant to allow all the points to be seen.

Finally, Figure 6 shows that the shape of the data is maintained as ideas arrive over time, which holds for every dataset we have examined. Altogether (invariance over aggregation, time, company, industry, and methodology details) this is compelling evidence that there is a root cause or generating function that gives rise to power law distributions in idea generation.

5. OTHER POWER LAWS

Despite their curious omission in most statistics courses, many phenomena demonstrate power law behavior and have robust theories behind them. We can learn from these examples and develop a credible model to account for the shape of our data, and then apply that model to suggest future actions which could make collaborative innovation even richer and more reliable.

For example, graphs of the size vs. the rank (largest to smallest) of landslides follow power laws [7] as do forest fires [8], human disease epidemics [9], internet website connections [10], and collaboration networks of movie stars [10]. Note the exceptional diversity of these examples: from nature and human behavior, timeless and modern. When such things share a common generating property, that property must be incredibly simple and powerful.

The simple and powerful generating function behind these phenomena is the Yule process [11] or simply "the rich get richer" [5]. It's more likely for rocks on a hillside to roll downhill if the rocks around them are rolling (they bump into each other). It's more likely for trees to catch fire if the trees around them are already on fire. It's more likely that you'll get sick if the people around you are sick, that you'll get asked to be in a new movie if you're already well known, or that you'll visit a website if it's already highly connected to other websites. It may help to think in the converse: if the rocks on a hillside are quiet, it's unlikely that any particular rock will start rolling downhill, if the forest is quiet, it's unlikely that a tree will just catch fire, and if you've just launched your website or have just arrived in Hollywood, it's not very likely you'll get a lot of connections tomorrow. Things may begin for non-power law reasons (lightning strikes the forest, a stranger with a disease visits your city, the famous star breaks her ankle and the ingénue gets the part), but once initiated the Yule process dominates and a power law results. If events are independent (like tossing a coin), classical Gaussian bell-curve statistics result, but if events have some positive dependence on what already exists, then power laws result. The mathematics is well developed by Newman [5].

6. RESULTS FROM SIMULATIONS

We use two models to simulate a "rich get richer" process, in which the audience for a challenge is represented by a grid of cells. In the first model, at each time point the probability that a cell's value is incremented ("this person submits an idea") is

$$p = p_0(1 + bA_{self} + cA_{group})$$

where p_0 = background probability (a constant), A_{self} = the current value of this cell, A_{group} = the average value of a neighborhood of cells, and b and c are empirically estimated constants. This reproduces the observed properties of our datasets (scale-free growth and comparable exponent) only when b and c are of comparable value. If either is zero the exponents are significantly different from what we observe, if both are zero the distribution is Gaussian rather than power law. Our second model is a similar joining model from Bagrow et al. [12], in which the probability that a cell's value is incremented is,

$$p = r(\text{rich-get-richer}) + (1-r)(\text{new-joiner})$$

where r is a constant which splits between two processes: in the rich-get-richer process only cells with positive values are incremented and in proportion to their current value, in the new-joiner process a new cell is chosen at random and its value incremented from zero to one. This model is proven to generate power laws [12], and matches our results when r = 0.5-0.67 (which corresponds to α =3.0-2.5 since α = 1+1/r).

Since our results are well modeled by a single parameter (the ratio of b to c in the first model, or r in the second), it's worth exploring factors which may affect the value of that parameter. Several options are shown in Table 1, exemplified by three online cooperative systems which span the range of observed power law exponents: Twitter [13], Wikipedia [6], and our data for business challenges. In fact our challenges are at the extreme of observed cooperative power law behavior; we know of no examples with lower r (or higher α), consistent with theory which says that r=0.5 (α=3) is a transition point beyond which power laws collapse to Gaussian distributions [12]. The other theoretical bound at r=1 (α=2), e.g. Twitter, is also well established [5].

Table 1: Ranges of parameters for collaboration systems and possible driving factors

	examples		
	Twitter [13]	**Wikipedia edits [6]**	**business challenges [this work]**
slopes of power laws	r=0.96, α=2.04	r=0.78, α=2.24	r=0.5-0.67, α=2.5-3.0
driving factors			
incentives	incentive to stay >> incentive to join		incentive to stay ~ incentive to join
duration	always available		available for a limited time
difficulty	easy to contribute		hard to contribute
specificity	general topics, lots to say		specific topics, limited individual knowledge

The interpretation of incentive to stay vs. incentive to join follows directly from our second model [12]. We suggest that duration may important because the rich-get-richer contribution is more time-sensitive than the new-joiner. Wilkinson [6] derives power laws by assuming a quitting model (vs Bagrow's joining model), and interprets the governing parameter to represent the difficulty of participation. We also suggest that specificity is a better interpretation than difficulty: no matter how strong the incentives to keep contributing, in a business challenge with very specific content few individuals will have more than a handful of relevant contributions. Also all of the above systems are

technically easy to master; it cannot be the web interface which results in such consistent differences, but rather the content.

Our models and the literature offer little guidance on which of the driving forces above are more correct, or which others may be missing. All, alone or in combination, can be relevant. We offer them as aids to understanding and interpreting observed systems, and in planning experiments or strategy to change outcomes.

7. NO NETWORKS NEEDED

Note that no "community" or "network" need be involved: as long as there is positive feedback from the group ($c>0$ and $c \sim b$, first model) whether that group consists of just a handful of nearest-neighbor cells or the entire population, or its equivalent, a substantial level of new-joiners ($1-r \sim 0.33-0.5$, second model), a good fit to observed data is seen. The data of Figure 4a corroborates this result: for confidentiality reasons this public-facing idea solicitation site has been run as a blinded "drop box" such that contributors cannot see any content or any statistics from other previous contributors. They get no feedback on others' identities or behavior. This is an intriguing result because a considerable number of power-law processes are claimed to rely on network feedback [10]. It may be that network effects could augment idea generation ("I am more likely to contribute if I see my buddies or team members do so") but we have shown that it is not essential: only positive feedback from oneself ("I put in an idea before, so I'm more likely to again.") and a general background new-joiner rule are necessary. Interestingly, network theory may provide support for its own irrelevance: Ohtsuki et al [14] show that all network topologies give the same results for feedback-driven cooperation, which may be telling us that the network is moot and a very simple Bagrow-type model [12] suffices.

8. EXCEPTIONS

As shown in Figure 7, there can be strong exceptions to power law behavior. In every such case, however, the data comes from a forced-participation process rather than a voluntary one. Figure 7 is from a senior management meeting in which approximately 100 executives were assembled in a conference room, and following some introductory presentations were told to open their laptops, connect to the appropriately configured software, and "in the next 15 minutes, put in at least 5 ideas for what you can do in your business line to support the strategy you heard presented this morning." Such events are highly effective, generating much more content of higher quality and diversity than traditional Osborne brainstorming [15,16]. They are also highly competitive and far from voluntary, so it is no surprise that the rank-frequency curve becomes anchored on the x-axis near the total possible audience size, and grows upwards from there. Over time we could expect this process to approach a Gaussian distribution. Our simulations show the same curve-over phenomenon as the percent participation increases above 30-40%; in other words the power law phenomenon is well behaved only for participation lower than this (which is the general case for voluntary web- or email-based processes of any type [17]).
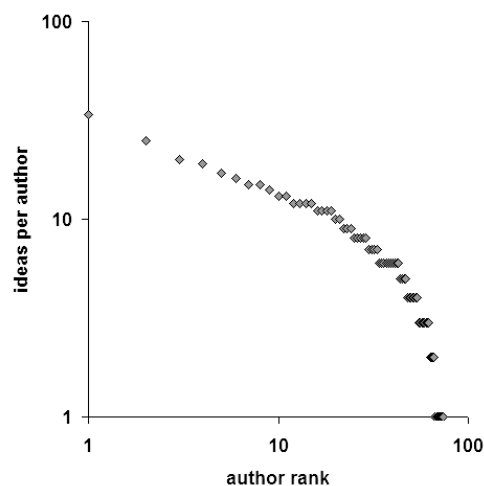

Figure 7. Rank-frequency plot from a forced-participation event.

The aggregate data of Figure 2 contain a number of such forced events which may account for its downward trend at high author rank.

9. CONSEQUENCES FOR YOUR INNOVATION STRATEGY

We have seen that voluntary idea generation shows power-law behavior at all scales and across diverse use-cases, companies, and industries. Simple simulations based on a positive feedback model reproduce its behavior. Why does this matter?

9.1 Where the Ideas Are

The six challenges in Figure 4 are a reasonably representative set. Across these challenges, on average, 80% of the participants put in just one contribution, which accounted for 57% of the total entries. Figure 8 shows this graphically: people who put in just 1 or 2 contributions accounted for 75% of the total.
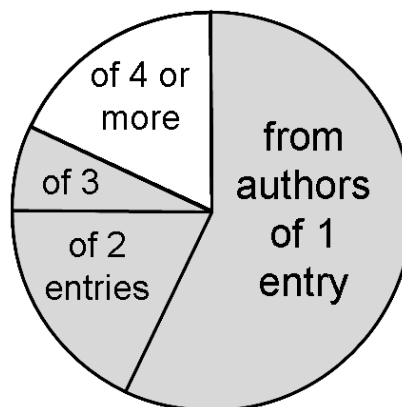

Figure 8. Where do ideas come from? Average of the six campaigns of Figure 4, showing that most entries come from people who just put in 1 or 2.

This is the principal take-home lesson of the long tail of collaborative innovation, namely that most of the content comes not from a highly active minority, but rather from the majority of individuals who contribute only rarely, once or twice per event. To the extent that we recognize and reward individual quantity (with prizes, bonus points, newsletter kudos and so forth), and thereby ignore or even denigrate the rare contributor, we risk losing a great deal of the potential in our systems. Our innovation strategy should focus on increasing the number and diversity of people that participate instead of pumping up the numbers of those who are already frequent players. We should recognize quality, not quantity, even if that is more work and takes longer.

10. MAXIMIZING VALUE, MINIMIZING NOISE

We believe that idea generation is driven by a Yule process, because with the great range of scales and businesses that make up our data we need something of its simplicity and power to be a common generating principle. This is more than academic: the positive feedback of a Yule process is something that we can use to advantage in managing our challenge campaigns.

Positive feedback processes cannot start all by themselves. If no careless camper or lightning ignites the first trees, a forest fire will never happen. If no earthquake or heavy snowfall sets things in motion, avalanches are very unlikely. By analogy we can make our campaigns start more quickly and with greater content by "priming the pump" (a different involuntary behavior), which at a minimum means seeding them with a handful of realistic ideas entered by the sponsors or facilitators. It is also highly effective to kick off a large voluntary campaign with a small, driven, face-to-face session similar to the management meeting described above under Exceptions ("I want 3 ideas from each of you in the next 10 minutes."). Subsequent voluntary contributors see peers' content the first time they enter the challenge, giving them both actual examples and social permission to contribute. There is little more intimidating than a blank page.

The feedback mechanism of idea generation can also be used to help manage negative content. Many challenge sponsors are concerned that their event could get sidetracked by irrelevant or even offensive open commentary. By analogy to the forest fire, the most effective time to quench a new blaze

is when it's very small: a quick bucket of water does the trick while an army of firefighters will struggle with a major blaze. Similarly, an attentive facilitator can stop irrelevant topics with a single calm but stern public comment. It is the public visibility of this cold water that makes it incredibly effective. There are many other best practices to mitigate this type of risk in a positive way which are beyond the scope of this article.

10.1 Lesson from Retail: Business as Usual is Not Good Enough

In retail, the "head" is defined as the number of items (books or songs or movies) that a traditional physical store would carry. The head stops when the next-most-popular book is not likely to sell enough to recoup the cost of offering it. The equivalent in business idea generation is calling meetings, creating a team, or launching a task force: these are proven approaches to generating new ideas and building out options. But just like shelving a book, calling a meeting or forming a physical team has real costs in scheduling, logistics, travel, and disruption of other work. There is also a strong social limit to effective team size with optima cited from 4 to 12 team members [18].

Imagine that we replaced the open electronic challenges of Figure 4 with teams limited to 10 people each. Even if we were able to include only the most prolific people, we would have left 92% of the ideas "down the tail," invisible to us. And yet by approaching business needs with small teams this is exactly what happens. We are not asserting that electronic media are superior to face-to-face small teams in all tasks; certainly focused teams are proven and effective in their motivation, 100% participation, and follow-through. But in this internet age where electronic challenge forums are readily available and easily implemented, to be restricted to tackling business needs only with small physical teams is just as myopic, risky, and self-limiting as assuming that your friendly bookshop or music store can ignore Amazon and iTunes forever.

11. ONGOING RESEARCH

We remain committed to ongoing research aimed at understanding and improving the efficiency and business benefit that can be obtained from scalable collaboration methods. Our interests include (i) developing a value metric for ideas, (ii) investigating the scaling properties not just of ideas but on page hits, written comments, and various voting and ranking techniques, and (iii) developing scalable, fair evaluation and decision strategies.

ACKNOWLEDGEMENTS

REFERENCES

[1]   Sobel D., *Longitude,* Fourth Estate, London, 1995

[2]   Taleb N., *The Black Swan,* Random House, New York, 2007.

[3]   Ormerod P., *Why Most Things Fail,* Pantheon, New York, 2005, 173-179.

[4]   Anderson C., *The Long Tail*, Hyperion, New York, 2006

[5]   Newman M.E.J., *Contemporary Physics* 46, 323-351, 2005, arXiv:cond-mat/0412004v3.

[6]   Wilkinson D., *Strong regularities in online peer production*, in *Proceedings of the 2008 ACM Conference on E-Commerce*, Chicago, IL, July 2008, at www.hpl.hp.com/research/scl/papers/regularities/

[7]   Guzzetti F., Malamu B., Turcott D., Reichenbach P., *Earth and Planetary Science Letters* 195, 169-183, 2002

[8]   Drossel B., Schwabl F., Phys. Rev. Lett. 69, 1629–1632, 1992, cited in [5]

[9]   Rhodes C., Anderson R., *Nature* 381, 600-602, 1996

[10] Barabasi A.L., Albert R., *Science* 286:509-512, 1999

[11] Yule G., Philos. Trans. R. Soc. London B 213, 21–87, 1925, cited in [5].

[12] Bagrow J., Sun J., ben-Avraham D., *J. Phys. A: Math. Theor.* 41, 185001, 2008, arXiv:0712.2220v2

[13] Heil W., Piskorski M., Harvard Business School, at http://blogs.harvardbusiness.org/cs/2009/06/new_twitter_research_men_follo.html# .The authors quote 26.7 average tweets per Twitter contributor, so from Newman [5] equation 12 we can estimate α=2.04.

[14] Ohtsuki H., Hauert C., Lieberman E., Nowak M., Nature 441, 502-505, 2006

[15] Spencer R., Innovation by the Side Door, in *Research-Technology Management*, 50: 5, pp 10-12, 2007

[16] Johansson F., *The Medici Effect*, *Harvard Business School Press*, 2004

[17] Nielsen J., *Participation Inequality*, at http://useit.com/alertbox/participation_inequality.html

[18] Useem J., *How to Build a Great Team*, *Fortune*, June 1, 2006, at http://money.cnn.com/2006/05/31/magazines/fortune/intro_greatteams_fortune_061206/index.htm

[19] The preferential attachment simulator is open-source and available at http://scaledinnovation.com#simulator