

The Wisdom and Manipulability of Threads

wisdom of crowds | magnitude estimation | decision-making | threads | dot-guessing games

Extended Abstract

Social decision-making is increasingly relying on digitized aggregates of people’s opinions and judgments. These aggregates are frequently maintained as threads, i.e. as sequences of posts on a website. While it has been shown that knowledge of thread aggregates can distort individual decision-making, it is unknown how the ability to see preceding posts in a thread may influence collective accuracy, i.e. the wisdom of threads. We investigate experimentally the accuracy and manipulability of threads, in which people make numerical estimates of varying difficulty and varying degrees of social information. Analyzing the data using gaussian mixture models (GMMs), we obtain novel insights into the intrinsic use of social information within each thread.

The data was collected from 7,814 participants playing a dot-guessing game [1] on Amazon Mechanical Turk, where each participant successively estimated the number of visible dots $d \in \{55, 148, 403, 1097\}$ in various images, while seeing $v \in \{0, 1, 3, 9\}$ preceding estimates (historical threads). An additional 3,934 participants were shown the same images while seeing the $v \in \{1, 3, 9\}$ *highest* estimates made so far (manipulated threads).

We compare the log-ratio of the individual thread medians, $\log(M_{(d,v)}/d)$, using a linear normal model to quantify differences between threads in terms of $\log(d)$ and v , the latter as a categorical variable. Figure 1a shows that the collective performance for historical threads declines significantly with increasing task difficulty (higher d), but improves when the social information is substantial (high v). In contrast to [2] and in accordance with [3], this finding supports the claim that crowds indeed may become wise under (pristine) social influence. For the manipulated series, Fig.1b, the manipulation gives a large positive bias for $v = 3, 9$ which increases with d , implying that when a task becomes more demanding, the amount of (biased) social information has a detrimental impact on thread performance.

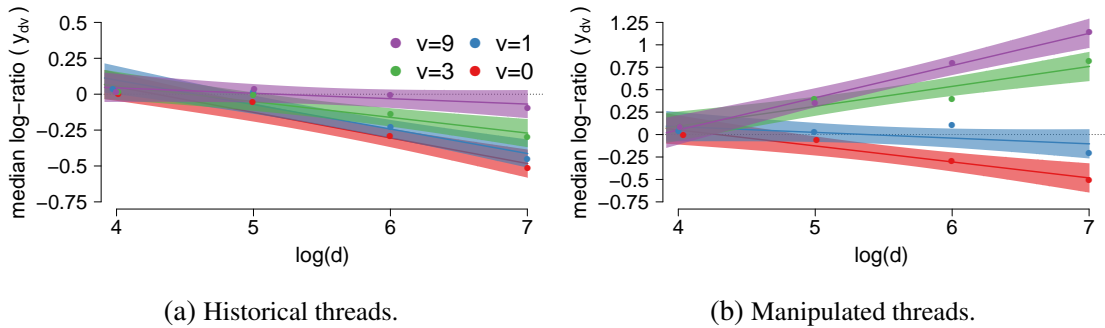


Figure 1: Relationship between median log-ratio and number of dots (lines) with 95% confidence bounds (shaded areas). The colors represent the four settings on number of visible estimates. There is a clear relation between number of dots and number of visible estimates in both thread types, but it is more pronounced for the manipulated series. The red lines corresponding to the control groups with $v = 0$ are the same in both plots.

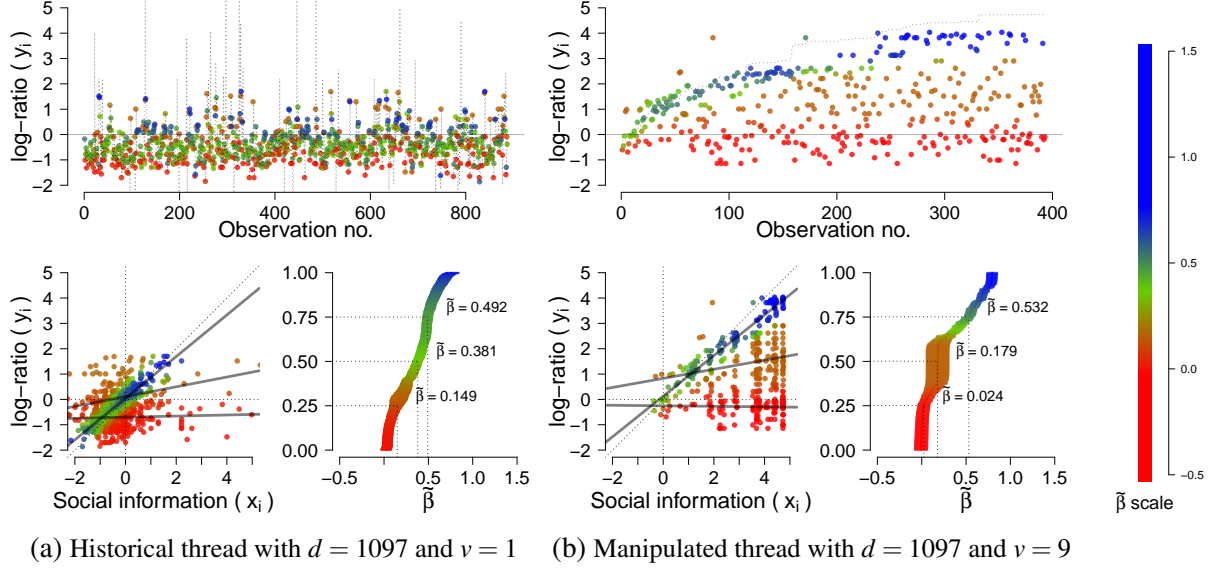


Figure 2: The left hand side shows three plots of a historical thread with $d = 1097$ and $v = 1$, and the right hand side shows three plots of a manipulated thread with $d = 1097$ and $v = 9$. Top plots shows the log-ratio estimates over time (observation no.), with the geometric mean of the social information shown by a dotted line. Bottom left plots show the log-ratio of the estimates as a function of the log-ratio of the social information, indicating how differently participants use their social information. Bottom right plots show the cumulative distribution of individual $\tilde{\beta}$'s with 95% intervals derived from the fitted models and added interquartile values of $\tilde{\beta}$.

Figure 2 displays the effect of social information, obtained by fitting a GMM to each thread using the geometric mean of visible estimates as a proxy for the available social information. The effect of social information on each participant is given by $\tilde{\beta}$, the weighted average of the individual states of the GMM, where a high value ($\tilde{\beta} > 0.6$, blue color) implies a large effect, a low value ($\tilde{\beta} \approx 0$, red color) implies little or no effect and a medium value ($\tilde{\beta} \approx 0.4$, green color) suggests a compromise between the two extremes. Figure 2a shows results for a historical thread where estimates are split among keepers (red), compromisers (green) and adopters (blue) throughout the thread. Figure 2b in contrast shows the results for a manipulated thread, where the same three groups are clearly visible, but the evolutionary trend is completely different compared to the historical thread. Contrasting Figures 2a and 2b clearly reveals different dynamics of threads in terms of manipulation or not. In general, our experiments show distinct distributions of overlapping groups (keepers, compromisers and adopters) in each thread, and in the case of manipulation suggest a substantial split into enthusiastic adopters versus skeptic keepers. The GMM framework is also applied to independent data from [4] allowing similar insights into qualitatively different tasks.

- [1] J. J. Horton, The dot-guessing game: A ‘fruit fly’ for human computation research, SSRN 1600372 (2010).
- [2] J. Lorenz, H. Rauhut, F. Schweitzer, D. Helbing, How social influence can undermine the wisdom of crowd effect, P Natl Acad Sci Usa 108 (2011) 9020–9025.
- [3] J. Becker, D. Brackbill, D. Centola, Network dynamics of social influence in the wisdom of crowds, P Natl Acad Sci Usa 114 (2017) E5070–E5076.
- [4] B. Jayles, H.-r. Kim, R. Escobedo, S. Cezera, A. Blanchet, T. Kameda, C. Sire, G. Theraulaz, How social information can improve estimation accuracy in human groups, P Natl Acad Sci Usa 114 (2017) 12620–12625.