

# Using the Wisdom of Crowds to Improve Hiring

October 11, 2016



The British statistician Francis Galton applied statistical methods to many different subjects during the 1800s, including the use of fingerprinting for identification, correlational calculus, twins, blood transfusions, criminality, meteorology and, perhaps most famously, human intelligence. Galton, who was an ardent eugenicist, believed that intelligence was a trait that only a minority of elite individuals possessed. The majority of common people, he believed, were not very competent decision-makers.

To put his theories to the test, Galton ran a famous experiment designed to analyze whether groups of common people were capable of making accurate choices. He asked 800 regular townsfolk at a county fair in Plymouth to guess the weight of an ox. People wrote down their estimates on bits of paper, which Galton then analyzed.

As it turned out, the median of all 800 guesses was very close to the correct answer. Much closer even than individual guesses from the oxen experts (farmers and butchers). The group estimate was 1,197 pounds and the actual weight of the ox was 1,198 pounds – a difference of just 0.08%.

This “wisdom of the crowd” demonstrated that, under the right conditions, groups of people can make more insightful decisions than individuals, sometimes even besting the experts.

In the UK, the [Behavioural Insights Team](#) (BIT) is now investigating ways to use this crowd-sourced wisdom to improve hiring decisions. People with different backgrounds and experiences will tackle problems differently, and this diversity of perspectives can help organizations make better decisions.

“Organisations spend eye-watering sums trying to attract the best talent because in many industries, the difference between the best and the good has real implications for the bottom line,” BIT contributors Kate Glazebrook, Theo Fellgett, and Janna Ter Meer write in a blog post.

Yet, many hiring decisions come down to superficial criteria, such as choosing to interview only graduates from certain universities or unconsciously favoring candidates based on traits like gender, ethnicity, or even favorite sports teams.

The BIT wanted to find out whether they could build better, more diverse teams by adopting a hiring strategy based on the wisdom of crowds. Research from APS Fellow [Philip Tetlock](#) (University of Pennsylvania) has demonstrated that people are better at forecasting future outcomes when they work together in collaborative teams.

“In fact, researchers have even shown that US defense intelligence analysts with access to classified information can be beaten by some rudimentarily-educated amateurs: largely because they come to conclusions too quickly and struggle to update their opinions in the face of new and conflicting information,” the BIT explains.

A team of psychological scientists, led by APS Fellow Adam Galinsky (Columbia University), recently summarized empirical arguments for more diverse teams in [Perspectives on Psychological Science](#): “Homogeneous groups run the risk of narrow mindedness and groupthink (i.e., premature consensus) through misplaced comfort and overconfidence. Diverse groups, in contrast, are often more innovative and make better decisions, in both cooperative and competitive contexts.”

So, when it comes to reviewing resumes and interviewing applicants, how big does the crowd need to be to maximize the benefits?

The BIT designed a simple online experiment in which around 400 reviewers rated four hypothetical job candidates based on responses to a generic recruiting question (i.e., “Tell me about a time when you used your initiative to resolve a difficult situation?”). The crowd had a clear favorite, and easily identified the best response.

“We took our data and ran statistical simulations to estimate the probability that different groups could correctly select the best candidate,” Glazebrook and colleagues explain. “We created 1,000 combinations of reviewers in teams of different sizes, ranging from one to seven people. We then pooled them by the size of the group and averaged their chance of selecting the right candidate.”

When there was a gap in quality between the best and second best responses, an individual picked the less qualified person around 16% of the time. However, with a group of three decision makers, the odds of choosing the lesser candidates dropped to 6%, and with a five person group it dropped to 1%. When the two candidates were very similar, individuals selected the best candidate around 50% of the time – basically, they had the same accuracy as tossing a coin. A crowd of seven, on the other hand, picked the superior candidate more than 70% of the time.

Of course, 400 reviewers for every job filled isn’t very practical. Ultimately, the evidence suggested that three reviewers was the optimal crowd size for recruitment, but more experiments are still in the works.

## References

Galinsky, A. D., Todd, A. R., Homan, A. C., Phillips, K. W., Apfelbaum, E. P., Sasaki, S. J., ... &

Maddux, W. W. (2015). Maximizing the Gains and Minimizing the Pains of Diversity A Policy Perspective. *Perspectives on Psychological Science*, 10(6), 742-748. doi: 10.1177/1745691615598513

Glazebrook, K., Fellgett, T., Ter Meer, J. (2016, February, 17). Would you hire on the toss of a coin? Retrieved from <http://www.behaviouralinsights.co.uk/labour-market-and-economic-growth/would-you-hire-on-the-toss-of-a-coin/>