



The wisdom of the crowd

Over one hundred years ago, Francis Galton reported an interesting observation about the trustworthiness of popular judgments [1]—the *vox populi* (i.e., the average of almost 800 guesses) in a competition to guess the weight of an ox on display was within 1% of the actual weight. Admittedly, as was subsequently pointed out by Perry-Coste [2], the participants in the competition were “butchers, farmers, or men otherwise occupied with cattle”. Hence, the guesses must have been informed by a great deal of practical expertise.

Nevertheless, the idea that large numbers of opinions, indifferently informed or uninformed, will, if sufficiently numerous, converge on accuracy has recently enjoyed something of a resurgence [3]. The purpose of this editorial is to suggest that the idea is quite wrong.

In science, as we know from Galileo, “the authority of thousands of opinions is not worth as much as one tiny spark of reason in an individual man”. Large numbers give no guarantee that we are converging on the truth. But there is a great emphasis today on achieving mass consensus. One notices how persistently it is asserted, by some, that there is a “consensus” accepting the phenomenon of anthropogenic global warming, or the overwhelming net benefit of genetically modified crops. The authority of consensus seems to be an implicit part of “post-normal science”. And the larger the numbers joining in the consensus, the greater its purported authority.¹

Traditionally, for example as practised by the seven-person Swiss Federal Council (Bundesrat; Conseil fédéral; Consiglio federale), consensus is reached by painstaking discussions in which all points of difference are hammered out and modified until there is mutual accord. One notices a similar *modus operandi* in the committees of the International Organization for Standardization (ISO). In contrast, “democratic consensus” is a very crude instrument—one simply counts the votes for each alternative, and the largest number wins. In closely fought matters, the majority might be very slight—the minority is still, by convention, forced to accept the view of the majority. In democratic counting, however, how one actually does it plays a very significant rôle in

determining the outcome. For example, in many countries voting for Parliamentary representatives is organized into constituencies. The political party with the greatest number of representatives often does not have the greatest popular vote aggregated over the entire country [4].

The online encyclopaedia Wikipedia aims at achieving consensus for each of its articles. The concept is that someone creates an article, which is then open to anyone to edit and “improve” it. All contributors are anonymous. In contrast, the editors of traditional encyclopedias sought the world’s greatest expert in the field and asked her or him to contribute an article, bearing the author’s signature. Manifestly, the concept has not been successful [5,6]. Many Wikipedia articles are very misleading,² and the problem seems to be getting worse.³

In 1894 Walter Weldon threw 12 dice 26,306 times and recorded the outcomes [7, 8]. The agreement with the binomial distribution is extraordinary poor [9], even though the system seems to be very simple and exhaustively defined.

In principle, for any random additive process, the greater the number combined to calculate the arithmetic mean, the more accurate the result. Galton’s observation appeared to corroborate this “common sense”. If Weldon’s dice were biased, one could “explain” his observations, although there was no independent corroboration of such bias.

In a random multiplicative process, the ratio between the average value of the product $\langle P \rangle$ and the most probable value P_{mp} diverges exponentially with the number of variables N as N goes to infinity [10]. The essence of knowledge is more likely to be multiplicative than additive, hence this is a more appropriate model, which guarantees the failure of the crowd-sourcing or populist approach.

In any “real”, complex system [11], typically there is no guarantee that the more data one gathers about the system and analyses, the better one will be able to understand it.⁴ Usually the opposite is true: the more data one has, the more incomprehensible the system. Hence the traditionally overweening desire of *Homo administratus* to drastically simplify the system being administered often leads to very poor results [12].

¹ One is reminded of the Irish applewoman’s principle in economics—you can afford to sell each apple at a loss if you only sell enough.

² There appears to be no actual evidence of the efficacy of the Wikipedia model. The null hypothesis, that it is no better than having a single expert, has not been disproven.

³ There is no need to lament the decline of Wikipedia. Search engines have now become so good, alongside increasing adeptness at using them (hopefully), that it is possible even for a generalist to quickly gather representative information about a topic from the Internet, and likely of better quality than what will be found in a Wikipedia article.

⁴ P.M. Allen, personal communication.

Hence, we have a sort of paradox—the mechanical gathering and averaging of data does not guarantee its convergence to “the truth”, even when it is a simple fact.⁵ On the other hand, drastic simplification of reality, which is mostly complex,⁶ leads to error, as expected [11] and as found empirically [12]. The underlying reason is that most of reality is the result of evolutionary processes. John Kay has constructed an amusing sequel to Galton’s observation, describing how the story might have evolved [13].

Given the enduring presence of the concept of evolution, Darwinian natural selection and so forth, to which not only scientists but most laymen would subscribe nowadays, it is surprising what little practical appreciation there is of its implications. Apart from the above, which vitiate most models of the “system dynamics” type, we also have phenomena such as the rise of microbial resistance to antibiotics and the resistance of weeds to herbicides, which are sometimes treated as unexpected phenomena that can somehow be overcome.

But, of course, life would be very boring if it were not so inherently unpredictable.

J.J. RAMSDEN

REFERENCES

1. F. Galton, *Vox populi*. *Nature* (Lond.) **75** (1907) 450–451.
2. F.H. Perry-Coste, The ballot-box. *Nature* (Lond.) **75** (1907) 509.
3. J. Surowiecki, *The Wisdom of Crowds*. New York: Anchor Books (2005). The book ends with the strong affirmation that “the crowd’s judgment is going to give us the best chance of making the right decision”.
4. The particular result on any occasion may depend on exactly how the country is divided up into constituencies. For an account of the paradoxes of democratic voting, see P. Hoffman, *Archimedes’ Revenge*, Part IV (pp. 223–271) New York: W.W. Norton (1988).
5. J.J. Ramsden, The future of Wikipedia. *Nanotechnol. Perceptions* **11** (2015) 131–135.
6. D. Cross, Wither Wikipedia? *Nanotechnol. Perceptions* **12** (2016) 50–52.
7. K. Pearson, X. On the criterion that a given system of deviations from the probable in the case of a correlated system of variables is such that it can be reasonably supposed to have arisen from random sampling. *Phil. Mag.* (Ser. 5) **50** (1900) 157–175.
8. Z. Labby, Weldon’s dice, automated. *Chance* **22** (2009) 6–13.
9. W. Feller, *Introduction to Probability Theory and its Applications*, vol. 1, 3rd edn, pp. 148–149. New York: Wiley (1968).
10. S. Redner, Random multiplicative processes. *Am. J. Phys.* **58** (1990) 267–273.
11. P.M. Allen and M. Strathern, Complexity, stability and crises. In: *Complexity and Security* (eds J.J. Ramsden and P.J. Kervalishvili), pp. 71–92. Amsterdam: IOS Press (2008).
12. J.J. Ramsden, Administrative failure. *Nanotechnol. Perceptions* **13** (2017) 151–160.
13. J. Kay, The parable of the ox. *Financial Times* (24 July 2012).

⁵ A further example—the experimental determination of the (bio)physicochemical parameters of a system by undergraduates given the task as a practical exercise will not usually be reliable, even though the experiment may be repeated dozens or even hundreds of times.

⁶ Allen has called complexity the “ontology of connexions”.⁴