The Wisdom of Crowds with Informative Priors

Pernille Hemmer (phemmer@uci.edu)
Mark Steyvers (msteyver@uci.edu)
Brent Miller (brentm@uci.edu)

Department of Cognitive Sciences,
University of California, Irvine
Irvine, CA, 92697-5100

Abstract

In some eyewitness situations, a group of individuals might have witnessed the same sequence of events. We consider the problem of aggregating eyewitness testimony, trying to reconstruct the true sequence of events as best as possible. We introduce a Bayesian model which incorporates individual differences in memory ability, as well as informative prior knowledge about event sequences, as measured in a separate experiment. We show how adding prior knowledge leads to improved model reconstructions, especially in small groups of error-prone individuals. This Bayesian aggregation model also leads to a "wisdom of crowds" effect, where the model's reconstruction is as good as some of the best individuals in the group.

Keywords: Eyewitness Testimony; Wisdom of Crowds; Rank Ordering; Bayesian Modeling; Serial Recall.

Introduction

Studies of eyewitness testimony have shown that human memory can be incomplete and unreliable (e.g., Loftus, 1975). In real world situations, there might be multiple eyewitnesses, all of whom witnessed the same set of events. This raises the possibility of recovering the true account of events by analyzing the similarities in the recalled memories across individuals. Different individuals might also recall different aspects of the events, such that an aggregate narrative, based on the group's memory, would be closer to the true sequence of events than that of any one individual. An investigator might try to manually reconstruct the aggregate narrative, or witnesses might be allowed to discuss the events in order to develop the group narrative. Communication between witnesses however, has been shown to lead to much worse performance (Gagnon and Dixon, 2008), and humans have been shown to be inconsistent in assessing group information from multiple sources (Stasser & Titus, 1985). To avoid these problems, we propose a model of aggregation that can integrate the recalled memories from a number of independent individuals , while also taking in other important factors, such as individual differences and prior knowledge, into account.

Research on the "Wisdom of Crowds" (WoC) has shown that an aggregation of independent judgments often leads to a group estimate that is closer to the ground truth than that of most of the individuals (Surowiecki, 2004). These group estimates are often simply found by taking the mean, median, or mode of responses (Galton, 1907; Surowiecki, 2004). Much of the previous literature on aggregation of judgments has focused on tasks where individuals estimate numerical quantities and probabilities (Budescu, Yu, 2007; Hogarth, 1978; Wallsten, Budescu, Erev, & Diederich, 1997). It is, however, often that case that eyewitness have to retrieve information more complex than single numerical estimates.

The WoC effect can also be demonstrated with more complex problem sets. For example, the WoC effect has been demonstrated with solutions to problem-solving situations such as finding minimum spanning trees for a set of nodes (Yi, Steyvers, Lee & Dry, in press). Steyvers, Lee, Miller, and Hemmer (2009) showed that order information from semantic memory can also be combined across individuals to give high accuracy in reconstructing the true order of items along some physical or temporal dimension; when individuals recalled the order of US presidents, or the order of rivers according to length, many of the individual orderings were error-prone, but the aggregate orderings were more accurate, on average. In Steyvers et al. (2009), a number of aggregation models for order information were tested. It was found that using Bayesian models that incorporated psychologically plausible representations, cognitive processes and individual differences outperformed basic heuristic aggregation approaches, such as taking the mode..

When errors across indivduals are uncorrelated (as they tend to be when individuals independently give their judgments) the errors will cancel out in the aggregate. Therefore, one expects the best results in WoC experiments with a large number of individuals. In eyewitness situations however, there is rarely a "crowd" available to witness the same set of events. In these cases, we have to rely on a small number of individuals (in many cases, just one) and significant errors might not cancel. Therefore, it might not be sufficient to just analyze the commonalities across the witness reports. We propose that it is better to combine the witness reports along with prior knowledge about the particular event sequence. Combining prior knowledge with noisy information has been shown in other domains to improve the recovered estimate (Hemmer & Steyvers, 2008; Konkle & Oliva, 2007; Kan, Alexander, Verfaelle, 2009).

We focus in this research on the problem of reconstructing event sequences. The goal is to reconstruct

the true ordering of a set of events by aggregating the recalled orderings from a small number of individuals, all of whom witnessed the same event sequence. The novelty of the current approach is that we incorporate informative prior knowledge in an aggregation model for order information in order to improve the aggregate estimate. This is especially helpful when aggregating across a small number of error-prone individuals.

We present our results as follows. We first report on behavioral experiments wherein we tested people's ability to reconstruct, from episodic memory, the order of stereotyped events (e.g., getting up in the morning), or random events (e.g., clay animation without a clear story line). We also report on experiments where we measured prior knowledge for the same set of events. We then describe a Bayesian approach that aggregates the orderings across individuals while taking prior knowledge into account.

Empirical Study on Serial Recall

Much research on serial recall has been done on random word and letter sequences that do not have any obvious organization. In such experiments, individuals are shown a sequence of words or letters, and the task is to recall the original temporal order as best as possible during a later test. Typical errors in the recalled orderings are transposition errors where the orderings are locally perturbed (Estes, 1997; Nairne, 1992) -- two events nearby in time tend be reconstructed as occurring nearby but the amount of perturbation noise depends on many factors such as time elapsed between study and test, stimulus characteristics and individual differences. Similar patterns have been observed in more naturalistic experiments, such as naming the day of the week an event occurred (Huttenlocher, Hedges, & Prohaska, 1990), as well as for autobiographical memory, such as ordering the events of September 11th (Altmann, 2003). With more naturalistic event sequences, prior knowledge about the event sequences can influence episodic memory. People have clear expectations for routine activities and are sensitive to the ordering of actions within an activity (Bower, Black & Turner, 1979).

We conducted a series of behavioral experiments using two types of event sequences. We used a number of *stereotyped* event sequences, such as getting up in the morning, or jumping on a bus, for which people have clearly defined expectations, and a number of *random* event sequence, such as clay animation sequences or Japanese pizza commercials, for which the temporal organization might be less structured. To assess the prior knowledge people have about these types of events, we first conducted a prior knowledge study where we asked participants to order the events in the most natural order possible without actually showing them the original, true event sequence. This allows us to estimate a model for the prior probability of each sequence.

In a separate experiment, we assessed serial recall for each of event sequences. It should be noted that our definition of serial recall differs from the standard use of the term in that our task only involves ordering the events, not recalling the items to be ordered, as in a standard serial recall task. In our task, we first showed a video of the original event sequence which was followed by a serial recall test in which individuals ordered image stills from the video as best as possible according to the original temporal sequence in which the events appeared. No communication between individuals was allowed in any of our tasks, and therefore the data consists of independent recollections from individuals.

Methods

Participants were undergraduate students at the University of California, Irvine. There were 16 participants in the prior knowledge experiment and 28 participants in the serial

prior

1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
4	3	13	1	9	5	3	12	11	10	10	2	5	1	1	10
A	B	A	A	A	B	B	A	B	B	A	A	D	B	A	B
B	C	C	B	C	D	A	D	A	A	C	B	B	A	B	A
E	A	E	C	D	A	D	G	C	E	E	D	A	C	D	E
D	D	F	D	E	E	C	C	E	C	D	C	C	D	C	F
C	E	G	E	G	C	E	E	D	G	G	E	E	E	E	D
G	G	H	G	F	G	G	F	J	F	F	G	G	F	F	G
F	F	I	F	H	F	F	H	I	H	H	F	F	G	G	H
H	H	D	H	I	H	H	I	H	J	I	H	H	H	H	I
I	I	J	I	J	I	I	J	F	I	J	I	I	I	I	J
J	J	B	J	B	J	J	B	G	D	B	J	J	J	J	C

serial recall

1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28
2	1	2	4	2	5	1	1	1	1	2	1	6	1	1	0	1	2	2	1	2	1	1	1	2	0	1	1
B	A	A	A	B	B	A	A	A	A	A	A	A	A	A	A	A	B	B	A	A	A	B	A	B	A	A	A
A	B	B	B	A	A	B	B	B	B	B	B	G	B	B	B	B	A	A	B	B	B	A	C	A	B	B	B
C	C	C	G	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	B	C	C	C	C
D	D	D	C	D	D	D	D	D	D	D	D	B	D	D	D	D	E	D	D	D	D	D	D	D	E	D	
E	E	G	D	E	H	E	E	F	F	F	E	D	F	E	E	E	D	E	E	F	E	E	E	E	D	E	
G	G	E	E	G	F	G	G	E	E	G	G	E	E	G	F	G	F	G	G	G	G	F	F	G	F	F	G
F	F	F	F	F	E	F	F	G	G	E	F	F	G	F	G	F	G	F	F	E	F	G	G	F	G	G	F
H	H	H	H	G	H	H	H	H	H	H	H	H	H	H	H	H	H	H	H	H	H	H	H	H	H	H	H
I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I
J	J	J	J	J	J	J	J	J	J	J	J	J	J	J	J	J	J	J	J	J	J	J	J	J	J	J	J

Figure 1. The sequence A-J shows the 10 images from the 'bus' video sequence in the correct temporal order. The two tables show the participant orderings in the prior knowledge and serial recall experiment. The first row is the participant id. The second row is the Kendall's tau distance between the true ordering and the recalled order for that participant.

recall experiment.

Materials. We sampled 6 videos from YouTube.com. Three videos depicted stereotyped events sequences (getting up in the morning, a wedding, getting on the school bus). Three videos depicted more random event sequences (a Japanese yogurt commercial, a Japanese pizza commercial, and a clay animation sequence). For each of the 6 videos 10 still images of individual scenes were drawn. See Figure 1 for an example.

Prior Knowledge Experiment. Participants were shown 10 image stills from a given event sequence (e.g., Wedding) and asked to order the 10 images based on their prior expectation of how the event in the slides might unfold. Importantly, in this experiment, participants were never shown the original video sequence from which the image stills were drawn. They responded using an interactive interface in which the images were randomly ordered on the screen and the instruction was to order the images in any way to make the sequence as natural as possible.

Serial Recall Experiment. Participants first viewed the original video sequence. Participants were then presented with the same interface as in the prior knowledge experiment. They were shown 10 image stills that they had to order in the original temporal order. For both the prior knowledge and memory experiment, the initial ordering of the 10 image stills, as well as the order of the 6 video sequences, was randomized across participants.

Results and Discussion

To evaluate the performance of participants, we measured the distance between the reconstructed and the correct ordering. A commonly used distance metric for orderings is Kendall's τ (Marden, 1995). This distance metric is the minimum number of adjacent pairwise swaps necessary to resolve any disagreements between the two orderings being compared. Values of τ range from $0 \leq \tau \leq (N-1)/2$, where N is the number of items in the order: N=10 for all of our event sequences. In our experiment, a τ=0 indicates that the participant responded with the exact correct ordering. A τ=1 indicates that one adjacent pair of items was swapped. When participants are using a random guessing strategy, their expected mean expected distance is $\tau = (N-1)/4 = 22.5$.

Figure 1 shows the raw data collected for the "bus" video sequence – a stereotyped event sequence. In the prior knowledge experiment, participants produced orderings that were much better than chance, suggesting that a priori, it is possible to guess the true ordering of events in these types of event sequences. In the memory experiment, 2 participants produced the correct ordering, and 15 more were within one swap of the true order. Note that very few identical orderings are produced between participants. We found that for all 3 random events, in both the prior knowledge experiment and the memory experiment, each participant produced a unique ordering. For the 3 stereotyped event sequences however, only one sequence led to unique orderings across all participants.

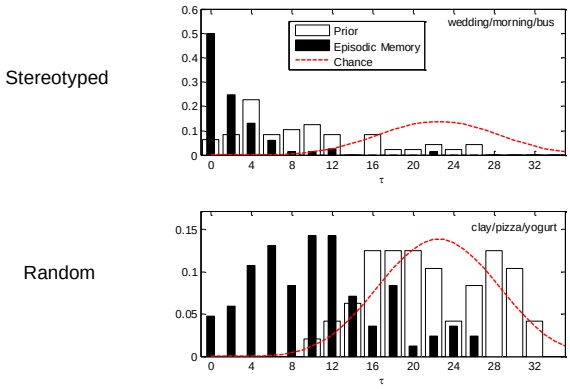

Figure 2. Distributions of Kendall τ distances.

Figure 2 shows the distributions of the Kendall τ distances for the serial recall and prior knowledge experiment. The top panel shows the distances for stereotyped event sequences and the bottom panel shows the distances for random event sequences. The dashed line shows the distribution of distances that can be expected from a random guessing strategy (this distribution can be calculated exactly, see Marden, 1995). For both the stereotyped and random event sequences, the distances are lower for the memory task than for the prior knowledge task. The distances are also lower for the stereotyped event sequences than for the random event sequences. Even when participants did not study the videos (the prior knowledge condition), they performed better than chance in the stereotyped condition, as compared to the random condition where prior knowledge performance led to a distribution of distances very similar to distances expected from chance performance. These results demonstrate that general knowledge about events can greatly contribute to the accuracy of recalling these events.

Modeling

We can conclude from our empirical study that prior knowledge can lead to improved average performance in recall. When ordering scenes from an event with strong prior expectations, the resulting orderings are relatively close to the true ordering. Of course, performance improves on average after observing the true event sequence and later recalling the sequence from memory. This raises the question of how one might incorporate an informative prior in a model for aggregating rank-ordered recall. Such priors might guard against errors from a small number of poorly performing individuals. In this paper, we explore very simple models to aggregate the orderings of individuals. The goal of the modeling is not to build a comprehensive model of recall that specifies all the representations and processes involved in storing and retrieving information from memory. Instead, we will focus on simple probabilistic models such as a Mallows model (e.g. Steyvers et al., 2009) that allow us to aggregate the retrieved orderings from a number of individuals using Bayesian inference. The current model incorporates two important differences to the

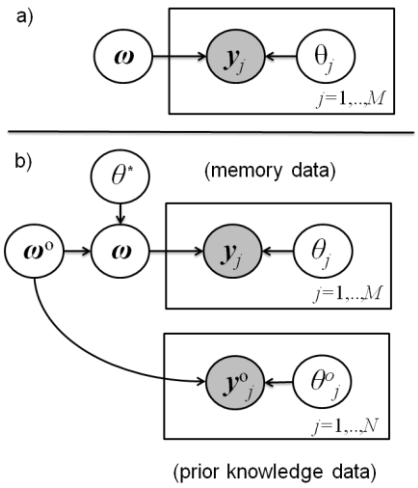

Figure 3. The graphical model representations for the Mallows model with an uninformative prior (a) and an informative prior about the group knowledge (b).

previous work by Steyvers et al. (2009). First, we generalize the model to allow for individual differences in memory performance. These individual differences are estimated by the model in a purely unsupervised fashion and do not require knowledge of past performance in other tasks or access to a known ground truth. With the individual differences, the model finds aggregates that are weighted towards solutions provided by the individuals that are estimated to have good memory performance.

Second, we develop a simple extension of Mallows models that allows for informative priors. This prior is estimated from the orderings produced in the prior knowledge experiment.

Mallows Model with an Uninformative Prior

In a basic Mallows model (Marden, 1995), all individuals are assumed to derive their orderings from a single underlying ordering, that we will refer to as the *group knowledge*. The group knowledge is a latent variable in the model that can be estimated from the data. Importantly, Mallows model assumes that each individual produces orderings centered on the group ordering with distant orderings less likely than orderings close to the group ordering. Although Mallows-type models have often been used to analyze preference rankings (Marden, 1995), they have not been applied, as far as we are aware, to ordering data from serial recall experiments. In our first extension of the standard model we allow for individual differences in memory performance. We evaluated this aggregation model by comparing the estimated group ordering to the ground truth. If the model is able to tap into the collective wisdom of a group of individuals, the estimated group ordering should be close to the true ordering.

Specifically, let $\boldsymbol{y}_j$ represent the ordering from individual j, and $\boldsymbol{\omega}$ the latent group ordering. In a Mallows model, the

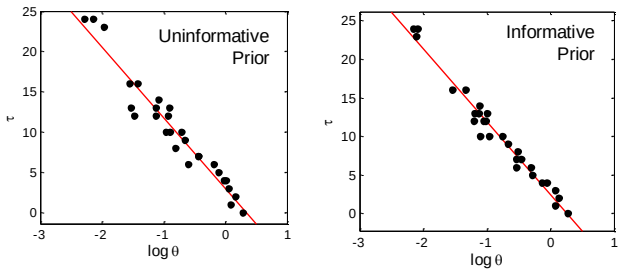

Figure 4. Calibration results for the two models for one event sequence.

probability of each individual ordering given the group ordering is given by

$$p(\boldsymbol{y}_j \,|\, \boldsymbol{\omega}, \theta_j) \propto e^{-\mathrm{d}(\boldsymbol{y}_j, \boldsymbol{\omega})\theta_j} \quad (1)$$

where for simplicity we have omitted the normalization constant. The function d returns the Kendall τ distance between two orderings. The scaling parameter θ_j determines how close the observed order for individual j is to the group ordering. It can be interpreted as an individual (inverse) noise parameter -- good individuals tend to closer to the group consensus (high θ) whereas poor performing individuals return more idiosyncratic orderings further away from the group knowledge (low θ). We will assume a Gamma prior on the individual noise levels: $\theta_j \sim \mathrm{Gamma}(\theta_0\lambda, 1/\lambda)$, where λ is a hyperparameter that sets the overall level of cohesion expected from the group. Notably, in this first model, we have assumed a uniform prior over group orderings, $\boldsymbol{\omega} \sim \mathrm{Uniform}(\Omega)$, where Ω is the set of all orderings. Therefore, a priori, the model assumes no preference for a particular group ordering.

Figure 3, panel a, shows a graphical representation of the model. Shaded nodes represent observed variables while nodes without shading represent latent variables. The arrows indicate the conditional dependencies between the variables and the plate represents the repeated sampling steps across M subjects in the memory experiment.

Mallows Model with an Informative Prior

We now introduce a simple variant of this model that allows for an informative prior. The idea is that the group knowledge is itself sampled from a Mallows model:

$$p(\boldsymbol{\omega} | \boldsymbol{\omega^0}, \theta^*) \propto e^{-\mathrm{d}(\boldsymbol{\omega}, \boldsymbol{\omega^0})\theta^*} \quad (2)$$

where $\boldsymbol{\omega^0}$ is the prior ordering from which the group ordering is derived, and θ^* is a scaling parameter. This prior stage in Mallows model at first might not seem to gain any additional information because it is not clear how the prior ordering can be constrained. However, we have data in the prior knowledge experiment in which N participants tell us what orderings they expect from certain scenes. Let $\boldsymbol{y}_j^0$ represent the prior ordering given by individual j in the prior knowledge experiment. We assume that these are produced by a Mallows model with $\boldsymbol{\omega^0}$ as the "center":

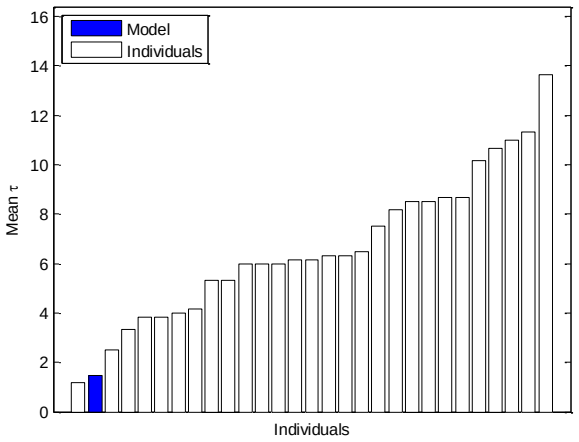

Figure 5. Performance of individuals and model (with informative prior) averaged over six event sequences.

$$p\left(\boldsymbol{y}_j^0 \middle| \boldsymbol{\omega^0}, \theta_j^0\right) \propto e^{-\mathrm{d}\left(\boldsymbol{y}_j^0, \boldsymbol{\omega^0}\right)\theta_j^0} \quad (3)$$

Figure 3, panel b, shows the corresponding graphical model. With this model, we are setting a prior on the group ordering -- when there is only data available from a few individual in the memory experiment, the group ordering will be influenced by the data from the prior knowledge experiment leading to group orderings that are a priori deemed likely. When data from more individual becomes available in the memory experiment, the prior knowledge data will have a diminishing influence on the group ordering which will be mostly determined by the memory data.

Modeling Results

All latent variables in the model were estimated using a MCMC procedure, separately for each event sequence. The result of the inference procedure is a probability distribution over group orderings, of which we take the mode as the single answer for a particular problem. Note that the inferred group ordering does not have to correspond to an ordering of any particular individual. The model just finds the ordering that is close to all of the observed memory orderings.

Figure 4 shows the calibration for the two models on a single event sequence (the clay animation video). Each panel shows the relationship between the inferred θ (related to the distance of each individual to the group ordering) and the Kendall's τ distance of the individual's answer to the ground truth. The plots show that individuals who are close to the group ordering tend to be closer to the ground truth. This means that the models can calibrate the performance levels of individuals, even in the absence of any explicit feedback or access to the ground truth.

Figure 5 shows the Kendall's τ distance for each individual in the memory experiment averaged over the six event sequences. Note that there are substantial individual differences with some individuals coming relatively close to the ground truth. The figure also shows the average model performance. Comparison between individual and model performance reveals a WoC effect: The model performs as well as some of the best individuals, with only one individual outperforming the model. Therefore, we can conclude there is a weak WoC effect (a strong WoC effect would correspond to a situation where the model outperforms all individuals in the group).

We now focus on applying the model to subsets of participants to mimic eyewitness situations that typically involve only small number of individuals. In the first analysis, we select a random set of *K* individuals from the original set of 28 individuals. We then apply the two models to the subset of individuals. Figure 6 shows model results for the model with the informative and uninformative prior separated for stereotyped and random event sequences. For random event sequences, where the prior is weak, there is no improvement in the aggregation between the two models (if anything, there is a small performance decrement for the model with the informative prior). For stereotyped event sequences however, people have strong prior expectations about the true ordering of events and there is a marked improvement in the aggregate response in the model with the informative prior. This improvement is most pronounced with low sample sizes (*K*=1 and *K*=2) when the prior can still exert an influence on the inferred group orderings. Note that when *K*=1, the model with the uninformative prior has no information other than the ordering given by a single individual – therefore, the aggregate solution given by the model is equivalent to the ordering provided by the individual. This results in an average tau of around 15. However, performance for the model with the informative prior is much better resulting in a tau of around 8, because the aggregate solution combines the single remembered ordering with the a priori likely orderings.

To better highlight the benefit of the prior information, we also conducted a model analysis where we selected the *worst* performing individuals in the sample. In this sampling procedure, we sample the *K* worst individuals where we vary *K* from 1 (the single worst performing individual) to 28 (all individuals combined). Figure 7 shows model results for both models separated for stereotyped and random event sequences. The relative performance benefits can be seen most clearly for the stereotyped event sequences for low sample sizes (*K*=1 and *K*=2). In these cases, the worst individuals recall event sequences that are a priori unlikely and the prior "corrects for" the noise in the available data.

Therefore, these analyses suggest that an aggregation model with informative priors can be used to guard against the most egregious errors committed by the worst individuals in the memory task.

Conclusions

We have presented two approaches for aggregating recalled sequences of events in order to reconstruct the true event sequence as best as possible. Individuals are likely to differ in their ability to recall event sequences and pay attention to different parts on an event sequences. Therefore, by

analyzing the consistencies in orderings across individuals, we can extract the collective wisdom in the group. We presented two aggregation approaches based on Mallows model that allow for individual differences. The models combine information at the group level with information at the individual level to explain orderings given by an individual. In the first approach, the model uses only the data from the individuals who all witnessed an event sequence. In the second approach, the model uses an additional source of data based on the prior knowledge about the events extracted from another group of individuals.

We demonstrated a weak WoC effect, where the average performance of the model was better than every individual, save one. We have also shown that a Mallows model with informative priors has a markedly improved ability to reconstruct the ground truth in cases where the event sequences are highly stereotyped and a small sample of poorly performing individuals is used for aggregation. This is particularly important in eyewitness situations where we typically have only a small number of individuals available.

References

Altmann, E. M. (2003) Reconstructing the serial order of events: A case study of September 11, 2001. *Applied Cognitive Psychology*, **17**, 1067-1080.

Bower, G. H., Black, J. B., & Turner, T. J. (1979). Scripts in memory. *Cognitive Psychology*, **11**, 177-220.

Budescu, D. V. & Yu, H. (2007). Aggregation of opinions based on correlated cues and advisors. Journal of Behavioral Decision Making, 20, 153-177.

Estes, W.K. (1997). Processes of Memory Loss, Recovery, and Distortion. Psychological Review, 104, 148-169.

Gagon, L.M. & Dixon, R.A. (2008). Remembering and retelling stories in individual and collaborative contexts. *Applied Cognitive Psychology*, *22*, 1275-1297.

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

Hemmer, P. & Steyvers, M. (2008). A Bayesian Account of Reconstructive Memory. In V. Sloutsky, B. Love, and K. McRae (Eds.) *Proceedings of the 30th Annual Conference of the Cognitive Science Society*.

Hogarth, R. M. (1978). A note on aggregating opinions. Organizational Behavior and Human Performance, 21(1), 40-46.

Huttenlocher, J., Hedges, L. V., & Prohaska, V. (1992). Memory for day of the week: A 5+2 day cycle. *Journal of Experimental Psychology: General*, *121*, 313–325.

Kan, I.P., Alexander, M.P. & Verfaellie, M. (2009). Contribution of prior semantic knowledge to new episodic learning in amnesia. *Journal of Cognitive Neuroscience*, *21*, 938-944.

Konkle, T., & Oliva, A. (2007). Normative representation of objects: Evidence for an ecological bias in perception and memory. In D. S. McNamara & J. G. Trafton (Eds.), *Proc.s of the 29th Annual Cognitive Science Society, (pp. 407-413), Austin, TX: Cognitive Science Society.*

Loftus, E.F. (1975). Leading questions and the eyewitness report. *Cognitive Psychology*, *7*, 560-572.

Marden, J. I. (1995). Analyzing and Modeling Rank Data. New York, NY: Chapman & Hall USA.

Nairne, J. S. (1992). The loss of positional certainty in long-term memory. *Psychological Science*, **3**, 199–202.

Stasser, G., Titus, W. (1985). Pooling of unshared information in group decision making: Biased information sampling during discussion. *Journal of personality and social psychology*, 48(6), 1467-1478.

Steyvers, M., Lee, M.D., Miller, B., & Hemmer, P. (2009). The Wisdom of Crowds in the Recollection of Order Information. In J. Lafferty, C. Williams (Eds.) *Advances in Neural Information Processing Systems, 23*. MIT Press.

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

Wallsten, T.S., Budescu, D.V., Erev, I. & Diederich, A. (1997). Evaluating and combining subjective probability estimates. Journal of Behavioral Decision Making, 10, 243-268.

Yi, S. K. M., Steyvers, M., Lee, M. D., Dry, M. J. (in press) Wisdom of the Crowds in Minimum Spanning Tree Problems. Proceedings of the 32nd Annual Conference of the Cognitive Science Society.

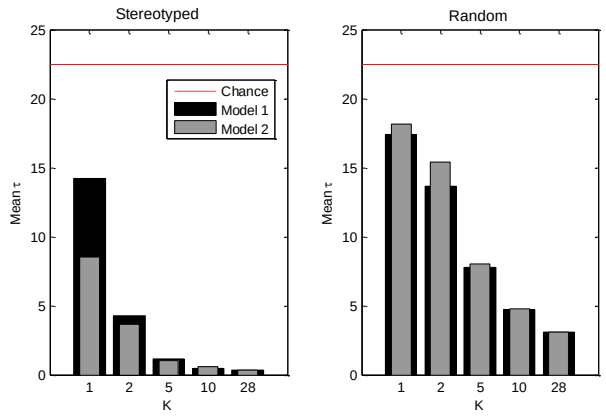

Figure 6. Results from the models with an uninformative prior (model 1) and informative prior (model 2) for *random* subsets of *K* individuals from the memory task.

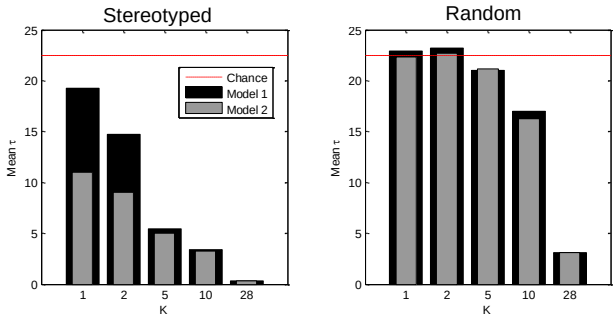

Figure 7. Results from the models with an uninformative prior (model 1) and informative prior (model 2) for subsets of the *worst K* individuals from the memory task.