One Among Many

The self in social context

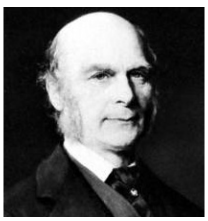
by Joachim I. Krueger, Ph.D.

Joachim Krueger is a social psychologist at Brown University who believes that rational thinking and socially responsible behavior are attainable goals.
See full bio

Beware of small majorities

Ignoring a third, dissenting, opinion can lead to bad decisions.

Published on January 14, 2010



In a recent post I argued that, although a lot can be said for the basic rationality of human (and non-human) conformity, there can also be trouble. The example du jour was a beauty contest. With too much choice copying among females (women choosing men whom other women choose), both females and males can suffer, on average.

Another limitation is the size of the majority that is being copied. Suppose you are trying to estimate the number of marbles in a glass jar. There are more marbles than you can count. Yet, you can make a guess using your impression of the size of the jar and the size of the individual marbles. Now suppose you are told that 100 other people have already made estimates independent of one another and that 95% of these estimates fall between 700 and 800. With the benefit of this information, your best strategy is to estimate that there are 750 marbles in the jar. If you estimated the number as 200, you would recognize yourself as an outlier, whose judgment should not be trusted. Using the aggregated information of others' estimates is like using the "poll-the-audience" lifeline on "Who wants to be a millionaire?" But what if you estimated the number of marbles to be 200 before knowing the others' estimates? Once you find out how far off you are from the rest of the group, you should not object when your estimate is removed as an outlier.

As a general rule, it becomes easier to identify outliers as the number of observations increases and as the variance or these observations decreases. Suppose there are only two others who both gave a high estimate, whereas you gave a low estimate. Should you gracefully concede on the idea that their estimates are probably more accurate than yours because they agree with each other, whereas you disagree? This is a tempting thought. Perhaps agreement reveals accuracy even when the number of agreeing persons is at its logical minimum.

I will now argue that agreement is a mere proxy of accuracy, and not a particular good one at that. It is true that if all judgments are accurate, they will all agree with one another. The inverse, however, is not true because judgments can be in agreement for reasons that have nothing to do with accuracy. One of these reasons is chance.

The alternative way to proceed is to use all three judgments (yours and the judgments of the two others) and compute the average. The average is the best estimate of the latent parameter that you are all trying to capture. According to this approach, each of the three judges is an independent measurement instrument and each individual judgment is a composite of information (truth) and error. The errors are assumed to be independent of one another, and averaging judgments strips them away.

We now have two competing recommendations on how to proceed if there are two high and one low judgment. (A) Remove the low judgment or persuade the outlying judge to change his mind and conform with the majority; (B) average the three judgments without prejudice against any individual one of them. Each method has its advocates. The main argument for A is that the low estimate is "obviously" and outlier and that agreement indicates accuracy [I already questioned this idea]. Moreover, supporters of A believe that consensus-seeking discussion among judges

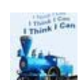
Subscribe to One Among Many
Subscribe via RSS

Recent Posts in One Among Many


Doing one's duty as a counterpoint to destructive obedience by Joachim I. Krueger, Ph.D.


Ignoring a third, dissenting, opinion can lead to bad decisions. by Joachim I. Krueger, Ph.D.


Women choose whom other women choose. by Joachim I. Krueger, Ph.D.


Confidence predicts success, unless you choose your own challenge. by Joachim I. Krueger, Ph.D.

Positive emotions underlie spirituality. by Joachim I. Krueger, Ph.D.

More of One Among Many blog

Most Popular Posts

Most Read | Most Emailed

1 There Are No Bad Racks
by Christopher Peterson, Ph.D.

2 
Online Mindfulness Meditation Therapy for PTSD, Boulder Colorado
by Peter Strong, Ph.D.

3

is always salutary. Through discussion, judges can come closer to the truth. But which truth? If the two high judges concede a little and the low judge concedes a lot, the end result might be the average that was already computed from the original judgments. If so, the group discussion was a waste. Alternatively, if only the outlying judge concedes (which is likely to happen under asymmetrical conformity pressure), the end result is what one would get by simply ignoring the outlier. Again, the group discussion was a waste of time and adrenaline. A third possibility is that the outlying judge concedes a bit more than the two agreeing judges put together. The end result is a group judgment that can be described as a weighted average (where an individual judgment is weighted in proportionate to its proximity to other judgments). This sounds good, but no one knows what the weights should be exactly. There are many points where the weighted judgment may end up that lie between the pure strategies A and B. Therefore, I will consider only A and B in the remainded of these remarks.

By using two statistical principles, we can determine whether A or B is the better strategy without appealing to intuition, plausibility, or tradition (we have always done it this way!). The first method is to ask how probable the set of three observed judgments if we assume that A or B is correct. Suppose the three judgments are 2, 2, and -2. Think of these numbers as a sample drawn from a population with a standard deviation of 1. In contrast to the standard normal distribution, however, the mean is not 0. Instead, the mean is either 2 if we assume that theory A is correct, or it is .667 (2/3) if theory B is correct. The joint probability of finding 2, 2, and -2 (or numbers more extreme) turns out to be .000008 under theory A and .00003 under theory B. The ratio of the latter over the former is 3.75, which means that (if both theories were regarded as equally likely to be true at the outset), theory B is almost four times more likely to be true than theory A. This result means that if you remove the outlying judgment (or persuade the dissident judge) to change her mind, you lose important information, whereby the resulting group judgment becomes worse.

The second method is to ask what would happen if more judgments were collected from other independent observers [note that there is no need to actually get those judgments!] We now assume that the population of numbers underlying all these judgments is a standard normal (M = 0, SD = 1). Hence, assuming the set of numbers associated with theory A after outlier removal or correction (2, 2, 2) is extremely positive. If another set of three judgments were sampled from the population, the resulting mean would most likely lie between 0 and 2, and closer to the latter inasmuch as the whole measurement process is reliable. As measurement is never completely free from error, we expect some regression to the mean. Now, assuming the set of numbers given by theory B (2, 2, -2), the mean of the second sample of three judgments would most likely lie between 0 and 2/3, and because the 2/3 is less extreme than 2, the size of the expected regression effect is smaller under theory B than under theory A.

The significance of this exercise is that it reveals that ignoring (or browbeating) outliers in a small sample does not correct the well-known regression effect in measurement, but that it makes it worse. The best estimate under theory B (2/3) is probably a bit higher than it would be after continued sampling. If anything, this estimate should be reduced. By cutting off the outlier, however, we move the group estimate from 2/3 to 2. By making it more extreme, we are making it more likely to be positively inflated.

Lest you think that this story is too abstract and that theories A and B don't make no never mind anyway, let me emphasize that they matter a great deal when small committees decide admissions, funding, promotions etc. Consider 100 candidates applying for money to do research. Each proposal is rated by three judges and each judge's scores are standardized. Only the top few can be funded. A proposal with ratings of 2, 2, and 2 is safe, but a proposal with ratings of 1, 1, and 1 is not. Now a third proposal is of the kind discussed above (2, 2, -2). According to theory B (simple averaging), this proposal does not make the cut. According to theory A (outlier removal), this proposal rises above the second one, and possibly keeps it from being funded. So group discussion can do a lot of damage. If, as in this example, relatively high scores are of greatest interest, proposals (people) with one negative outlier will be selectively favored. In a funding or promotion context, no one is interested in cases with two low scores and one high score.

The approach of simple averaging is not biased against good cases with low variance (1, 1, 1), whereas outlier-removal is. There is one final caveat. Group discussion by itself tends to make judgments more extreme.

In other words, if the good low-variance case (1, 1, 1) were also discussed, some of the damage done by the outlier-removal ploy might be controlled.

BTW, the gentleman in the photo is Sir Francis Galton.

Print | Send to friend

Have a comment? Start the discussion here!

Tags: consensus, decision-making, judgment, outliers