Modeling the effect of chained study in transitive inference

Douglas B. Markant (dmarkant@uncc.edu)
Department of Psychological Science
University of North Carolina at Charlotte
9201 University City Blvd., Charlotte, NC 28223 USA

Abstract

A hallmark of human memory is the ability to integrate discrete experiences into cognitive maps. A fundamental form of this integration is transitive inference (TI), in which overlapping premises (e.g., $A < B$, $B < C$) are integrated into a unified representation of a relational hierarchy ($A < B < C$). Few existing theories provide a mechanistic account of this construction of relational knowledge and how it is shaped by different training conditions. This study builds on recent findings that TI is facilitated by chaining of overlapping premises, with a new behavioral experiment confirming an advantage over non-overlapping sequences matched for premise frequency and spacing. A subsequent simulation study shows that the chaining effect is captured by a particle filter which performs approximate Bayesian inference about the latent hierarchy. These results provide a better understanding of how chaining shapes the construction of relational knowledge in the face of uncertainty and forgetting.

Keywords: transitive inference; relational learning; particle filter

People have a remarkable ability to organize separate, but related, experiences into internal cognitive maps, which then support flexible inferences about relationships that have not been directly experienced, as when taking a shortcut for the first time (Peer, Brunec, Newcombe, & Epstein, 2021) or making predictions about unobserved links in a social network (Son, Bhandari, & FeldmanHall, 2021). Transitive inference (TI) is a fundamental form of this ability applied to items organized in a linear hierarchy (e.g., $A < B < C$). In TI people learn the relations between adjacent items by encoding *premise pairs* (e.g., $A < B$, $B < C$) and are tested on the ability to make transitive inferences about novel pairs (e.g., to infer that $A < C$). The capacity for TI is present at a young age (Bryant & Trabasso, 1971) and is evident in a wide range of species (Vasconcelos, 2008), highlighting its important role in extracting relational knowledge from related experiences.

There is ongoing debate over the cognitive mechanisms that are involved in TI, with numerous theories based on simple forms of associative or reward-driven learning (Frank, Rudy, & O'Reilly, 2003; Wynne, 1995) as well as retrieval-based inference at the time of test (Kumaran & McClelland, 2012). While multiple cognitive processes may independently support TI depending on the nature of the task at hand, a growing body of evidence suggests that human learners rely on a constructive process during learning, integrating premises into a unified mental map of the hierarchy (Hummel & Holyoak, 2001; Jensen et al., 2015).

Evidence for relational integration during learning comes from findings that performance improves when training involves *chained* sequences of overlapping premises across trials (see examples in Figure 1B). In a card sorting TI task where people rank the items as they are presented with premises, chaining leads to more accurate solutions compared to random sequences among preschool children (Andrews & Halford, 1998; Halford, 1984) and adults with deficits in relational reasoning due to prefrontal lobe damage (Waltz et al., 1999) or dementia (Waltz et al., 2004). Chaining is thought to reduce the difficulty of relational integration in these tasks because they allow learners to combine new information with a single (immediately preceding) premise, making it easy to incrementally build a cognitive map of the hierarchy (see also Foos, Sabol, Smith, & Mynatt, 1976).

Most prior investigations of chained study in TI have involved external representations of the hierarchy (i.e., arranging cards) and small numbers of premises, leaving it unclear how premise order affects learning of more complex hierarchies over extended periods of training. However, recent results from Markant (2020) suggest that chaining is beneficial in a more complex task with 9 item hierarchies. In that study people preferred to chain premises when they had control over the training sequence, and the sequences they generated were more effective than random sequences when presented to other learners. Although these results suggest that chaining supports relational integration, the effects might also be attributable to other differences in the training sequences (including the relative frequency and spacing of premises from different parts of the hierarchy), highlighting the need for a more controlled examination of the effects of chaining in TI.

Current study

The present study had two main goals. The first goal was to directly compare training sequences that were matched for premise frequency and spacing but differed in whether premises overlapped across trials. Participants completed a TI task in which they studied relations between adjacent individuals in a social hierarchy. They were then tested on their ability to reconstruct the hierarchy and their accuracy on a TI test. The experimental manipulation concerned the order of premises during training: In the Chains con-

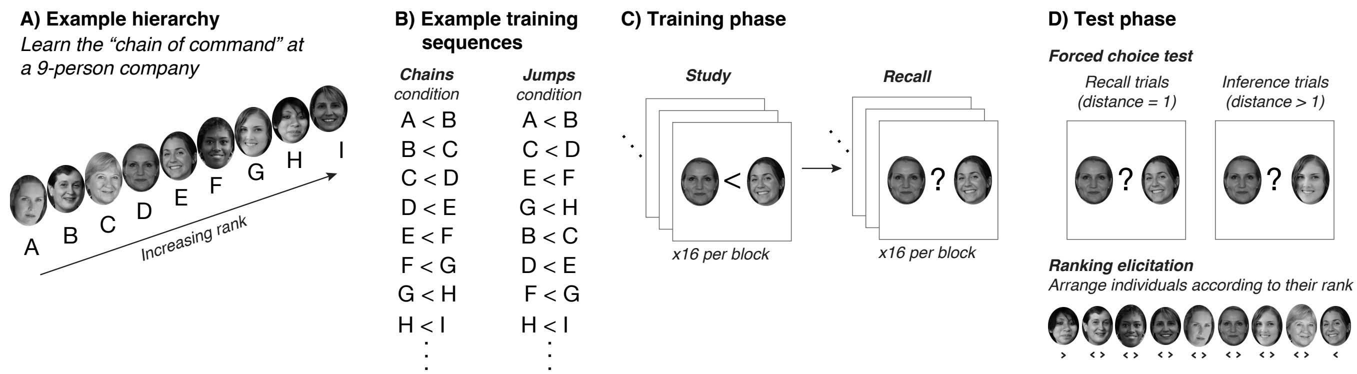

Figure 1: Task design.

dition the premises overlapped from trial to trial, while in the Jumps condition the premises were presented with the same frequency and spacing but never overlapped in successive trials. Based on the results of Markant (2020) I predicted that chained sequences would facilitate the integration of premises into a unified representation, leading to more accurate knowledge of the hierarchy.

The second goal of the study was to examine whether computational models of TI can account for the advantage of chained study. I present a simulation study comparing two models from Kumaran, Banino, Blundell, Hassabis, and Dayan (2016) which formalize the TI problem as estimating the positions of individuals along a latent, continuous dimension. The principal aim of the simulation study was to determine whether these models can account for observed differences in learning of the hierarchy without assuming any other differences between conditions.

Experiment

The experiment was conducted online in two sessions separated by approximately 5 days. The present paper will focus on the results of the first session only, which included the training and test phase of the TI task (Figure 1).

Participants

Students were recruited via an email announcement. $N = 44$ people completed the study (age $M = 23.36$ years, SD = 4.88; 59% female, 27% male, 14% no sex indicated). Participants received $8 for successful completion of both sessions, plus an additional bonus of up to $4 based on performance in the test phase (average bonus of $2.71, $SD = 1.69$). The first session took an average of 26 minutes ($SD = 9.5$).

Materials and Procedure

Participants learned about a 9-item social hierarchy made up of individuals represented by face images (Figure 1A) drawn from the 10k Adult Faces Database (Bainbridge, Isola, & Oliva, 2013). Participants were instructed to the learn the "chain of command" at a fictional company by memorizing the relationships between adjacent pairs of individuals (e.g., that person A is directly supervised by person B). The instructions included an example of a transitive inference across two pairs of individuals who did not appear later in the task. Participants were therefore fully informed about the nature of the underlying hierarchy and the learning goal.

Training sequences. The experimental manipulation determined the order of presentation of the 8 premises (Figure 1B). In the Chains condition sequences were composed of overlapping chains of premises (e.g., $A < B$, $B < C$, $C < D$...), while in the Jumps condition there was no overlap between premises in successive trials. The training sequences in the two conditions were otherwise matched for the relative frequency of premises (with all premises presented equally often) and the spacing of repeated presentations of the same premise. In both conditions the sequence of 8 premises was presented twice in each training block and the direction of the sequences (forward or backward through the hierarchy) alternated across blocks, with the direction in the first block randomized for each participant.

Training phase. The training phase had up to 10 blocks. Each block began with 16 study trials in which premises were individually displayed for 2.5 s (Figure 1C). This was followed by 16 recall trials (two trials for each premise pair) presented in random order. Each premise was displayed and participants were instructed to click on the person who was higher ranked. No feedback was provided until the end of the block, at which point participants were told the proportion of correct responses.

Participants had to complete a minimum of 3 training blocks. The training phase ended either after 10 blocks or when participants reached a criterion of 100% correct responses in a block, indicating that they chose the higher-ranked individual for every premise pair twice.

Test phase. The training phase was followed by a brief distractor task in which participants solved a set of arithmetic problems. They then completed the test phase which included a standard forced choice test and a ranking elicitation. There were 72 trials in the forced choice test, with two repetitions of every possible pairing of items from the hier-

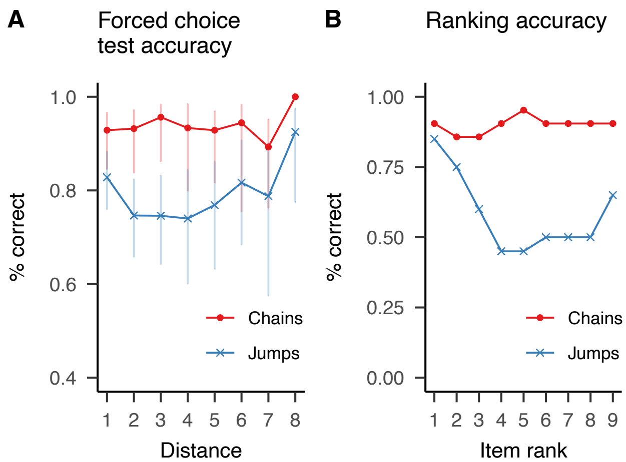

Figure 2: *A:* Accuracy on the forced choice test as a function of the distance between the items in the hierarchy. Error bars indicate bootstrapped 95% CIs. *B:* Proportion of participants who ranked items correctly by their actual rank.

archy (Figure 1D). *Recall trials* involved premises that were experienced during the study phase, whereas *inference trials* involved novel pairings of non-adjacent items. Test trials were presented in the same manner as the recall trials from the training phase, with participants instructed to select the person who was ranked higher in each test pair. No feedback was presented during the test phase. Following the forced choice test, participants ranked the nine individuals according to their positions in the hierarchy. The images were displayed in a random order and the position of each item could be changed by clicking on arrow buttons.

Results

Three participants were excluded for failing to reach the training criterion, leaving $N = 41$ (21 in the Chains condition, 20 in the Jumps condition). The conditions did not differ in the number of blocks to criterion (Chains: $M = 4.48$, $SD = 2.27$; Jumps: $M = 5.4$, $SD = 1.73$; $F(1,39) = 2.13$, $p = .15$).

Forced choice test accuracy. Responses on the forced-choice test were scored according to whether the higher-ranked individual in each test pair was chosen and were modeled using mixed effects logistic regression with trial type (recall vs. inference) and study condition (Chains vs. Jumps) as fixed effects and random intercepts for participants. An analysis of deviance indicated a significant effect of study condition ($\chi^2(1) = 12.28$, $p < .001$). There was no overall effect of test type ($\chi^2(1) = 3.54$, $p = .06$) but there was a significant interaction ($\chi^2(1) = 4.10$, $p = .04$). Accuracy in the Chains condition was significantly higher for both recall trials (Chains: $M = .93$, $SD = .13$; Jumps: $M = .83$, $SD = .15$; $OR = 5.78$, 95% CI [1.46, 22.90], $z = 2.50$, $p = .01$) and inference trials (Chains: $M = .94$, $SD = .14$; Jumps: $M = .77$, $SD = .22$; $OR = 11.59$ [3.20, 42.0], $z = 3.73$, $p < .001$).

Figure 2A shows accuracy as a function of distance between the individuals in a test pair. Performance in the Chains condition was very high, with many participants achieving perfect accuracy on inferences at all distances. In both conditions there was also evidence of a symbolic distance effect, such that performance on inference trials improved with greater distances between the items.

Ranking accuracy. Overall ranking accuracy was calculated as the proportion of nine items that were ranked in the correct position. Ranking accuracy was higher in the Chains condition ($M = .90$, $SD = .29$) than the Jumps condition ($M = .58$, $SD = .41$; $\chi^2(1) = 50.97$, $p < .001$). Ranking accuracy by position is shown in Figure 2B. In the Chains condition approximately 90% of participants ranked each individual correctly across all positions in the hierarchy. In the Jumps condition ranking accuracy was highest for the endpoints but lower for individuals in the middle of the hierarchy.

Modeling the effect of chained study

The behavioral results demonstrate a clear effect of premise order on learning of the hierarchy, with chained sequences leading to high accuracy on both the forced choice and ranking tests. Training sequences that were matched for premise frequency and spacing but with no overlap between premises led to poorer recall of the premises in the final test, lower accuracy on inference trials, and more errors in the elicited rankings. There was also evidence for a symbolic distance effect (Moyer & Bayer, 1976) such that accuracy was highest for more distant inferences (Figure 2A). These findings strongly suggest that participants performed the task by constructing an integrated cognitive map of the hierarchy as they learned, and that chained study facilitated this process.

In the remainder of the paper I examine whether these effects can be captured by existing computational models of TI. A number of approaches have been taken to model TI, including mechanisms related to associative learning, reinforcement learning, and retrieval-based inference (see Jensen, Terrace, & Ferrera, 2019). A detailed review of these "model-free" approaches are beyond the scope of the present paper, but recent work suggests that they often fail to account for human performance in TI tasks (Jensen et al., 2015). For this reason I focus on two models originally used by Kumaran et al. (2016) in the context of TI: 1) a particle filter and 2) RL-Elo.

Particle filters are a type of sequential Monte Carlo approximation to Bayesian inference (Speekenbrink, 2016) in which a posterior distribution is approximated with a set of discrete samples (particles). A key advantage of this approach is the ability to model how information processing constraints (e.g., limits on time or memory capacity) impact behavior by varying the number of particles N. Particle filters can capture order effects and other deviations from the predictions of optimal Bayesian models in domains such as causal learning (Abbott & Griffiths, 2011), change detection (Brown & Steyvers, 2009; Yi, Steyvers, & Lee, 2009), category learning (Lloyd et al., 2019; Sanborn et al., 2010), and conditioning (Daw & Courville, 2008).

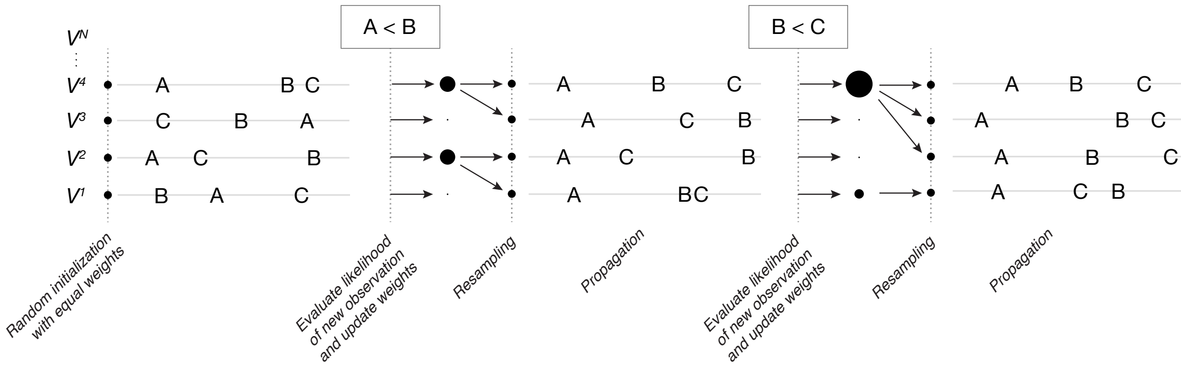

Figure 3: Depiction of the particle filter representation of three items A, B, and C. Each particle serves as a hypothesis about items' positions. When the premise A $<$ B is observed, the weights are updated and used to resample the particles, leading to the proliferation of particles that are consistent with the premise. In the propagation step, the positions are randomly perturbed. After observing B $<$ C, the resulting particle set includes many particles with the correct ordering of the three items.

Kumaran et al. (2016) used a particle filter to model performance in a TI task involving the learning of a social hierarchy. For this task, the model learns items' positions along a latent, continuous dimension, using a limited set of particles to approximate the posterior distribution over positions at each point during training. The authors also compared the particle filter to a model with a similar underlying representation but which uses a learning rule from reinforcement learning (RL-Elo, described in detail below). Behavior was best-described by the particle filter overall, although RL-Elo was found to be a credible alternative during later stages of learning.

In the following I examine whether the effect of chained study can be seen in the performance of the particle filter and RL-Elo models when trained with the same sequences as in the behavioral study. The goal was to explore whether an advantage from chained sequences emerges for either model without assuming any other differences between conditions. Given that participants required an average of 5 blocks (10 repetitions of each premise) to reach the training criterion, I also compared model performance under varying levels of random trial-to-trial fluctuation in the estimates of items' positions in the hierarchy (akin to forgetting).

Particle filter

The particle filter model is based on the bootstrap filter (Doucet, De Freitas, & Gordon, 2001) and largely follows Kumaran et al. (2016). The positions of the nine individuals are represented as a vector of values V along a continuous dimension (see Figure 3). Each particle V^k can be viewed as a different hypothesis about items' positions and is associated with a weight w^k. Particles are randomly initialized with $V_0^k \sim \mathcal{N}(0, \sigma_0)$ and uniform weights $w_0^k = 1/N$ for $k = (1, ..., N)$. For each trial t during training, the particles are reweighted according to the likelihood of the premise observed on that trial, such that $w_t^k = g(y_t|V_{t-1}^k)\, w_{t-1}^k$ (see below for description of the likelihood function g). As a result, those particles for which the underlying values V^k match the order implied by the premises will have greater weight.

Two other common steps in particle filters are resampling and propagation. In the resampling step, particles are resampled with replacement according to the normalized weights and the weights are reset to $w_t^k = 1/N$. In the propagation step, the particles are subjected to random drift according to the transition probability $V_{t+1}^k \sim N(V_t^k, \sigma_d)$. Resampling and propagation allow for the removal of particles that are inconsistent with the observed premises and exploration of alternative solutions. Other applications of particle filters differ in the frequency of resampling and propagation, e.g. by only resampling when the diversity in the particle set drops below a threshold (Kumaran et al., 2016), which may have consequences for the kind of order effects that emerge (Abbott & Griffiths, 2011). For simplicity I assume that both resampling and propagation occur on every trial.

Likelihood function: Local vs. global updating. I explored two variants of the model based on whether the likelihood evaluates local information (limited to the premise observed on the current trial) or global information (accounting for the positions of items that are not observed on the current trial). On each study trial the learner observes a premise $y_t : x_i < x_j$ and evaluates its likelihood under each particle according to:

$$g(y_t|V_t^k) = \alpha \mathbb{1}(V_t^k) + \frac{1}{1 + e^{-\beta(V_{t,j}^k - V_{t,i}^k)}}, \qquad (1)$$

where $\mathbb{1}(V) = 1$ if item x_i is immediately below item x_j (with no intervening items) in the ranking implied by V, and 0 otherwise.

When $\alpha = 0$, the likelihood is simply a sigmoid function of the difference in the positions of items x_i and x_j, with higher likelihood for particles under which x_j has a higher value than x_i, controlled by a scale factor β. This follows Kumaran et al. (2016) and leads to a form of *local updating*, in that the likelihood depends only on the items observed in a given study trial and doesn't explicitly take into account the positions of any other items in the hierarchy.

When $\alpha > 0$, the likelihood is increased by α when x_j is ranked directly above x_i and there are no intervening items. This reflects the fact that participants in the behavioral study were informed that the premises represented adjacent pairs of individuals in the chain of command. Accounting for this constraint should lead to a form of *global updating* where the the latent positions of other items would affect the likelihood even though they are not presented on the current trial.

RL-Elo

The RL-Elo model (Kumaran et al., 2016) was inspired by the Elo rating system for ranking chess players based on the outcomes of pairwise matchups. Like the particle filter, it estimates individuals' positions V on a latent dimension. However, RL-Elo relies on a single point estimate of items' positions, and therefore lacks the representation of uncertainty of the particle filter. RL-Elo also differs in the mechanism for updating estimates through experience, using incremental value updating in response to prediction errors about the higher-ranked individual in each premise.

Values for all items are initialized to $V_0 = 0$ at the start of training. After observing a premise $y_t : x_i < x_j$, the current estimates V_t are used to calculate the probability of item x_j being ranked above x_i:

$$p(y_t|V_t) = \frac{1}{1+e^{-\beta(V_{t,j}-V_{t,i})}}, \qquad (2)$$

which is again a sigmoidal function of the difference between items' values with a scaling parameter β. Values for the higher- and lower-ranked items in the premise are then updated in response to prediction errors as follows:

$$V_{t+1,i} = V_{t,i} + \delta(p(y_t|V_t) - 1) + \varepsilon$$
$$V_{t+1,j} = V_{t,j} + \delta(1 - p(y_t|V_t)) + \varepsilon,$$

where δ is a learning rate between 0 and 1. This implies that when a prediction error occurs (e.g., because the estimated value of x_i is higher than x_j), the value of the higher-ranked item is increased and the value of the lower-ranked item is decreased. Finally, items' positions were subject to random fluctuation through the addition of noise with $\varepsilon \sim \mathcal{N}(0, \sigma_d)$.

Simulation study

The models were trained on the sequences presented to participants in the Chains and Jumps conditions, up to the average number of training blocks completed by participants (5 blocks). For both models, predicted accuracy on the forced choice test was based on an optimal decision rule such that an item was judged to be ranked higher if its underlying value was greater than the alternative for a given test pair.

For the particle filter, performance was simulated across a grid of values for the number of particles N and the SD of the proposal distribution σ_d. I considered two variants of the particle filter based on local updating ($\alpha = 0$) and global updating ($\alpha = 10$) as described above. For RL-Elo the simulations covered a range of values for the learning rate δ and the width of the noise distribution σ_d. For all models the scaling factor β was fixed to 1. Performance was averaged across 5000 iterations for each combination of parameter values.

Results The top row of Figure 4 shows the overall predicted test accuracy. For both particle filter models (left two columns), increasing numbers of particles (N) leads to higher accuracy and approaches ceiling for $N = 1000$. The bottom row of Figure 4 shows the difference between the Chains and Jumps conditions, with higher accuracy in the Chains condition across a wide range of parameter values for both local ($\alpha = 0$) and global ($\alpha = 10$) updating. The effect is largest when the number of particles is relatively low (peaking around $N = 40$). As the number of particles increases the advantage of chained study declines, and when $N = 1000$ the effect disappears as accuracy nears ceiling for both conditions. As a preliminary evaluation of the model's ability to capture human performance, I found the parameter values that minimized the RMSE between participants' test accuracy and the model predictions. The predicted accuracy for the resulting set ($N = 40$, $\sigma_d = 1$, $\alpha = 10$) is shown in the inset of Figure 4. In addition to the difference between the training conditions, the model predicts symbolic distance effects as seen in the behavioral results. The greatest source of error is for recall of the studied premise pairs (distance = 1), for which the model underestimates the actual performance in both conditions.

Across a wide range of parameter values for RL-Elo there was little evidence for differences in performance between the Chains and Jumps conditions. The only exception was a small advantage for chained study when there was no noise ($\sigma_d = 0$), but the high levels of accuracy in both conditions suggest this would not be able to account for the behavioral results.

Comparing the particle filter and RL-Elo across values of σ_d illustrates the different effects of random drift on learning under the two models. Increasing σ_d under RL-Elo leads to rapid declines in performance, with drift acting strictly as a form of forgetting. For the particle filter, random drift during propagation also leads to some forgetting in that it reduces the influence of earlier premises on items' estimated positions. However, propagation is crucial for maintaining diversity among the particles and allowing the model to explore for better solutions throughout training. Even for large numbers of particles, the particle filter performs poorly when σ_d is low because the particles remain anchored to their random initial positions, while for values of $\sigma_d \geq 1$ the model performs at similarly high levels of performance.

Discussion

Past work on TI indicates that constructive processes are at work during learning, as learners piece together a unified mental representation of the hierarchy as premises are observed. The results of the behavioral study indicate that this process is easier when premises are presented in chained sequences, while training sequences without any overlap be-

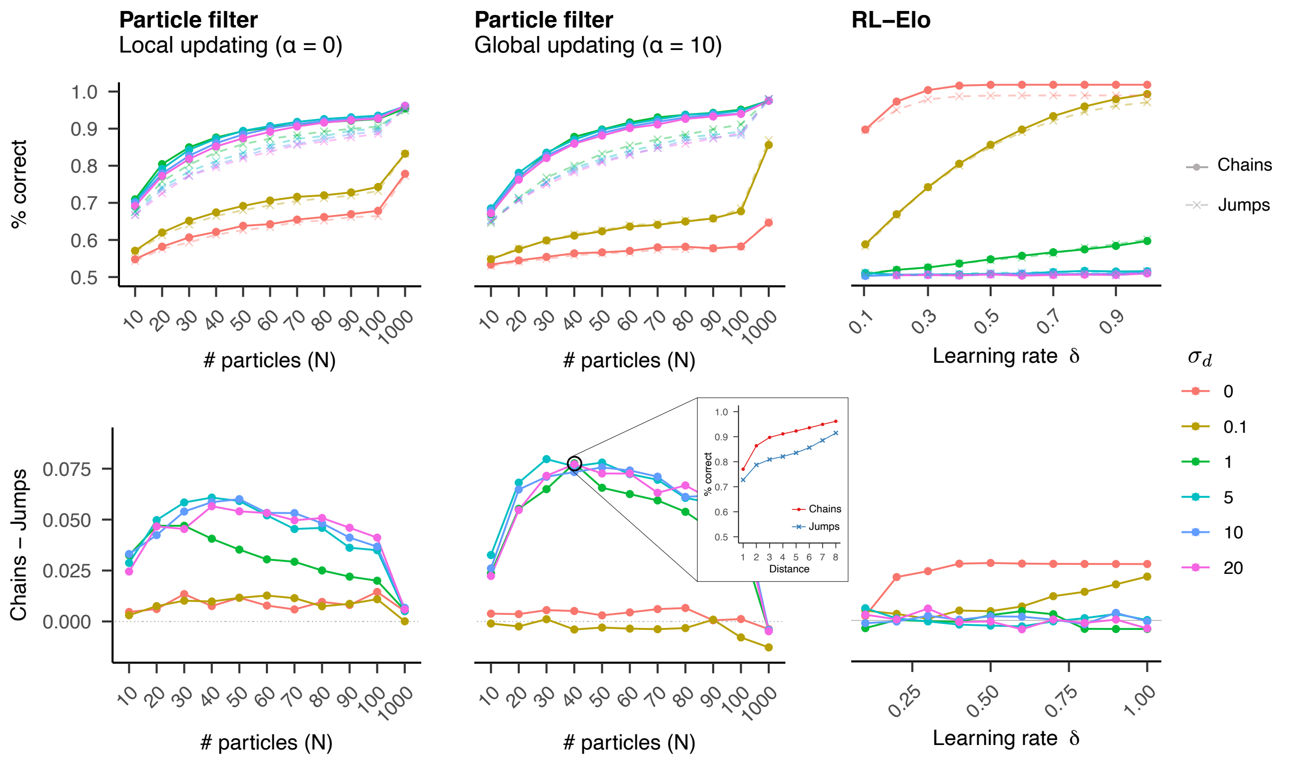

Figure 4: *Top:* Accuracy on the forced choice inference test from the model simulation study for the Chains and Jumps conditions. *Bottom:* Difference in predicted inference accuracy between the Chains and Jumps condition.

tween successive premises leads to poorer relational learning. This finding echoes earlier work showing benefits of chaining in serial order learning (Foos et al., 1976; Foos, 1984) and simpler variants of the TI task (Andrews & Halford, 1998; Halford, 1984; Waltz et al., 1999, 2004).

A major contribution of the present work is to show that the advantage of chained study is naturally accounted for by a particle filter without assuming any other differences between conditions. Importantly, the effect is tied to the approximate nature of the particle filter, disappearing for models with large numbers of particles that provide a finer approximation of the full posterior. For smaller values of N, random trial-to-trial fluctuation in items' positions during propagation has an outsize impact on the distribution of particles and can dilute the effect of earlier premises. For instance, when learners in the Jumps condition encounter the premise $B < C$, multiple trials have passed since seeing the premise $A < B$, by which point the estimated positions of A and B could have reversed in some particles. Chained sequences appear to reinforce the correct ordering of nearby items that are most susceptible to the effects of such variability when the number of particles is relatively low.

The advantage from chained study was larger for the global updating model ($\alpha = 10$) which assumed that the estimated positions of other items were taken into account, thereby favoring particles in which the items in the current premise were adjacent in the implied order. However, the benefit from chaining was still apparent for the local updating model ($\alpha = 0$) where the likelihood depends only on the relative positions of the items in the current premise (regardless of whether they are immediately adjacent in the implied order). This suggests that a similar mechanism could explain the advantage from chained study in settings where people are unlikely to consider unobserved elements during training, such as when they are not informed about the underlying hierarchy (Markant, 2021).

Although further work is needed to fit the model to the behavioral data, inspection of the best-fitting particle filter from the grid search (inset of Figure 4) suggests that the model can reproduce the key features of the results, including the difference between training conditions and the symbolic distance effect. A notable exception is that it underestimates recall accuracy for studied premises. Under the particle filter, adjacent items will tend to be closest in the learned representation of the hierarchy, making it difficult to account for the U-shaped pattern in the Jumps condition (Figure 2A). One possibility is that people rely on a form of direct memory for studied premises that is independent of the integrated map represented by the particle filter (Russin, Zolfaghar, Park, Boorman, & O'Reilly, 2021). An important question for future work is how direct memory for studied premises might influence constructive processes during relational learning and inference.

References

Abbott, J. T., & Griffiths, T. L. (2011). Exploring the influence of particle filter parameters on order effects in causal learning. In L. Carlson & T. F. Shipley (Eds.), *Proceedings of the 33rd Annual Meeting of the Cognitive Science Society.* Austin, TX: Cognitive Science Society.

Andrews, G., & Halford, G. S. (1998). Children's ability to make transitive inferences: The importance of premise integration and structural complexity. *Cognitive Development*, *13*(4), 479–513.

Bainbridge, W. A., Isola, P., & Oliva, A. (2013). The intrinsic memorability of face photographs. *Journal of Experimental Psychology: General*, *142*(4), 1323.

Brown, S., & Steyvers, M. (2009). Detecting and predicting changes. *Cognitive Psychology*, *58*(1), 49–67.

Bryant, P. E., & Trabasso, T. (1971). Transitive inferences and memory in young children. *Nature*, *232*, 456–458.

Daw, N., & Courville, A. (2008). The pigeon as particle filter. *Advances in Neural Information Processing Systems*, *20*, 369–376.

Doucet, A., De Freitas, N., & Gordon, N. (2001). An introduction to sequential Monte Carlo methods. In *Sequential Monte Carlo methods in practice* (pp. 3–14). Springer.

Foos, P. W. (1984). The construction of linear orderings under conditions of increased memory load. *Bulletin of the Psychonomic Society*, *22*(5), 406–408. doi: 10.3758/BF03333859

Foos, P. W., Sabol, M. A., Smith, K. H., & Mynatt, B. T. (1976). Constructive Processes in Simple Linear-Order Problems. *Journal of Experimental Psychology: Human Learning and Memory*, *2*(6), 759–766.

Frank, M. J., Rudy, J. W., & O'Reilly, R. C. (2003). Transitivity, flexibility, conjunctive representations, and the hippocampus. II. A computational analysis. *Hippocampus*, *13*(3), 341–354.

Halford, G. S. (1984). Can young children integrate premises in transitivity and serial order tasks? *Cognitive Psychology*, *16*(1), 65–93.

Hummel, J. E., & Holyoak, K. J. (2001). A process model of human transitive inference. In *Spatial Schemas in Abstract Thought* (pp. 279–305).

Jensen, G., Muñoz, F., Alkan, Y., Ferrera, V. P., & Terrace, H. S. (2015). Implicit Value Updating Explains Transitive Inference Performance: The Betasort Model. *PLOS Computational Biology*, *11*(9), e1004523. doi: 10.1371/journal.pcbi.1004523

Jensen, G., Terrace, H. S., & Ferrera, V. P. (2019). Discovering Implied Serial Order Through Model-Free and Model-Based Learning. *Frontiers in Neuroscience*, *13*. doi: 10.3389/fnins.2019.00878

Kumaran, D., Banino, A., Blundell, C., Hassabis, D., & Dayan, P. (2016). Computations underlying social hierarchy learning: distinct neural mechanisms for updating and representing self-relevant information. *Neuron*, *92*(5), 1135–1147.

Kumaran, D., & McClelland, J. (2012). Generalization through the recurrent interaction of episodic memories: A model of the hippocampal system. *Psychological Review*, *119*(3), 573.

Lloyd, K., Sanborn, A., Leslie, D., & Lewandowsky, S. (2019). Why higher working memory capacity may help You learn: Sampling, search, and degrees of approximation. *Cognitive Science*, *43*(12). doi: 10.1111/cogs.12805

Markant, D. B. (2020). Active transitive inference: When learner control facilitates integrative encoding. *Cognition*, *200*, 104188. doi: 10.1016/j.cognition.2020.104188

Markant, D. B. (2021). Chained study and the discovery of relational structure. *Memory & Cognition*. doi: 10.3758/s13421-021-01201-1

Moyer, R. S., & Bayer, R. H. (1976). Mental comparison and the symbolic distance effect. *Cognitive Psychology*, *8*(2), 228–246.

Peer, M., Brunec, I., Newcombe, N., & Epstein, R. (2021). Structuring Knowledge with Cognitive Maps and Cognitive Graphs. *Trends in Cognitive Sciences*, *25*(1), 37–54.

Russin, J., Zolfaghar, M., Park, S. A., Boorman, E., & O'Reilly, R. C. (2021). Complementary Structure-Learning Neural Networks for Relational Reasoning. In T. Fitch, C. Lamm, H. Leder, & K. Teßmar-Raible (Eds.), *43rd Annual Meeting of the Cognitive Science Society.* Austin, TX: Cognitive Science Society.

Sanborn, A., Griffiths, T., & Navarro, D. (2010). Rational approximations to rational models: Alternative algorithms for category learning. *Psychological Review*, *117*(4), 1144–1167.

Son, J.-Y., Bhandari, A., & FeldmanHall, O. (2021). Cognitive maps of social features enable flexible inference in social networks. *Proceedings of the National Academy of Sciences*, *118*(39), e2021699118.

Speekenbrink, M. (2016). A tutorial on particle filters. *Journal of Mathematical Psychology*, *73*, 140–152.

Vasconcelos, M. (2008). Transitive inference in non-human animals: An empirical and theoretical analysis. *Behavioural Processes*, *78*(3), 313–334. doi: 10.1016/j.beproc.2008.02.017

Waltz, J. A., Knowlton, B. J., Holyoak, K. J., Boone, K. B., Back-Madruga, C., McPherson, S., ... Miller, B. L. (2004). Relational integration and executive function in Alzheimer's disease. *Neuropsychology*, *18*(2), 296.

Waltz, J. A., Knowlton, B. J., Holyoak, K. J., Boone, K. B., Mishkin, F. S., de Menezes Santos, M., ... Miller, B. L. (1999). A system for relational reasoning in human prefrontal cortex. *Psychological Science*, *10*(2), 119–125.

Wynne, C. D. L. (1995). Reinforcement accounts for transitive inference performance. *Animal Learning & Behavior*, *23*(2), 207–217. doi: 10.3758/BF03199936

Yi, M. S., Steyvers, M., & Lee, M. (2009). Modeling Human Performance in Restless Bandits with Particle Filters. *The Journal of Problem Solving*, *2*(2), 81–101. doi: 10.7771/1932-6246.1060