



Minerva Access is the Institutional Repository of The University of Melbourne

Author/s:

Ponsonby, AL;Mattingly, K

Title:

Evaluating New Ways of Working Collectively in Science with a Focus on Crowdsourcing

Date:

2015-07-01

Citation:

Ponsonby, A. L. & Mattingly, K. (2015). Evaluating New Ways of Working Collectively in Science with a Focus on Crowdsourcing. *Ebiomedicine*, 2 (7), pp.627-628. <https://doi.org/10.1016/j.ebiom.2015.06.005>.

Persistent Link:

<https://hdl.handle.net/11343/260915>

License:

[CC BY-NC-ND](#)



Commentary

Evaluating New Ways of Working Collectively in Science with a Focus on Crowdsourcing

Anne-Louise Ponsonby^{a,*}, Karl Mattingly^b

^a Environmental and Genetic Epidemiology Research Group, Population Health, Murdoch Childrens Research Institute, Royal Childrens Hospital, University of Melbourne, Flemington Road, Parkville, Victoria 3052, Australia

^b slowCapital Pty Ltd, Melbourne, Australia

Collective intelligence is intelligence arising from multiple individuals, working either independently or not. Crowdsourcing is the form of group aggregate work characterised by the aggregate of the individual decisions that are made independently. A key feature is that no member to member communication exists. For humans, the independent aggregation of multiple opinions through simple averaging, majority rules or market based algorithms leads to a marked improvement in decision accuracy. This was first recognised by Francis Galton who analysed the opinions of 787 people about the weight of an ox and found combining their numerical estimates resulted in a median estimate that was remarkably near the true weight of the ox (Surowiecki, 2004). The key feature of this work is not only aggregate decision accuracy but also that a minority of individuals such as 2% or 4% can perform better than the group average (Surowiecki, 2004). There is a greater opportunity for diversity and member errors that are uncorrelated with other members' individual errors. Another advantage is that the collective group size can be very large maximising cognitive diversity, a key element in enhancing group performance (Page, 2007). Even as the average amount of expertise decreases when the crowd grows, it may more than make up for it with increased diversity (Page, 2007).

Here, Candido-do-Reis et al. report an evaluation of the test performance of the post aggregation performance of 98,293 citizen scientists who scored over 180,000 sub-images, pathology images for cancer cells, identification and oestrogen receptor status, by comparing their performance against trained pathologists (Candido-dos-Reis et al., 2015). The investigation found that very similar results were obtained for both citizen scientists and the trained pathologists. The study was well conducted although the actual task differed a little between the citizen scientists and the pathologists in that the colours were transformed for the former (Candido-dos-Reis et al., 2015). It was noted that the citizen scientist performance was not improved by weighting for user performance score (Candido-dos-Reis et al., 2015). However this has not been the case in other crowdsourcing studies. For example, a study examining the use of probabilistic coherence weighting to aggregate judgements of multiple forecasters was able to show improvements of up to 30% over the established benchmark of a simple

equal weighted averaging of forecasts (Karvetski et al., 2013). In the paper by Candido-do-Reis et al., the human only crowdsourcing post-aggregation performance was superior to the machine-learning model (Candido-dos-Reis et al., 2015) and this is an area of current interest. In some settings, a hybrid of both human and machine crowdsourcing may be beneficial (Nagar and Malone, 2012) and hybrid forms require further evaluation.

The important concept of crowdsourcing has been taken up for market-based prediction markets, which operate with the additional benefit of a crowd-generated feedback signal or price, rather than a simple average. Evidence is mounting that prediction markets outperform not only the judgements of individual experts but also simple averages from the crowd. For example in a two year test on geopolitical questions, a prediction market performed 40% better than simple averages (Twardy et al., 2014). There has been a call for prediction market methodology to be more widely utilised in science (Pfeiffer and Almenberg, 2010) and this is now occurring through the SciCast platform (Twardy et al., 2014).

So where does this lead traditional open group work? Open work with member to member, face to face, or electronic communication is by nature limited in size, thus limiting cognitive diversity. Optimal group size will decrease as communication needs increase with higher task complexity (Mattingly and Ponsonby, 2014). Intra group social influences can be problematic as it can reduce the diversity of member opinions without reducing collective error (Mattingly and Ponsonby, 2014). Individuals' decisions may become correlated due to open interaction and a new set of social biases at the group level may be introduced such as group think, in-group bias and bandwagon effect (Mattingly and Ponsonby, 2014).

However, sometimes optimal communication and co-operation can only occur through the interactive open work group. Woolley et al. reported on group performance in over 700 individuals working in groups of 2 to 5 in a range of face to face tasks; a collective intelligence emerged from the open group beyond the sum total of individual input (Woolley et al., 2010). This collective intelligence was more strongly correlated with a proportion of conversation turn taking and social sensitivity rather than the average or maximum intelligence of individual group members. It was concluded that it may be easier to range the intelligence of a group rather than the individual (Woolley et al., 2010). Incentives for group participation need to be considered. In Citizen Science intrinsic motivation such as scientific curiosity and altruism are often engaged

DOI of original article: <http://dx.doi.org/10.1016/j.ebiom.2015.05.009>.

* Corresponding author.

E-mail address: anne-louise.ponsonby@mcri.edu.au (A.-L. Ponsonby).

URL: <http://www.mcri.edu.au> (A.-L. Ponsonby).

but for other types of crowdsourcing activities extrinsic rewards, such as monetary prizes are sometimes offered; such as by the data science platform of **Kaggle**.

Thus, when setting up a work group for science or medicine, one needs to consider whether a group aggregate or open work group is required, or a hybrid of both. The characteristics of the group should be explicitly considered: purpose, communication, size, members, incentives, techniques to support decision making, avoid bias and organisational structure. Here, Candido-do-Reis et al. provide a comprehensive formal evaluation of crowdsourcing with collective individual inputs compared with more traditional work practices in the setting of a cancer pathology study. Such evaluations are to be commended as we move towards new ways of working together collectively.

Conflicts of Interest

Anne-Louise Ponsonby is an Advisory Board Member of slowCapital. Karl Mattingly is a Director of slowCapital. slowCapital is a software company that is developing an internet platform that crowdsources forecasts.

References

- Candido-dos-Reis, et al., 2015. Crowdsourcing the general public for large scale molecular pathology studies in cancer. *EBioMedicine*, <http://dx.doi.org/10.1016/j.ebiom.2015.05.009> (<http://www.sciencedirect.com/science/article/pii/S2352396415300165>).
- Kaggle, d. The Home of Data Science <https://www.kaggle.com/> (Accessed 3rd June 2015).
- Karvetski, C., Olson, K., Mandel, D., Twardy, C., 2013. Probabilistic Coherence Weighting for Optimizing Expert Forecasts. *10* (4), 305–326.
- Mattingly, K., Ponsonby, A.L., 2014. A consideration of group work processes in modern epidemiology. *Ann. Epidemiol.* *24* (4), 319–323.
- Nagar, Y., Malone, T.W., 2012. Improving predictions with hybrid markets. AAAI Technical Report FS 12-06 Machine Aggregation of Human Judgement. Association for the Advancement of Artificial Intelligence, USA.
- Page, S.E., 2007. *The Difference: How the Power of Diversity Creates Better Groups, Firms, Schools, and Societies*. Princeton University Press, Princeton.
- Pfeiffer, T., Almenberg, J., 2010. Prediction markets and their potential role in biomedical research—a review. *Biosystems* *102* (2–3), 71–76.
- Surowiecki, J., 2004. *The Wisdom of Crowds: Why the Many are Smarter Than the Few and How Collective Wisdom Shapes Business, Economies, Societies and Nations*. Random House, New York.
- Twardy, C., Hanson, R., Laskey, K., Levitt, T.S., Goldfeder, B., Siegel, A., Dambrosio, B., Maxwell, D., 2014. SciCast: collective forecasting of innovation. *Collect. Intell.* (<http://blog.scicast.org/about/scicast-prediction-market-bibliography/> (Accessed 3 June 2015)).
- Woolley, A.W., Chabris, C.F., Pentland, A., Hashmi, N., Malone, T.W., 2010. Evidence for a collective intelligence factor in the performance of human groups. *Science* *330* (6004), 686–688.