The Wisdom of Crowds with Communication

Brent J. Miller (brent.miller@uci.edu)
Mark Steyvers (mark.steyvers@uci.edu)

Department of Cognitive Sciences
2201 Social & Behavioral Sciences Gateway Building
University of California
Irvine, CA 92697

Abstract

The average estimates of a group of individuals are generally better than the estimates of the individuals alone, a phenomenon commonly referred to as the wisdom of crowds. This has been shown to work for many types of simple tasks, but has generally been performed on subjects that do not communicate with one another. We report group aggregation performance for more complex tasks, involving reconstructing the order of time-based and magnitude-based series of items from memory. In half of these tasks, subjects receive the previous subject's final ordering in a serial fashion. The aggregate for communicating subjects is better than that for independent subjects. We introduce a Bayesian version of a Thurstonian model to show how each subject combines their individual, private knowledge with the previous individual's ordering. The model also shows that individuals can produce estimates in the shared information condition that are better for aggregating than independent estimates.

Keywords: Bayesian Modeling; Rank Ordering; Wisdom of Crowds; Iterated Learning.

Introduction

When Galton first surveyed English fair-goers in 1906, it was a novel curiosity that their estimates of the dressed weight of an ox, when averaged, closely approximated the true weight (Galton, 1907). Subsequently, many demonstrations have shown that aggregating the independent judgments of a number of individuals often results in an estimate that is close to the true answer. This phenomenon has come to be known as *The Wisdom of Crowds* (Surowiecki, 2004). This *Wisdom of Crowds* (WoC) is currently used in a number of real-world applications, such as prediction markets (e.g., Dani et al., 2006) and political polling (e.g., Silver, 2008).

Previous WoC research has focused on independent subjects, where there is no communication between individuals (Galton, 1907; Bruce, 1935; Lorge et al., 1957; Surroweicki, 2004). Many authors have suggested that subject independence is important for creating the WoC (Surroweiki, 2004; Sjöberg, 2006; Vul et al., 2008). Many of these WoC demonstrations have involved estimating single numerical quantities. For such quantities, robust estimates of the central tendency of individual estimates (such as the mean) are an effective aggregation method (Yaniv, 1997).

In the current research, we wish to see if there are situations where allowing communication between subjects can improve WoC aggregation. In contrast to previous research, we are interested in exploring WoC in high-dimensional problems, such as rank ordering, where items need to be placed in a sequence. We have shown previously that the WoC effect can still be observed in these problems, even though simple distributional estimates such as the mean are no longer sensible, or even possible (Steyvers, Lee, Miller, and Hemmer, 2009).

There are many ways in which subjects can communicate. In traditional "face to face" group decision making, subjects share information by meeting and talking with one-another, and attempt to reach a group decision through discussion. This process becomes more complicated when subjects need to generate individual answers for aggregation, however. There are methods in the decision making literature designed to limit and control group interaction to aid in post-hoc statistical aggregation. Some of these also have the advantage of improving overall group performance (Kerr & Tindale, 2004). These include forcing subjects to communicate electronically (Gallupe et al., 1994), and limiting the amount of information that is shared (Rowe & Wright, 1999).

One method for controlled group interaction that is also amenable to indirect, computer-based information sharing is iterative communication. This kind of communication is used in iterated learning models, such as those used to study the evolution of syntactic structure (Kirby, 2001). Similar to the children's game "telephone", each subject is placed in some arbitrary order from first to last. The first subject begins by making an estimate, which is passed on to the second individual. The second subject makes an estimate that is passed on to the third, and so on through the entire group. In this way, every subject (except the first) gives and receives exactly one estimate. Unlike in iterated learning, each subject has independent knowledge in addition to the received estimate. They combine this knowledge with the given estimate to create a new estimate that is passed on to the next individual.

Beppu and Griffiths (2009) have done work involving iterated learning models with individual knowledge. By varying the types of independently known and communicated information, they were able to create an iterative environment where the last subject in the group could recreate a distribution that early individuals had little knowledge of.

Unlike Beppu and Griffiths, we are not focused on the answer from the last individual. While their subjects were

estimating a two-dimensional function, our subjects were involved in making higher-dimensional rank ordering estimates. Because of this, the informativeness of each subject's shared estimates was much lower, and our iterative subjects were not likely to converge on the correct answer and stay there. To preview our results, the last iteratively communicating subject's estimate for each question in our experiment was not much better than the average response. The average aggregate for these iteratively communicating subjects, however, was a better estimate than the average aggregate for the independent subjects who had not received one.

In this paper we present empirical and theoretical research on communication and the WoC phenomenon for rank order aggregation. We conduct an empirical study where subjects are asked to rank order the occurrence of events (e.g., US presidents by term of office) or the magnitude of some physical property (e.g., rivers by length). Importantly, for all of the questions there is a known ground truth. The ground truth might only be partially known to the tested individuals. Subjects answer half the questions independently, and half with iterated communication.

We compare the performance of WoC aggregation between independent subjects and "iterated" subjects. We also develop a probabilistic model to describe the behaviors of both types, based on a Thurstonian approach that represents the items for each problem as distributions on an interval dimension. We use Markov chain Monte Carlo to estimate parameters from these models, and make qualitative and quantitative comparisons between our proposed models and the actual experimental data.

Experiment

Method

Subjects were 172 undergraduate students at the University of California, Irvine. The experiment was composed of 17 questions, identical to those used in Miller et al. (2009) and Steyvers et al. (2009). All were general knowledge questions regarding: population statistics (4 questions), geography (3 questions), dates, such as release dates for movies and books (7 questions), U.S. Presidents, material hardness, the 10 Commandments, and the first 10 Amendments of the U.S. Constitution. For half of the questions (picked randomly for each subject), subjects started with the final ordering from a previous subject, allowing subjects to communicate iteratively. The first participant in the iterative condition received a random ordering. For the other half, subjects received a randomly ordered list of items as a non-communicative control, maintaining independence between subjects. Because of the uneven number of questions, we made the arbitrary choice to split the questions 9/8 between the iterative and independent conditions. For each question, there was a group of people who answered that question with communication, and a group who answered it without, and all individuals participated in both conditions.

All questions had a ground truth obtained from Pocket World in Figures and various online sources. An interactive interface was presented in a web browser on computer screens. Subjects were instructed to order the presented items (e.g., "Order these books by their first release date, earliest to most recent"), and responded by dragging the individual items on the screen using the computer mouse and "snapping" them into the desired locations in the ordering. When subjects were satisfied with their response, they clicked on a "submit" button. Once their response had been submitted, it was not possible to return to that question. The independent subject questions, where subjects received randomized starting lists, were presented first. These questions were then followed by a transition screen informing them that, "Each of the lists you will see next have already been sorted by *at least* one other subject. It is up to you to determine how well that/those subject(s) did, and to make any changes you feel are necessary." The subjects were then presented with the iterative questions, where the starting lists were sorted by the previous subject. Each iterative question carried a reminder that items had already been sorted by at least one subject (no information about the number of previous subjects was communicated). The assignment of questions between both conditions, and the order of the questions, was selected randomly.

Table 1: *Subject performance statistics*.

	Independent				Iterative				
Problems	PC	Group Ave.		Best	PC	Group Ave.		Best	Last
Book Releases	0.000	5	12	3	0.000	7	11	4	11
Euro. City Pop.	0.000	11	16	1	0.000	10	14	7	7
US City Pop.	0.025	12	15	0	0.021	10	12	0	6
World City Pop.	0.000	12	19	6	0.000	17	18	8	13
Landmass	0.000	5	10	2	0.000	5	6	1	7
Population	0.000	11	15	1	0.012	9	12	0	7
Hardness	0.000	12	17	7	0.000	13	14	7	11
Holidays	0.099	4	9	0	0.161	2	5	0	3
Movies	0.000	2	12	1	0.033	1	5	0	10
Oscar Movies	0.022	3	14	0	0.000	4	6	1	9
Osc. Best Movies	0.000	2	15	1	0.000	1	7	1	6
Presidents	0.103	3	11	0	0.218	1	5	0	2
Rivers	0.000	10	17	3	0.000	10	12	3	10
East-most States	0.036	4	9	0	0.088	1	4	0	1
Super Bowls	0.000	21	21	4	0.000	13	15	4	8
10 Amendments	0.081	5	14	0	0.000	4	11	1	9
10 Command.	0.022	14	19	0	0.108	8	11	0	1
AVERAGE	0.023	8.0	14.4	1.7	0.038	6.8	9.8	2.2	7.1

Results

Individual Subjects We first evaluated subjects' responses based on whether or not they reconstructed the correct ordering. Table 1 shows the proportion of individuals that got the ordering exactly right. Proportion Correct (PC) was evaluated for each of the ordering task questions, for both the independent and iterative conditions. On average, about two and four percent of subjects recreated the correct rank ordering respectively. We also analyzed the performance of subjects with a more fine-grained measure, using Kendall's τ distance. This distance metric is used to count the number

of pair-wise disagreements between the reconstructed and correct ordering. The larger the distance, the more dissimilar the two orderings are. Values of τ range from: $0 \leq \tau \leq N(N-1)/2$, where N is the number of items in the order (ten for all of our questions). A value of zero means the ordering is exactly right, a value of one means that the ordering is correct except for two neighboring items being transposed, and so on up to the maximum possible value of forty-five (indicating that the list is completely reversed). A score of 22.5 indicates roughly random performance.

Table 1 also shows the average and the best τ values for each of the seventeen sorting task questions. For the independent condition, one or more subjects got the ordering exactly right for seven questions, as indicated by a τ of zero for 'Best'. For the iterative condition, one or more subjects also got the ordering exactly right for seven questions, five of which were the same questions as in the independent condition. The best individuals on each question solved the problem exactly, or were within a few pair transposes, for most questions, indicating very good performance. Generally, subjects performed better in the iterative condition than they did in the independent condition, particularly the worst subjects. Performance gains were realized for all questions, not just the difficult ones (as indicated by high τ scores), with the notable exception of book release dates.

Figure 1 shows the performances of subjects averaged across questions in the independent condition, compared to the performances of subjects averaged across questions in the iterative condition. The subjects are ordered from worst to best performance. It is important to note that the best performances on the right are for individuals that did not have to complete the most difficult questions under that condition (independent or iterative). For this reason, the subjects are ordered in each condition for their performance within that condition, and there is no implicit relationship between the n^{th} subject in the independent condition and the independent condition. Subjects perform better in the iterative condition, and their estimates were approximately 6 τ closer to the ground truth, on average.

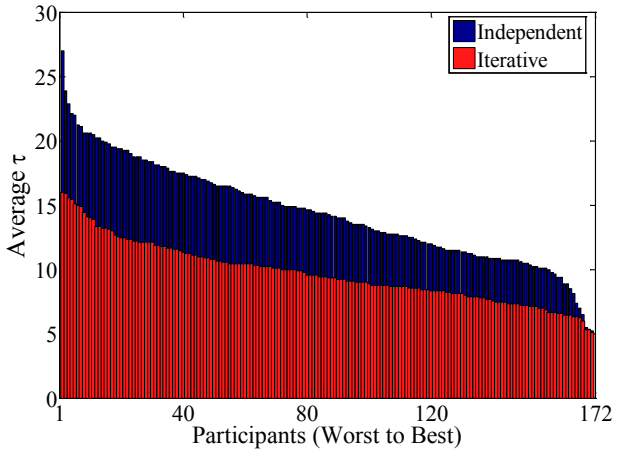

Figure 1. Average Subject Performance over Independent and Iterative Conditions.

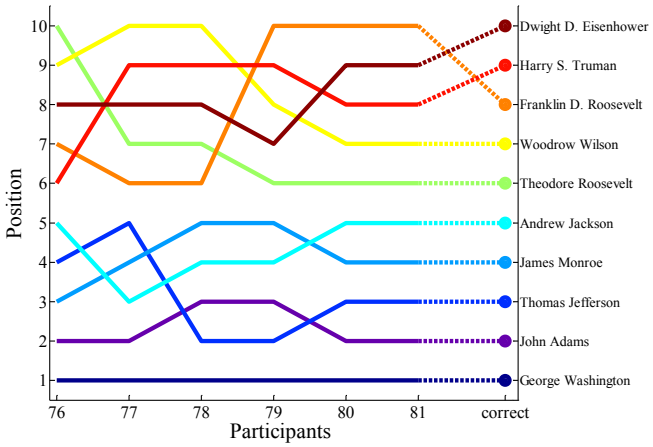

Figure 2. Subject Ranking for U.S. Presidents.

As this is a task that is effectively assessing prior knowledge, it is interesting to note that three out of the four questions with the best individual performances in both conditions occurred for the U.S. Presidents, Eastern-most States, and Dates of Holidays questions. These three questions relate to educational and cultural knowledge that seems most likely to be universally shared by our undergraduate subjects. The other two questions, the 10 Amendments and the 10 Commandments from the independent and iterative tasks respectively, are the kinds of more specialized knowledge one would expect particular individuals to have expertise in.

It is also interesting to note that the Book Release Dates question was not improved by iterative communication. Some of the books in question, such as Harry Potter and the Sorcerer's Stone, were also optioned for movie releases. It appears that subjects in the iterative condition were divided between the ordering books based upon movie releases and the actual book release dates, and so were unable to communicate much useful information between each other as they were drawing upon two contradictory sources of knowledge.

It was clear that subjects were able to communicate useful information in the iterative condition, and that they did so in an iterative fashion that was well suited our task. Figure 2 shows one such example from the U.S. Presidents question. The 76^{th} subject's final response has a τ of 9 compared to the ground truth, but by the 80^{th} subject's response, that had been reduced to 2. Each intervening subject was able to add some information, either by bringing an item closer to, or placing it in, its correct position.

It is important to note that iterative responses did not always converge so neatly however, and that when they did, they did not remain there. Table 1 lists the last responses from each iterative chain for each question (Iterative, Last). Even for questions on which subjects did relatively well overall, such as Movies, the final answer from subjects was often far from best answer. The final answer is never the exact correct ordering and final ordering equivalent to the best ordering only once. While iterative learning tasks usually converge to the best answer, the lack of convergence

in our data makes it unclear who had the best answer for the iterative (and independent) questions, barring access to the ground truth. This necessitates an extra method to select the best subjects.

Group Aggregate In order to test the collective performance of aggregation, we used the Borda count method, a popular technique for combining preferential candidate lists in voting theory. We have shown previously that this is a representative heuristic that performs among the best WoC methods for aggregating lists (Miller et al., 2009). In the Borda count method, weighted 'counts' are assigned such that the first item in a list of *N* items receives a count of *N*. The second item in the list receives a count of *N*-1, and so forth until the last item receives a count of 1. These counts are summed across the items for all subjects, and the item with the highest count is placed first in the aggregated list, with all subsequent items ranked according to their respective count totals. We treat that aggregate as its own subject, and calculate a τ score relative to the ground truth.

Table 1 reports the performance for Borda aggregation for each of the problems for each of the conditions (Independent/Iterative, Group). The average performance of the aggregate is better than the average performance of subjects, and is among the best performances for all subjects, in agreement with our findings in Miller et al. (2009). It is important to note that the "Best" column is the score of the best individual for that particular problem, and that no one individual significantly outscored the others or the aggregate across both conditions.

Comparing across conditions, we can see that the aggregate performed better for subjects that were allowed to communicate iteratively than for those that could not. This shows that the communicative behavior seen between subjects in Figure 2 does in fact improve overall group knowledge.

The size of the group also matters for group performance. Figure 3 shows the average performance of the aggregate over varying group sizes for both the iterative and independent conditions. With one subject, performance can be dramatically improved with just one more estimate. For a large enough group however, the benefit of adding an additional subject can be negligible.

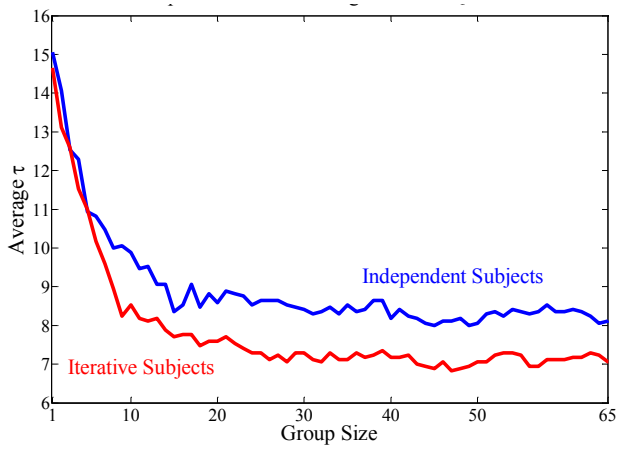

Figure 3. Aggregate Subject Performance Averaged over Questions for the Independent and Iterative Conditions.

We can see that the aggregate for the iterative condition clearly outperforms the independent aggregate, as shown in Table 1. We can also see that the performance gained by adding a subject diminishes rapidly for independent crowds after around 15 subjects. For the iterative group however, gains are made much more rapidly, and so further subjects become less important after about the 10th subject. This suggests that subjects are able to successfully combine their own information with that of previous subjects, and thus create a knowledgeable crowd with fewer subjects. We will explore this notion further in the next section.

Modeling

We designed two probabilistic generative models, based on a Thurstonian approach, to detail a process by which independent and iterative subjects might use item-level knowledge to create their rank orderings. In the iterative model, subjects must combine their item-level knowledge with that of previous individuals. We will first specify these models, and then compare their performance to our empirical results.

A Thurstonian Approach

A feature of a Thurstonian approach is that the ground truth is explicitly represented at the item level, as a latent set of coordinates on an interval dimension. Specifically, each item *i* is represented as a value $\boldsymbol{\mu_i}$ along this dimension, where $\boldsymbol{i \in \{1, \ldots, N\}}$. The interval representation of items is justifiable given that all the problems in our study involve one-dimensional concepts (e.g., the relative timing of events, or the lengths of items). See Figure 4a for an example.

Each individual is assumed to have access to the ground truth for each item, $\boldsymbol{\mu_i}$. Each individual, however, does not have precise knowledge about the exact location of each item, $\boldsymbol{\mu_i}$. For each individual *j*, we represent the individual's uncertainty about the location of item *i* with a Normal distribution, centered about $\boldsymbol{\mu_i}$, with a standard deviation $\boldsymbol{\sigma_{ij}}$. See Figure 4b for an illustration.

This individual item-level uncertainty, σ_{ij}, is different from previous Thurstonian models we have presented, where item uncertainty was represented only for the group as a whole (Steyvers et al., 2009). This allows us to represent the individual item-level differences in knowledge that affect communication in our iterative condition.

In order to make an ordering estimate, individuals draw samples for each item, x_{ij}, and the ordering of the samples becomes the ordering for the individual, $\mathbf{y}_j$. See Figures 4b and 4c for examples. The smaller the uncertainty for an item, σ_{ij}, the more knowledge a subject has about that that item, and the more likely they are to sample values close to the ground truth. Different subjects have different

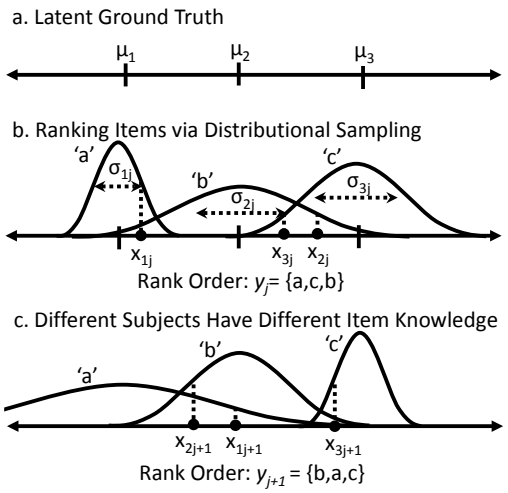

Figure 4. Illustration of the Thurstonian Model.
a) The Thurstonian interval dimension representation
b) Item-level uncertainty, sampling and order for subject *j*
c) The sampling and order for subject *j+1*

knowledge, making each participant's accuracy in ranking individual items different.

Independent Condition: Generative Model

Figure 5a shows the Thurstonian generative model for independent subjects for a single question, using graphical model notation (see Koller, Friedman, Getoor, & Taskar, 2007, for an introduction). The nodes represent variables, and the arrows indicate the conditional dependencies between the variables. Stochastic and deterministic variables are indicated by single and double-bordered nodes respectively. Observed data is represented by shaded nodes. The rectangles, or "plates", represent independent replications of sampling steps over items or subjects.

In the independent subject generative model, each individual j draws samples, x_{ij}, from the ground truth for each item, μ_i, with individual item-level uncertainty σ_{ij}. The final observed rank ordering, $\mathbf{y}_j$, is determined by the rank order of the subjects samples:

$$x_{ij} \sim \text{Norm}(\mu_i, \sigma_{ij})$$
$$\mathbf{y}_j = \text{Rank}(\mathbf{x_j})$$

Iterative Condition: Generative Model

Figure 5b shows the generative model for subjects in the iterative condition. Each subject j, for each item i, has access to samples, x_{ij}, from their knowledge of the ground truth, μ_{ij}, and their own item-level uncertainties, σ_{ij}. Samples are drawn from a normal distribution, identical to the independent model. Subjects' access to their samples and uncertainties allows them to quantify their item-level knowledge.

Subjects also have access to the previous subject's final ordering, $\mathbf{y}_{j-1}$. They combine these sources of information to generate their own estimate, $\mathbf{y}_j$. To do this, subjects infer

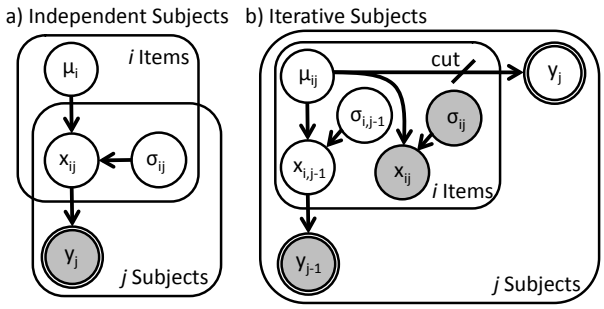

Figure 5. Independent and Iterative Generative Models.

the latent ground truth for items, μ_{ij}, that would most likely give rise to these observed values. The ground truth for item i is dependent upon the observed individual samples and uncertainty for that item, as well as for the latent samples and uncertainty of the previous participant. Using Bayes rule:

$$P(\mu_{ij}|x_{ij}, \sigma_{ij}, x_{i,j-1}, \sigma_{i,j-1}) \propto P(x_{ij}|\mu_{ij}, \sigma_{ij})\, P(x_{i,j-1}|\mu_{ij}, \sigma_{i,j-1})$$

$$\mu_{ij} \sim \text{Uniform}(0,1)$$

As the distributions on the right side of the equation are normally distributed, we can reduce it to the following sampling distribution:

$$P(\mu_{ij}|x_{ij}, \sigma_{ij}, x_{i,j-1}, \sigma_{i,j-1}) \propto \text{Normal}(\mu^*, \sigma^{*2})$$

Where μ^* and σ^* are essentially weighted averages:

$$\mu^* = \left(\frac{x_{ij}}{\sigma_{ij}^2} + \frac{x_{j-1}}{\sigma_{j-1}^2}\right) \Big/ \left(\frac{1}{\sigma_{ij}^2} + \frac{1}{\sigma_{ij-1}^2}\right), \sigma^{*2} = 1 \Big/ \left(\frac{1}{\sigma_{ij}^2} + \frac{1}{\sigma_{ij-1}^2}\right)$$

Participants infer the ground truth for items, μ_{ij}, by combining their observed and inferred samples of the ground truth, x_{ij} and $x_{i,j-1}$, weighted by the uncertainty for those samples, σ_{ij} and $\sigma_{i,j-1}$. This weighted averaging of each subject's own knowledge and the knowledge communicated to them yields a final result that is a combination of both. As illustrated by the cut notation in the graphical model, the final subject ordering, $\mathbf{y}_j$, is determined by the rank order of the averaged posterior samples for items for μ_{ij}.

Simulation

We simulated 65 subjects for 17 questions, using both the independent and iterative generative models. In order to ensure a fair comparison between conditions, subjects generated by the independent model were used for the observed individual knowledge parameters in the iterative model.

We use the same inverse-gamma prior for all subjects' item-level uncertainty, σ_{ij}:

$$\sigma_{ij}^2 \sim \text{InvGamma}(\alpha, \beta)$$

Where α and β are shaping parameters that were hand-fitted to create distributions of subject ranking performance, similar to those in the experimental data. We set α = 0.2 and β = 1.

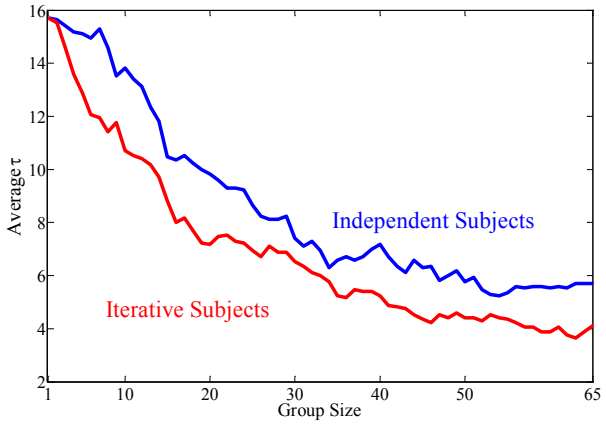

Figure 6. Aggregate Subject Performance Averaged over Questions for the Independent and Iterative Models.

Inference for all of the latent parameters in the iterative model was achieved using a Markov Chain Monte Carlo sampling procedure. We ran 1 chain for each subject, with 110 iterations and a burnin of 10.

Results

Figure 6 shows the aggregate performance for 65 simulated independent and iterative subjects for 17 questions, generated by the respective generative models. As in the empirical data, the iterative subjects outperform their independent counterparts in aggregate (albeit with a lower asymptote). The concept of our model, that each subject combines their own internal representation of the ordering with the previous ordering, based on some sense of how uncertain they are about their values for each item, seems to provide a good fit to the data, and seems to suggest that this is the strategy that participants are employing.

Conclusion

In this paper, we have shown that iterative communication between subjects can lead to better estimates in reconstructing the ground truth for rank ordering tasks, both individually and in the aggregate. We suggested that this is due to different item-level knowledge among subjects, and that subjects are integrating the given information from previous subjects with their own knowledge in a manner that preserves and adds knowledge. We developed a generative model for iterative subjects in a Thurstonian framework that utilized this principle, and demonstrated that simulated iterative subjects were able to out-perform simulated independent subjects in a way that was qualitatively similar to our empirical data. The model seems to provide a good qualitative account of the updating of individual knowledge with the communicated knowledge of other individuals. Further refinements to this model, including using a 2-stage generative process to introduce group-level uncertainty in the ground truth subjects have access to, will be necessary to make a better match between simulated and empirical data.

References

Beppu, A., & Griffiths, T. L. (2009). Iterated learning and the cultural ratchet. *Proceedings of the Annual Conference of the Cognitive Science Society* (pp. 2089-2094).

Bruce, R. (1935). Group Judgments in the Fields of Lifted Weights and Visual Discrimination. *The Journal of Psychology*, 1, 117-121.

Dani, V., Madani, O., Pennock, D.M., Sanghai, S.K., & Galebach, B. (2006). An Empirical Comparison of Algorithms for Aggregating Expert Predictions. *Proceedings of the Conference on Uncertainty in Artificial Intelligence* (UAI).

Gallupe, R. B., Bastianutti, L. M., & Cooper, W. H. (1994). Unblocking brainstorms. *Journal of Applied Psychology*, 76, 1, 137-142.

Galton, F. (1907). Vox Populi. *Nature*, 75, 450-451.

Kerr, N. L., & Tindale, R. S. (2004). Group performance and decision making. *Annual Review of Psychology*, 55(1), 623-655.

Kirby, S. (2001). Spontaneous evolution of linguistic structure: An iterated learning model of the emergence of regularity and irregularity. *IEEE Journal of Evolutionary Computation*, 5, 102-110.

Koller, D., Friedman, N., Getoor, L., & Taskar, B. (2007). Graphical models in a nutshell. In L. Getoor & B. Taskar (Eds.), *Introduction to statistical relational learning*. Cambridge, MA: MIT Press.

Lorge, I, Fox, D., Davitz, J., & Brenner, M., (1957). A Survey of Studies Contrasting the Quality of Group Performance and Individual Performance. *Psychological Bulletin*, 55, 337-372.

Miller, B., Hemmer, P., Steyvers, M., & Lee, M. D. (2009) The wisdom of crowds in rank ordering tasks. *Proceedings of the 9th International Conference of Cognitive Modeling*.

Rowe, G., & Wright, G. (1999). The Delphi technique as a forecasting tool: issues and analysis. *International Journal of Forecasting*, 15(4) 353-375.

Silver, N. (2010). FiveThirtyEight: Nate Silver's Political Calculus. http://fivethirtyeight.blogs.nytimes.com/about-fivethirtyeight/. Accessed January 2011.

Sjöberg, L. (2009). Are all crowds equally wise? A comparison of political election forecasts by experts and the public. *Journal of Forecasting*, 28(1), 1-18.

Steyvers, M., Lee, M. D., Miller, B., & Hemmer, P. (2009). The wisdom of crowds in the recollection of order information. In Y. Bengio, D. Schuurmans, J. Lafferty, C. K. I. Williams, & A. Culotta (Eds.), *Advances in Neural Information Processing Systems*, 22 (pp. 1785–1793).

Surowiecki, J. (2004). *The Wisdom of Crowds*. New York, NY: W. W. Norton & Company, Inc.

Vul, E., & Pashler, H. (2008). Measuring the crowd within: Probabilistic representations within individuals. *Psychological Science*, 19(7), 645-647.

Yaniv, I. (1997). Weighting and trimming: Heuristics for Aggregating Judgments under Uncertainty. *Organizational behavior and human decision processes*, *69*(3), 237-249.