Human brains construct individualized global rankings from identical few-shot local learning

Donging Liu[1-4, #], Muzhi Wang[1-5,#], Huan Luo[1-4,*]

[1]School of Psychological and Cognitive Sciences , Peking University; Beijing, 100871, China

[2]PKU-IDG/McGovern Institute for Brain Research, Peking University; Beijing, 100871, China

[3]Beijing Key Laboratory of Behavior and Mental Health, Peking University; Beijing, 100871, China

[4]Key Laboratory of Machine Perception (Ministry of Education), Peking University, Beijing, 100871, China

[5]Applied Computational Psychiatry Lab, Max Planck UCL Centre for Computational Psychiatry and Ageing Research, Queen Square Institute of Neurology, UCL, London, WC1B 5EH, United Kingdom

[#] These authors contributed equally to this work.

*__Correspondence__: huan.luo@pku.edu.cn (H.L.)

Short title: Constructive relational learning in humans

Keywords: transitive inference; subjective ranking; relational learning; few-shot; MEG

Abstract

Ranking—a typical relational structure—helps people organize complex information and overcome cognitive load, yet in real-world settings it is often inferred from few-shot learning of partial comparisons. How the human brain makes relational inferences under such sparse conditions remains unknown. In a preregistered study combined with Magnetoencephalography (MEG) recordings, we show that even after identical few-shot local pair learning, individuals tend to construct distinct, stable global rankings that diverge from the actual ranking order. MEG recordings reveal that frontoparietal neural representations are reorganized to reflect each person's idiosyncratic ranking, distinct from innate ranking knowledge. These findings challenge the prevailing view that identical inputs yield shared outcomes across subjects and highlight the constructive and generative nature of human relational learning—human brains actively infer and impose structure under sparse data constraints.

Introduction

Ranking—a common relational structure among items—is pervasive in daily life (e.g., rankings of tennis players, soccer clubs, or universities). It allows people to organize and compress complex information, hence reducing cognitive load and facilitating efficient decision-making (Dehaene et al., 2022; Iñiguez et al., 2022; Ramawat et al., 2023). A complete ranking typically requires either scalar values for each item or the outcomes of all pairwise comparisons. In reality, however, such complete data are rare; instead, individuals often infer a global ranking from a limited subset of comparisons (Jensen et al., 2021; Nelli et al., 2023; Sorscher et al., 2022; Tenenbaum et al., 2011; Xu et al., 2024). For instance, given match results for only certain pairs of players or teams, humans can estimate the overall order, including player pairs that have never competed directly. Remarkably, humans can make such inferences from only a few samples—a hallmark of few-shot learning, yet the mechanisms that support this ability under sparse conditions remains largely unknown.

A classical paradigm for relational learning is transitive inference (Acuna, Sanes, et al., 2002; Bryant & Trabasso, 1971; Piaget, 1923; Vasconcelos, 2008), in which relationships between non-adjacent items (e.g., A < C) are inferred from those between adjacent items (e.g., A < B, B < C). This capacity is seen not only in humans but also in non-human primates, rats, and birds (Daisley et al., 2021; Davis, 1992; Gillan, 1981; McGonigle & Chalmers, 1977). A related phenomenon, list linking (Nelli et al., 2023; Treichler & Van Tilburg, 1996; Wei et al., 2014), occurs when two separately learned lists (e.g., A < B < C and D < E < F) are rapidly integrated into a unified sequence after learning a single pair that spans the lists (e.g., C < D). Various computational models have been proposed to explain relational learning, including implicit value updating, asymmetric reinforcement learning, and probabilistic preference learning (Ciranka et al., 2022; Delius & Siemann, 1998; Jensen et al., 2015; Mannella & Pezzulo, 2024; Wynne, 1995). Neural network models equipped with active reactivation mechanisms can also produce similar behavior patterns (Kay et al., 2024; Miconi & Kay, 2025; Mitchell, 2007; Nelli et al., 2023), offering biologically

plausible accounts. Brain imaging studies further show that relational learning reshapes neural similarity patterns between items in the Prefrontal cortex, Parietal cortex, and other regions (Acuna, Eliassen, et al., 2002; Brunamonti et al., 2016; Di Bello et al., 2024; Hinton et al., 2010; Munoz et al., 2020a, 2020b; Ramawat et al., 2022; Van Opstal et al., 2009; Xu et al., 2024).

Despite methodological and theoretical differences, most models and studies predict that partial local learning yields a common ordering across individuals—either a noisy or systematically shifted version of the ground truth—when provided with the same inputs (Ciranka et al., 2022; Jensen et al., 2015; Li et al., 2017; Nelli et al., 2023). In this view, participants reconstruct a largely shared global order. However, relational learning is inherently a stochastic inference process, especially under noisy few-shot conditions (Lake et al., 2015; Tenenbaum et al., 2011). Such inference is likely shaped by inductive biases or heuristics —arising from feature weighting, memory constraints, and learning strategies—that vary across individuals (Lippl et al., 2024; Tenenbaum et al., 2011). We therefore propose an alternative account: few-shot learning of partial pairs does not yield a shared ranking but instead leads individuals to construct distinctive yet internally consistent global orders that deviates from the ground truth.

We tested these hypotheses in a preregistered few-shot relational learning experiment combined with Magnetoencephalography (MEG) recordings. we posit that, after briefly learning a limited set of non-adjacent pairs, participants would construct their idiosyncratic global rankings, leading to two predictions: (1) most participants would make stable errors on certain local pairs, but the specific error pairs would vary across individuals; and (2) within each participant, these local pair errors would be globally congruent rather than independent, together constituting a self-consistent but incorrect global ranking. A large-scale behavioral study confirmed both predictions, showing that participants derived distinct global rankings despite identical learning inputs. MEG recordings replicated the behavioral effects and further revealed that, after few-shot learning, neural similarity between items aligned with each participant’s own subjective ranking rather than with those of others.

Together, these findings challenge the prevailing view and provide converging behavioral and neural evidence to support the hypothesis that few-shot learning of partial pairs induces the subjective construction of global rankings in the human brain—a phenomenon largely overlooked in current models or theories of relational learning.

Results

Preregistered few-shot learning paradigm

We developed a preregistered few-shot learning paradigm (http://osf.io/dpxq4) to investigate how humans infer artificially defined rankings of items (Fig. 1a, left; left-to-right corresponds to low-to-high ranks) from brief learning of a limited set of non-adjacent local pairs (Fig. 1a, left, connecting lines). The experiment comprises two phases (Fig. 1a; right): few-shot learning phase (brown box) and testing phase (blue box). Participants first learned the relationship of eight non-adjacent image pairs (Fig. 1b; e.g., "Suitcase" scores three units higher than "Envelope"), with each pair presented only once per block across four blocks. In the testing phase, participants judged all 28 possible image pairs without feedback, selecting the image they believed had a higher score (Fig. 1c). Notably, subjects learned and were tested only on pairwise relationships, without explicit instruction to construct a full ranking of the eight items. Most crucially, all subjects were trained on the identical eight pre-determined local pairs (connecting lines in Fig. 1a), ensuring that any inter-subject differences in global ranking (as hypothesized in the preregistration) could not be attributed to variations in the learning sets.

Overall, participants demonstrated good ranking performance after few-shot learning. As expected, the directly learned image pairs (Fig. 1d, light blue) showed high accuracy, confirming successful learning. Critically, participants also performed relatively well on the pairs that were never directly experienced (Fig. 1d, dark blue), indicating their transitive inference ability from few-shot exposure to partial, non-adjacent pairs. Moreover, the inferred rankings exhibited hallmark signatures of ordinal ranking, including the "serial position effect"—enhanced accuracy for items at the beginning and end of the sequence (Fig. 1e)—and the "symbolic distance effect"—reduced accuracy for item pairs with smaller rank differences (Fig. 1f).

Figure 1 about here

Participants exhibit consistent, subject-specific local ranking errors

We next tested the core preregistered hypotheses by examining the consistent error patterns in participants' rankings after few-shot learning. Notably, eight of the 77 subjects performed well (accuracy > 50 %) on all local pairwise tests and were thereby excluded, leaving 69 subjects for subsequent analyses. Figure 2a shows their grand average ranking accuracy for all 28 local pairs. As expected, performance was higher for pairs with larger ranking distances (i.e., farther from the diagonal) and lower for those with smaller differences (closer to the diagonal).

Given that all participants learned the same local pairs, one might expect them to construct similar global rankings and exhibit comparable consistent error patterns. Surprisingly, however, we observed a striking subject-specific consistency in individual error patterns (Fig. 2h; individual error profiles). That is, each participant tended to make their own characteristic and repeated errors on specific pairs—for example, consistently inferring A < B when the correct relation was A > B. Two preregistered analyses were conducted to validate the findings (see Fig. 2de).

First, as preregistered, we examined the accuracy distribution across subjects for each local pair. For example, consider the 3rd vs. 4th pair (Fig. 2a, highlighted in red dotted box), which had a group mean accuracy of 0.67. If subjects inferred similar global rankings, we would expect a unimodal distribution centered around 0.67 (Fig. 2b, upper). However, the accuracy showed a bimodal distribution across subjects (Fig. 2b, lower), indicating that some participants consistently made correct inference ($3^{rd} < 4^{th}$), while others made robust errors ($3^{rd} > 4^{th}$). To quantify this bimodal pattern, we fit a Beta distribution to the accuracy profile of each pair (Fig. 2c). As shown in Fig. 2d, easy pairs (large ranking distances; lower-left corner) followed high-accuracy profiles (brown; $\alpha > 1$, $\beta < 1$). Critically, difficult pairs (i.e., small ranking distance, near the diagonal) tended to show bimodal distributions (green; $\alpha < 1$, $\beta < 1$) instead of unimodal distribution ($\alpha > 1$, $\beta > 1$). The results are thus consistent with our *preregistered Hypothesis I*—bimodal distribution of local pair accuracy across subjects—indicating that participants made distinct but consistent errors, rather than shared and random ones.

In addition to assessing cross-subject distribution, we further quantified the error consistency at the individual level. As shown in Fig. 2e (left), participants exhibited high levels of error consistency (above 0.8; one-sample t-test, $p < 0.001$). Remarkably, 54 of the 69 participants (54/69, 78%) made 100% consistent errors on at least one local pair, and the number rose to 63 participants (63/69, 91%) for 80% error threshold (Fig. 2e, right).

As a comparison, we simulated the ranking behavior of all 69 participants using *Q-learning* (Ciranka et al., 2022; Jensen et al., 2015), the canonical reward-prediction error (RPE) model, with individual ranking accuracies matched to those observed empirically. We then applied the same Beta distribution fitting to the simulated data and examined accuracy distributions across subjects. As shown in Figure 2f, easy pairs (large ranking distances; lower-left corner) produced the expected high-accuracy profiles (brown; $\alpha > 1$, $\beta < 1$), closely resembling experimental results. In contrast, difficult pairs (small ranking distances, near the diagonal) yielded unimodal distributions (grey; $\alpha > 1$, $\beta > 1$), diverging from the bimodal patterns observed in the empirical data (Fig. 2d). Consistently, simulated participants exhibited low error consistency (Fig. 2g, left), and only a small fraction (7 of 69) showed consistent errors at the 80% threshold on at least one local pair (Fig. 2g, right). These results indicate that the subject-specific local ranking errors (i.e., the bimodal distribution) observed experimentally cannot be accounted for by the classical Q-learning model (see Fig. S1 for similar results with *betasort* and *betaQ* models).

In summary, these findings support preregistered Hypothesis I – despite receiving identical few-shot local training, participants exhibited stable, subject-specific errors on certain local pairs—reflecting individualized distortions in their constructed rankings (Fig. 2i)—that cannot be readily explained by existing computational models.

Figure 2 about here

Participants form self-consistent, subject-specific global rankings

We next tested our *preregistered hypothesis II* and investigated the origin of the observed consistent local pair errors. One possibility is that these errors reflect isolated, local distortions that occur randomly within each subject. Alternatively, they may arise from a coherent yet distorted global ranking that participants construct. Whereas the former hypothesis predicts random, uncorrelated errors, the latter suggests that local errors within each individual should be internally consistent—constrained by a unified, albeit incorrect, global ranking.

To test this, we performed a self-consistency analysis. As illustrated in Figure 3a, the relationships among three items (ground truth: A < B < C) can fall into one of three categories: correct triad (e.g., A < B, B < C, A < C; in line with A < B < C); self-consistent but incorrect triad (e.g., A < B, B > C, A < C; in line with A < C < B), and self-inconsistent and incorrect triad (e.g., A < B, B < C, A > C; violating transitivity). We applied this triplet-based analysis across all possible item triads for each participant to assess the internal consistency of their inferred rankings.

As previously noted, eight participants were classified into the correct-ranking group. Of the remaining 69, 64 (92.8%) exhibited self-consistent but incorrect rankings, whereas only 5 (7.2%) showed self-inconsistent errors (Fig. 3b, left). Also, as preregistered, we quantified how self-consistent participants' judgements were by calculating the self-consistency coefficients based on how many self-inconsistent triads participants made (see Methods). We tested the group-level self-consistency by shuffling the accuracy matrix of each participant. The participants showed significantly higher self-consistency than these from permutations ($p < 0.001$, permutation-based test; $M = 0.68$, 95% quantile of the surrogation distribution: 0.70; from 0 to 1, from completely inconsistent to perfectly consistent). These results align with *preregistered Hypothesis II*, supporting that individuals made self-consistent local ranking errors. In other words, these local pairwise errors are not independent but are instead congruent within a global framework. In contrast, Q-learning model simulations did not reproduce the high prevalence of self-consistent errors (Fig. 3b, right; also see Figure S1 for similar results with *betasort* and *betaQ* models).

Given the predominance of self-consistent errors, we next reconstructed each participant's global ranking based on their own 28 pairwise judgments (one subject example in Fig. 3c), which revealed individualized global rankings that deviated from the ground-truth ranking (fig. 3e). In fact, Q-learning simulations produced substantially higher inter-subject similarity in global rankings than behavioral results (Fig. 3d, two-sample *t*-test, $t = -66.77$, $p < 0.001$), suggesting that classical learning models predict convergence toward a common ranking rather than the observed individualized global rankings (see Figure S1 for results with *betaQ* and *betasort* models).

In summary, these findings confirm preregistered Hypothesis II: although all participants underwent identical few-shot learning of partial local pairs, they constructed self-consistent, subject-specific global rankings that deviated from the true order and varied across individuals—patterns that cannot be readily explained by existing computational models.

Figure 3 about here

MEG experiment and replication of behavioral findings

Finally, we conducted MEG recordings (Elekta NeuroMag system) on a new cohort of participants ($N = 30$; 3 subjects excluded due to poor learning performance) to examine the neural basis of subject-specific global rankings formed through few-shot learning. The experiment consisted of three phases—pre-learning, few-shot learning, and post-learning (Fig. 4a). In the pre-learning phase, participants performed a real-world size ranking task on eight images, judging on each trial whether the current item was larger or smaller than the previous one (Fig. 4b; e.g., Sofa is larger than Drum). In the few-shot learning phase, same as the earlier behavioral protocol, they learned novel ranks for the same images by experiencing eight non-adjacent local pairs. The post-learning phase matched the pre-learning phase in timing and stimuli but required judgments based on the newly learned ranking (Fig. 4b). Thus, the only

difference between pre- and post-learning was the ranking rule: real-world size versus newly acquired ranks.

All key behavioral findings were replicated. Participants performed well on both ranking tasks (Fig. 4c) and showed classic serial position and distance effects (Fig. S2). Importantly, 24 of the 27 participants (24/27, 88%) made 80% consistent errors on at least one local pair (Fig. 4d), and 26 exhibited self-consistent but incorrect global rankings (Fig. 4e). These results support that subjects constructed individualized, self-consistent global rankings after few-shot learning.

Figure 4 about here

Neural activities align with subject-specific global rankings after few-shot learning

We hypothesized that neural correlates of subject-specific global rankings would emerge in the post-learning phase following few-shot learning, while real-world size representations—being innate—would be present in both pre- and post-learning phases. To test this, we employed representational similarity analysis (RSA), assuming that images closer in rank would evoke more similar neural responses.

As shown in Fig. 4f, we constructed two representational dissimilarity matrices (RDMs): one based on real-world size (RDM1; green), identical across participants, and one based on each subject's reconstructed global ranking (RDM2; purple). For each subject and time point, neural dissimilarities between image pairs were simultaneously regressed onto both RDMs, separately for the pre- and post-learning phases. We quantified the neural representation of real-world size as the average of the correlation coefficients (β□) across both phases, reflecting its presence during both phases (Fig. 4g). The neural representation of subject-specific rankings was defined as the difference in correlation coefficients (β□) between post- and pre-learning phases, isolating learning-related changes (Fig. 4i).

Neural signatures of real-world size appeared around 250 ms after image onset (Fig. 4g; cluster-based permutation test, corrected: 245–265 ms, p = 0.001; 890–1005

ms, p = 0.013). Crucially, neural correlates of subject-specific rankings emerged at a later latency (Fig. 4i; 685–775 ms, p = 0.032). Sensor-level searchlight analysis revealed distinct spatial distributions: real-world size effects were observed at posterior and lateral sensors, while subject-specific ranking effects localized to mid-parietal regions (Fig. 4h). A permutation test, shuffling RDM2 across subjects, confirmed the subject-specific nature of global rankings—neural responses aligned with each subject's own ranking, not with others' (Fig. 4j; p = 0.01).

Together, these results provide converging behavioral and neural evidence that subjects construct individualized global rankings after few-shot learning of partial local pairs.

Discussion

In this preregistered study, we uncover a striking phenomenon: even after receiving identical learning inputs, human participants construct distinct and stable global rankings that deviate from the true order. MEG recordings further show that neural similarity patterns between items are reorganized to align with each individual's ranking, rather than with rankings constructed by others. This subject-specific representation emerges in the frontoparietal cortex and is distinct from neural signatures associated with innate ranking knowledge (e.g., real-world object size). These findings challenge the prevailing view that identical local inputs yield similar global rankings and could not be fully captured by existing models of relational learning. Together, our findings underscore the constructive nature of human relational learning—human brains actively infer and impose structure when faced with sparse data.

Numerous models have been proposed to explain transitive inference—the canonical relational learning paradigm in which relationships between non-adjacent pairs are inferred from adjacent pairs (Jensen et al., 2019). One class posits that inference relies on an internal "number line" representation of items, enabling implicit global updates on each trial—a schema-based rather than model-free framework (Di Antonio et al., 2024). Another models transitive learning as reinforcement learning with asymmetric update rules, adjusting only the winner or loser per trial and producing the value compression often observed in behavioral economics (Ciranka et al., 2022). Moreover, transitive inference has also been quantified as probabilistic preference learning, e.g., via the one-parameter Mallows model, which infers rankings from the distribution of possible global orderings given observed pair (Mannella & Pezzulo, 2024). While our findings are not readily explained by these models, they align with the broader notion that even isolated pairwise learning triggers the construction of a unified mental number line that organizes and updates items globally, yielding internally consistent yet incorrect subjective rankings. In this view, value updating is not confined to local pairs but propagates automatically across the entire

set, constrained by the inferred relational structure.

Our MEG data show that item-wise neural similarity patterns are reconfigured to reflect each participant's idiosyncratic ranking. This newly acquired ranking representation emerges later in the frontoparietal cortex, whereas innate knowledge rankings (e.g., real-world object size) appear earlier in temporal regions. Prior neurophysiological studies in monkeys and neuroimaging work in humans have established the central role of the frontoparietal cortex in transitive inference (Brunamonti et al., 2016; Di Bello et al., 2024; Munoz et al., 2020a, 2020b; Sorscher et al., 2022). More recently, a "list linking" fMRI study reported rapid reorganization of object representations along an elongated neural manifold in dorsal frontoparietal areas following linking-pair exposure (Nelli et al., 2023). It has also been proposed that neural representations along a "geometric" mental line arise as a monotonic mapping between sensory input and behavioral output, suggesting a possible role for motor planning in transitive inference (Di Antonio et al., 2024). Our findings are consistent with these accounts, reinforcing the pivotal role of the frontoparietal cortex in storing newly learned relationships. Crucially, they extend prior work by providing direct neural evidence that these patterns encode subjective—rather than shared—global orderings.

A recent study reported that monkeys exhibit superstitious learning by imposing spurious orderings on untrained stimulus sets, irrespective of random reward feedback (Jin et al., 2022). This subjective ranking has been interpreted as a model-based strategy operating in unstructured environments, with internal models formed independently of trial-and-error learning. Our findings parallel but also diverge from this account. Unlike the random, unstructured relationships examined previously, our stimuli contained a hidden ranking structure that participants had to infer from limited exposure to partial pairs. Moreover, no feedback was provided, requiring pure inference rather than reinforcement. These results extend prior work by showing that subjective ranking is not restricted to unstructured inputs but also emerges in structured contexts, where it can induce idiosyncratic distortions of the true ranking.

What drives this subject-specific ranking after identical local learning? A likely

key factor is the few-shot learning task (Nelli et al., 2023): each of the eight non-adjacent pairs was learned only four times—far fewer than in most transitive inference studies and vastly less than what artificial neural networks require (Miconi & Kay, 2025). Such sparse exposure, combined with human memory constraints (Buschman et al., 2011; Cowan, 2010), likely produces noisier, more fragile representations that are easily reorganized under individual inductive biases (Mitchell, 2007; Palminteri et al., 2017). While relational learning and inference is inherently an ambiguous and underdetermined process (Jin et al., 2022; Lippl et al., 2024), few-shot learning amplifies these idiosyncratic influences. Importantly, this scenario has high ecological validity—humans routinely infer rankings and make decisions from incomplete information and minimal feedback in daily life. Our results therefore point to a largely overlooked mechanism for relational inference under sparse inputs.

Our task diverged from standard transitive inference (Ciranka et al., 2022; Vasconcelos, 2008) setup in two major ways: (1) participants learned non-adjacent rather than adjacent pairs, and (2) instead of making binary choices with feedback, they were shown the scalar scores of both items in each pair. This design resembles the “list linking” task used in previous studies (e.g., judging which object is “more brisp”) (Nelli et al., 2023), except that here participants selected the item with the higher score. Importantly, they learned only local pair relationships and were never instructed to construct a full ranking. From a reinforcement learning perspective (Sutton & Barto, 1998), this design should have simplified the task: values could be updated directly from the presented scores rather than inferred indirectly from binary-choice feedback, thereby facilitating more rapid and consistent global rankings across participants. Yet, the substantial between-subject variability we observed suggests that task structure alone was not the principal driver of individualized rankings.

What determines the particular “incorrect” global ranking adopted by each subject? Several factors may contribute when individuals attempt to learn relational information beyond their processing capacity. Participants may differentially weight item features, using them as idiosyncratic anchors to impose latent structural templates during relational inference (Acuna, Sanes, et al., 2002; Tenenbaum et al.,

2011). They may also rely on distinct memory compression strategies, storing a reduced and simplified version of the relational graph, which inevitably introduces systematic distortions or biases (Acuna, Sanes, et al., 2002; Jin et al., 2022). Consistent with this view, a recent neural network model of transitive inference showed that variation in a single conjunctive parameter in the output layer could produce individual differences in ranking behavior (Lippl et al., 2024).

In summary, out study highlights the constructive and generative nature of relational learning when faced with sparse inputs, whereby incomplete and noisy inputs are re-registered into an internally constructed, individually idiosyncratic ordering schema.

Acknowledgements

This work was supported by the National Science and Technology Innovation STI2030-Major Project 2021ZD0204100 (2021ZD0204103 to H.L.) and the Science Fund for Creative Research Groups of the National Natural Science Foundation of China (T2421004). We thank the Center for MRI Research at Peking University in Beijing, China, for assistance with data acquisition.

Author contributions

D.L., M.W., and H.L. originally conceived the study. D.L. and M.W. performed the experiments and analysed the data. D.L., M.W., and H.L. wrote the paper.

Declaration of interests

The authors declare no competing interests.

Data and materials availability

All data and codes will be made publicly available upon acceptance of the paper.

Declaration of generative AI and AI-assisted technologies in the manuscript preparation process

The authors used ChatGPT for language editing and clarity improvements. After using this tool/service, the author(s) reviewed and edited the content as needed and take(s) full responsibility for the content of the published article.

Figure legends

Figure 1

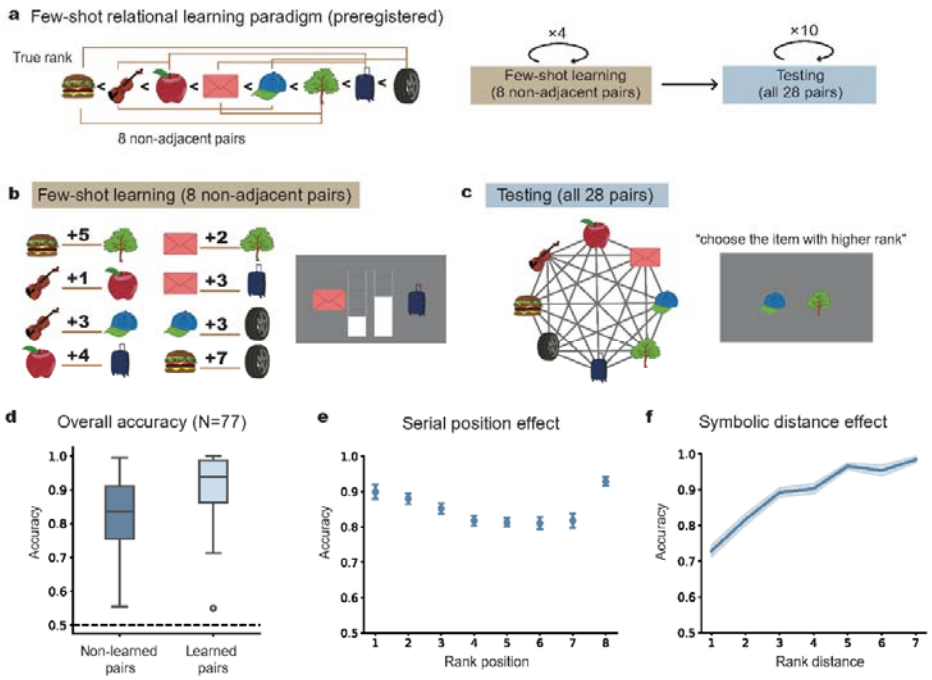

Figure 1. Preregistered few-shot learning paradigm. **a.** *Left*: predefined artificial ranking (true rank) of eight images (left-to-right corresponding to low-to-high ranks), unknown to subjects. Eight non-adjacent local pairs (connecting lines) were selected for few-shot learning (identical pairs for all subjects). *Right*: experiment consists of two phases: few-shot learning phase (brown box; each of the 8 pairs presented once per block across four blocks) and testing phase (blue box; each of the 28 pairs tested 10 times). Notably, subjects learned and were tested only on pairwise relationships, without explicit instruction to construct a full ranking of the eight images. **b.** Few-shot learning phase. *Left*: eight non-adjacent pairs and their score distance (indicated by the number); *Right*: sample trial display, i.e., "Envelope" scores three units lower than "Suitcase". **c.** Testing phase. *Left*: all 28 image pairs (connecting lines) were tested without feedback. *Right*: sample trial display, i.e., selecting the item with a higher score. **d.** Grand averaged ranking accuracy, for learned pairs (dark blue; 8 pairs in few-shot learning) and non-learned pairs (light blue; remaining 20 pairs not directly learned). **e.** Grand averaged accuracy as function of rank position (serial position effect). **f.** Grand averaged accuracy as function of ranking distance (distance effect).

Figure 2

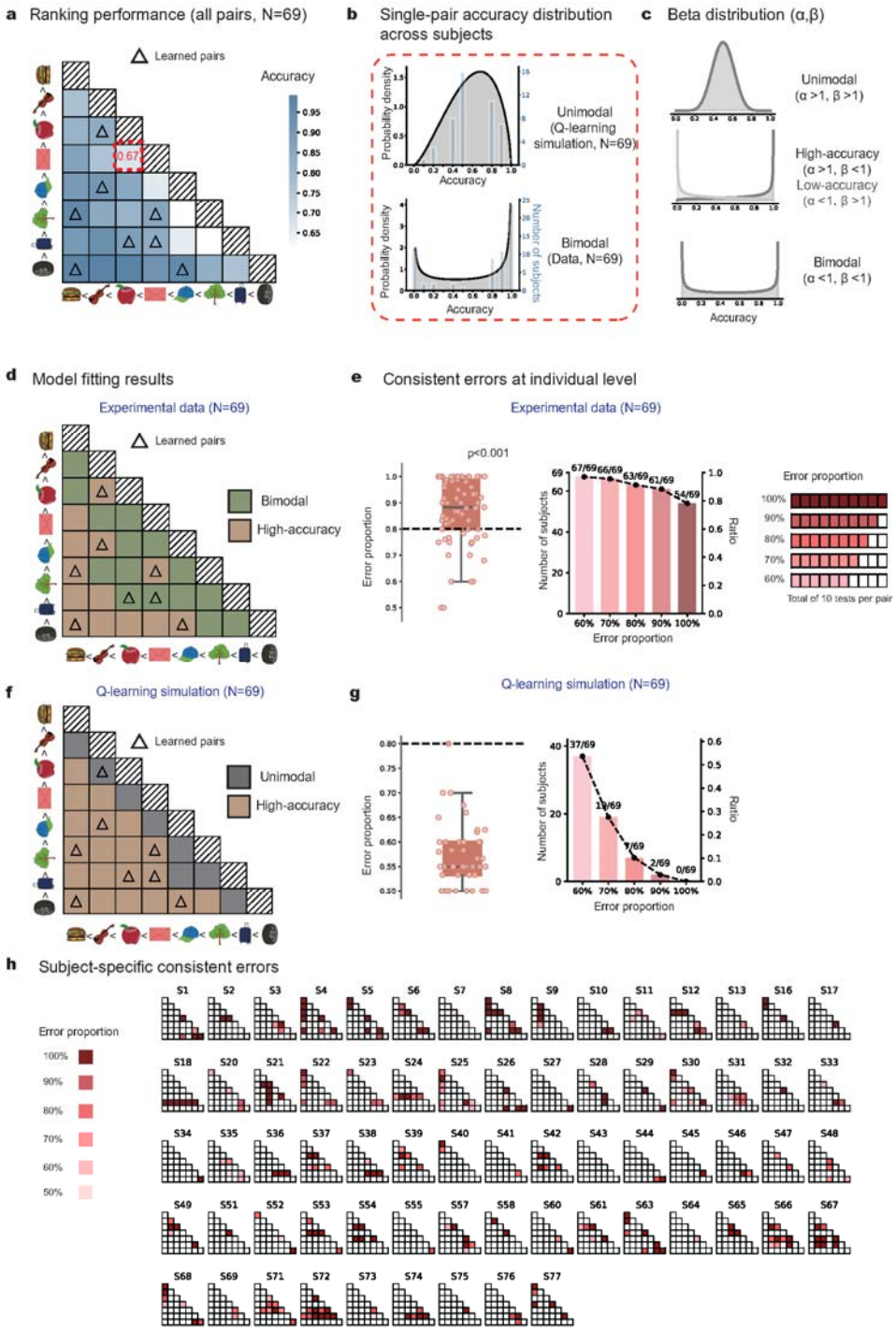

Figure 2. Consistent, subject-specific local errors in testing phase (*preregistered hypothesis I*). **a.** Grand averaged accuracy matrix for all the 24 tested pairs, with row and column denoting corresponding items per pair (deep-to-light color represent high-to-low accuracy). Eight directly-learned pairs (few-shot learning) were marked by triangles. Red dotted box highlights the exemplar pair in Fig. 2b. **b.** Illustration of

unimodal and bimodal accuracy distribution cross subjects for the exemplar pair (3rd vs. 4th) with a group mean accuracy of 0.67. *Upper*: unimodal distribution based on Q-learning model simulations (matched individual ranking accuracies), predicting converging global ranking across subjects. *Lower*: observed bimodal distribution, i.e., some subjects consistently made correct inference ($3^{rd} < 4^{th}$), while others made robust errors ($3^{rd} > 4^{th}$). **c.** Beta-distribution model fitting of cross-subject accuracy distribution with two parameters (α, β): Unimodal distribution (upper); High- or low-accuracy profile (middle); Bimodal distribution (bottom). **d.** Beta-distribution model fitting results for each pair (brown: high-accuracy, $\alpha > 1$, $\beta < 1$; green; bimodal distribution, $\alpha < 1$, $\beta < 1$), consistent with *preregistered Hypothesis I.* **e.** Individual-level error consistency. *Left*: grand averaged error consistency(>0.8, one-sample *t*-test, $t = 4.406$, $p < 0.001$). dots denote individual data. *Right*: number and proportion of subjects making consistent error on at least one local pair for different thresholds (0.6 to 1.0). **f.** Same as **d** but for Q-learning model simulation (brown: high-accuracy, $\alpha > 1$, $\beta < 1$; grey; unimodal distribution, $\alpha > 1$, $\beta > 1$). Note that simulation results support unimodal distribution that diverges from experimental findings. **g.** Same as **e** but for Q-learning model simulation. Note the substantially decreased error consistency at individual levels. **h.** Local error pattern for each subject. Blank tiles in the lower triangular matrices denote correct pairs (accuracy > 50%). Red tiles denote error pairs (deep-to-light color represent high-to-low error proportion).

Figure 3

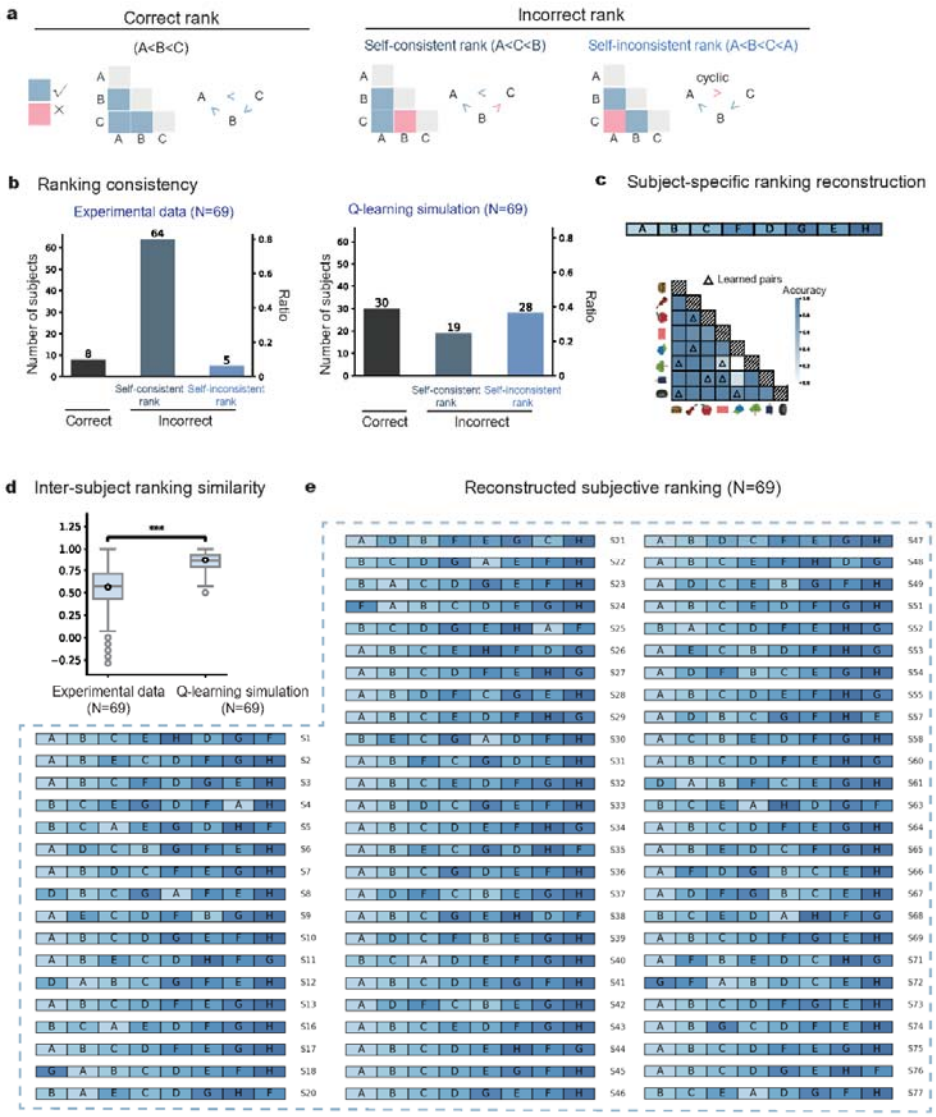

Figure 3. Self-consistent, subject-specific global rankings (*preregistered hypothesis II*). **a.** Illustration of self-consistency ranking analysis for a triple sequence (A < B < C) . Left: correct ranking (A < B, B < C, A < C; consistent with ground truth); Middle: self-consistent but incorrect ranking (A < B, B > C, A < C; consistent with A < C < B); Right: self-inconsistent and incorrect ranking (A < B, B < C, A > C; violating transitivity). The triplet-based analysis was conducted across all possible triads to assess the internal consistency of inferred rankings. **b.** *Left:* Number of subjects for each self-consistency category. *Right:* Number of subjects for each self-consistency category based on Q-learning model simulation. **c.** Exemplar subjective ranking reconstruction. *Upper*: reconstructed individualized global ranking. *Lower*: corresponding accuracy matrix for this subject. Triangles mark the directly learned pairs. **d.** Inter-subject ranking similarity for experimental data and Q-learning model simulation (two-sample t-test, $p < 0.001$). Q-learning simulations showed higher inter-

subject similarity in global rankings than behavioral result. **e.** Reconstructed subjective-specific rankings for all the 69 subjects.

Figure 4

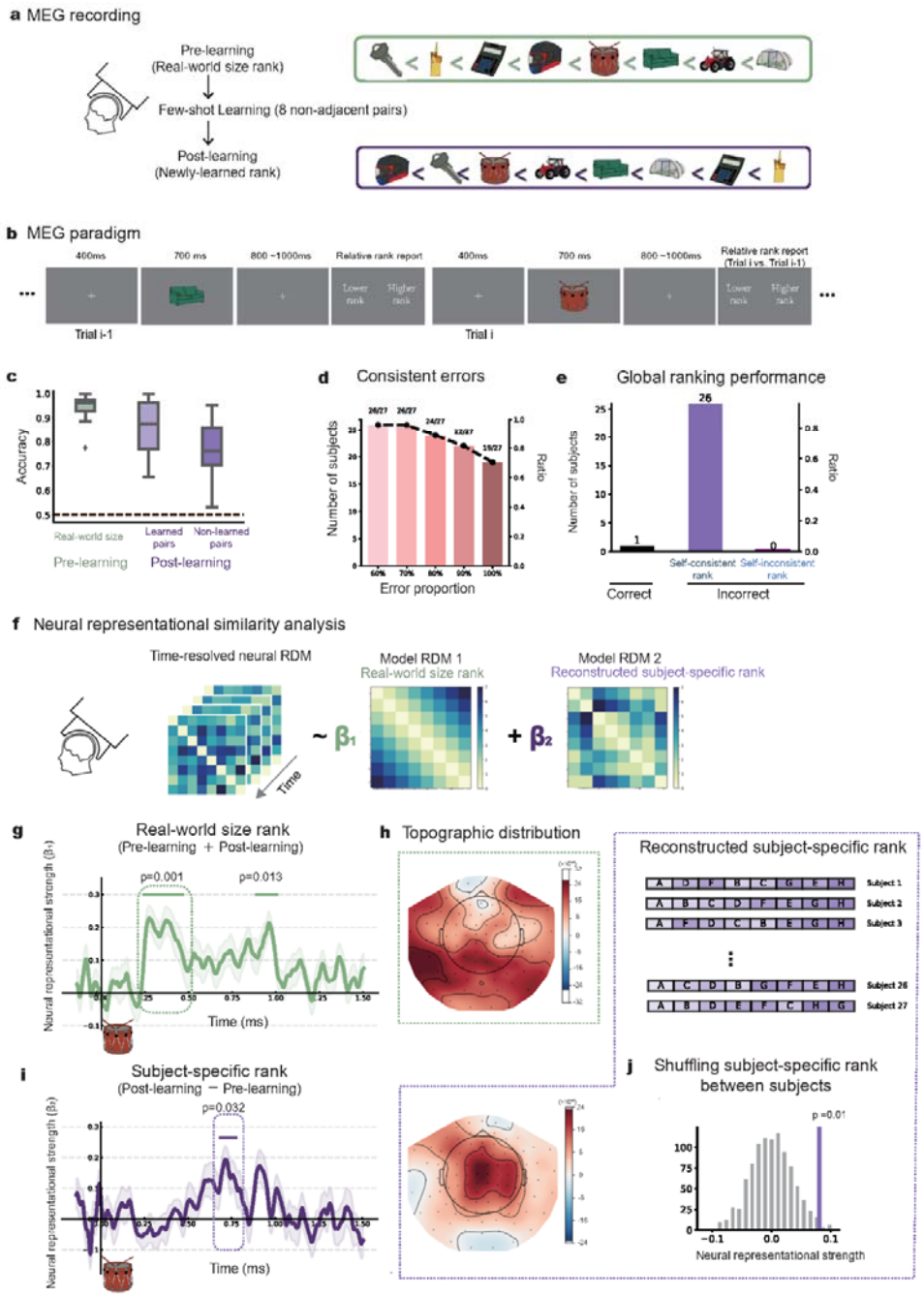

Figure 4. Neural correlates of subject-specific global ranking after few-shot learning (MEG experiment). **a.** MEG experiment consists of three phases—pre-learning, few-shot learning, and post-learning. In the pre-learning phase, participants performed a real-world size ranking task on eight images (e.g., Sofa is larger than Drum). Next, they learned new ranks for the same images through eight non-adjacent local pairs (Few-shot learning), and tested in the post-learning phase (e.g., Sofa ranks higher than Drum). **b.** MEG experimental procedure. Participants judged on each trial whether the current item was larger or smaller than the previous one in terms of real-world size

(pre-learning phase) or newly learned ranking (post-learning phase). Note that pre- and post-learning phases shared the same stimuli and procedure, only differing in ranking rule. **c.** Grand averaged ranking accuracy for real-word size task (pre-learning phase; green) and newly learned ranking test (post-learning phase; light purple: learned pairs; dark purple: non-learned pairs). **d.** Number and proportion of subjects making consistent error on at least one local pair for different thresholds. **e.** Number of subjects for each self-consistency category. **f.** Two representational dissimilarity matrices (RDMs) were constructed, based on real-world size (RDM1; green; identical across subjects) and each subject's reconstructed global ranking (RDM2; purple), respectively, which were used to regress neural dissimilarities, separately for pre- and post-learning phases. **g.** Grand averaged time-resolved RSA results for real-world size as a function of time after image onset. Horizontal lines denote significant time window (cluster-based permutation test, $p < 0.05$). **i.** Same as **h** but for newly-learned ranking. **h.** Topographic distributions for real-world size rank (upper) and subjective rank (bottom) using sensor-level search light analysis. **j.** Permutation test (n=1000, $p < 0.05$) examining neural correlates of subject-specific global ranking (dotted purple box in Fig. 4i) by shuffling RDM2 across subjects (upper).

References

Acuna, B. D., Eliassen, J. C., Donoghue, J. P., & Sanes, J. N. (2002). Frontal and parietal lobe activation during transitive inference in humans. *Cerebral Cortex*, *12*(12), 1312-1321. https://doi.org/10.1093/cercor/12.12.1312

Acuna, B. D., Sanes, J. N., & Donoghue, J. P. (2002). Cognitive mechanisms of transitive inference. *Experimental Brain Research*, *146*(1), 1-10. https://doi.org/10.1007/s00221-002-1092-y

Brunamonti, E., Mione, V., Di Bello, F., Pani, P., Genovesio, A., & Ferraina, S. (2016). Neuronal Modulation in the Prefrontal Cortex in a Transitive Inference Task: Evidence of Neuronal Correlates of Mental Schema Management. *J Neurosci*, *36*(4), 1223-1236. https://doi.org/10.1523/jneurosci.1473-15.2016

Bryant, P. E., & Trabasso, T. (1971). Transitive Inferences and Memory in Young Children. *Nature*, *232*(5311), 456-&. https://doi.org/DOI 10.1038/232456a0

Buschman, T. J., Siegel, M., Roy, J. E., & Miller, E. K. (2011). Neural substrates of cognitive capacity limitations. *Proceedings of the National Academy of Sciences*, *108*(27), 11252-11255. https://doi.org/10.1073/pnas.1104666108

Ciranka, S., Linde-Domingo, J., Padezhki, I., Wicharz, C., Wu, C. M., & Spitzer, B. (2022). Asymmetric reinforcement learning facilitates human inference of transitive relations. *Nature Human Behaviour*, *6*(4), 555-564. https://doi.org/10.1038/s41562-021-01263-w

Cowan, N. (2010). The Magical Mystery Four. *Current Directions in Psychological Science*, *19*(1), 51-57. https://doi.org/doi: 10.1177/0963721409359277

Daisley, J. N., Vallortigara, G., & Regolin, L. (2021). Low-rank

chicks are better at transitive inference reasoning. *Communications Biology*, *4*(1). https://doi.org/ARTN 1344

10.1038/s42003-021-02855-y

Davis, H. (1992). Transitive inference in rats (Rattus norvegicus). *Journal of Comparative Psychology*, *106*(4), 342-349. https://doi.org/10.1037/0735-7036.106.4.342

Dehaene, S., Al Roumi, F., Lakretz, Y., Planton, S., & Sablé-Meyer, M. (2022). Symbols and mental programs: a hypothesis about human singularity. *Trends in Cognitive Sciences*, *26*(9), 751-766. https://doi.org/10.1016/j.tics.2022.06.010

Delius, J. D., & Siemann, M. (1998). Transitive responding in animals and humans: Exaptation rather than adaptation? *Behavioural Processes*, *42*(2), 107-137. https://doi.org/https://doi.org/10.1016/S0376-6357(97)00072-7

Di Antonio, G., Raglio, S., & Mattia, M. (2024). A geometrical solution underlies general neural principle for serial ordering. *Nature Communications*, *15*(1). https://doi.org/10.1038/s41467-024-52240-6

Di Bello, F., Mione, V., Pani, P., Brunamonti, E., & Ferraina, S. (2024). Prefrontal cortex contribution in transitive inference task through the interplay of beta and gamma oscillations. *Communications Biology*, *7*(1). https://doi.org/10.1038/s42003-024-07418-5

Gillan, D. J. (1981). Reasoning in the chimpanzee: II. Transitive inference. *Journal of Experimental*

Psychology: Animal Behavior Processes, *7*(2), 150-164. https://doi.org/10.1037/0097-7403.7.2.150

Hinton, E. C., Dymond, S., von Hecker, U., & Evans, C. J. (2010). Neural correlates of relational reasoning and the symbolic distance effect: involvement of parietal cortex. *Neuroscience*, *168*(1), 138-148. https://doi.org/https://doi.org/10.1016/j.neuroscience.2010.03.052

Iñiguez, G., Pineda, C., Gershenson, C., & Barabási, A.-L. (2022). Dynamics of ranking. *Nature Communications*, *13*(1), 1646. https://doi.org/10.1038/s41467-022-29256-x

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. https://doi.org/10.1371/journal.pcbi.1004523

Jensen, G., Munoz, F., Meaney, A., Terrace, H. S., & Ferrera, V. P. (2021). Transitive inference after minimal training in rhesus macaques (Macaca mulatta). *J Exp Psychol Anim Learn Cogn*, *47*(4), 464-475. https://doi.org/10.1037/xan0000298

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

Jin, Y., Jensen, G., Gottlieb, J., & Ferrera, V. (2022). Superstitious learning of abstract order from random reinforcement. *Proceedings of the National Academy of Sciences*, *119*(35), e2202789119. https://doi.org/doi:10.1073/pnas.2202789119

Kay, K., Biderman, N., Khajeh, R., Beiran, M., Cueva, C. J., Shohamy, D., Jensen, G., Wei, X.-X., Ferrera, V. P., & Abbott, L. F. (2024). Emergent neural dynamics and geometry for generalization in a transitive inference task. *PLOS Computational Biology*, *20*(4), e1011954. https://doi.org/10.1371/journal.pcbi.1011954

Lake, B. M., Salakhutdinov, R., & Tenenbaum, J. B. (2015). Human-level concept learning through probabilistic program induction. *Science*, *350*(6266), 1332-1338. https://doi.org/10.1126/science.aab3050

Li, V., Herce Castañón, S., Solomon, J. A., Vandormael, H., & Summerfield, C. (2017). Robust averaging protects decisions from noise in neural computations. *PLOS Computational Biology*, *13*(8), e1005723. https://doi.org/10.1371/journal.pcbi.1005723

Lippl, S., Kay, K., Jensen, G., Ferrera, V. P., & Abbott, L. F. (2024). A mathematical theory of relational generalization in transitive inference. *Proceedings of the National Academy of Sciences*, *121*(28), e2314511121. https://doi.org/10.1073/pnas.2314511121

Mannella, F., & Pezzulo, G. (2024). Transitive inference as probabilistic preference learning. *Psychonomic Bulletin & Review*, *32*(2), 674-689. https://doi.org/10.3758/s13423-024-02600-6

McGonigle, B. O., & Chalmers, M. (1977). Are monkeys logical? *Nature*, *267*(5613), 694-696. https://doi.org/10.1038/267694a0

Miconi, T., & Kay, K. (2025). Neural mechanisms of relational learning and fast knowledge reassembly in plastic neural networks. *Nature Neuroscience*, *28*(2), 406-414. https://doi.org/10.1038/s41593-024-01852-8

Mitchell, T. M. (2007). The Need for Biases in Learning Generalizations.

Munoz, F., Jensen, G., Kennedy, B. C., Alkan, Y., Terrace, H. S., & Ferrera, V. P. (2020a). Learned

Representation of Implied Serial Order in Posterior Parietal Cortex. *Scientific Reports*, *10*(1), 9386. https://doi.org/10.1038/s41598-020-65838-9

Munoz, F., Jensen, G., Kennedy, B. C., Alkan, Y., Terrace, H. S., & Ferrera, V. P. (2020b). Learned Representation of Implied Serial Order in Posterior Parietal Cortex. *Scientific Reports*, *10*(1). https://doi.org/10.1038/s41598-020-65838-9

Nelli, S., Braun, L., Dumbalska, T., Saxe, A., & Summerfield, C. (2023). Neural knowledge assembly in humans and neural networks. *Neuron*, *111*(9), 1504-1516.e1509. https://doi.org/10.1016/j.neuron.2023.02.014

Palminteri, S., Lefebvre, G., Kilford, E. J., & Blakemore, S.-J. (2017). Confirmation bias in human reinforcement learning: Evidence from counterfactual feedback processing. *PLOS Computational Biology*, *13*(8), e1005684. https://doi.org/10.1371/journal.pcbi.1005684

Piaget, J. (1923). La pensée symbolique et le pensée de l'enfant. [Symbolic thought and the thought of the child.]. *Archives de Psychologie*, *18*, 273-304.

Ramawat, S., Marc, I. B., Ceccarelli, F., Ferrucci, L., Bardella, G., Ferraina, S., Pani, P., & Brunamonti, E. (2023). The transitive inference task to study the neuronal correlates of memory-driven decision making: A monkey neurophysiology perspective. *Neuroscience & Biobehavioral Reviews*, *152*, 105258. https://doi.org/https://doi.org/10.1016/j.neubiorev.2023.105258

Ramawat, S., Mione, V., Di Bello, F., Bardella, G., Genovesio, A., Pani, P., Ferraina, S., & Brunamonti, E. (2022). Different Contribution of the Monkey Prefrontal and Premotor Dorsal Cortex in Decision Making During a Transitive Inference Task. *Neuroscience*, *485*, 147-162. https://doi.org/https://doi.org/10.1016/j.neuroscience.2022.01.013

Sorscher, B., Ganguli, S., & Sompolinsky, H. (2022). Neural representational geometry underlies few-shot concept learning. *Proceedings of the National Academy of Sciences*, *119*(43), e2200800119. https://doi.org/10.1073/pnas.2200800119

Tenenbaum, J. B., Kemp, C., Griffiths, T. L., & Goodman, N. D. (2011). How to Grow a Mind: Statistics, Structure, and Abstraction. *Science*, *331*(6022), 1279-1285. https://doi.org/10.1126/science.1192788

Treichler, F. R., & Van Tilburg, D. (1996). Concurrent conditional discrimination tests of transitive inference by macaque monkeys: List linking. *Journal of Experimental Psychology: Animal Behavior Processes*, *22*(1), 105-117. https://doi.org/10.1037/0097-7403.22.1.105

Van Opstal, F., Fias, W., Peigneux, P., & Verguts, T. (2009). The neural representation of extensively trained ordered sequences. *NeuroImage*, *47*(1), 367-375. https://doi.org/https://doi.org/10.1016/j.neuroimage.2009.04.035

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

Wei, C. A., Kamil, A. C., & Bond, A. B. (2014). Direct and Relational Representation During Transitive List Linking in Pinyon Jays (

). *Journal of Comparative Psychology*, *128*(1), 1-10. https://doi.org/10.1037/a0034627

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

Xu, B., Wu, J., Xiao, H., Münte, T. F., & Ye, Z. (2024). Inferior parietal cortex represents relational structures for explicit transitive inference. *Cerebral Cortex*, *34*(4). https://doi.org/10.1093/cercor/bhae137

Materials and Methods

Participants

In the behavioral experiment, we initially recruited 40 participants through online social media platforms of Peking University as specified in our preregistration, confirming our preregistered hypothesis. To further validate and replicate our initial findings, we recruited an additional 40 participants (total N=80). Three participants were excluded for performing below chance level (50% accuracy) across all 10 testing blocks, resulting in a final sample of 77 participants. For the MEG experiment, we recruited 30 participants from the same population. Three participants were excluded for below-chance performance, yielding a final sample of 27 participants.

The study was approved by the Institutional Review Board of School of Psychological and Cognitive Sciences at Peking University (#2021-10-15). Participants provided written informed consent and received monetary compensation for their time (45 CNY for behavioral experiments, 250 CNY for MEG experiment).

Behavior experimental design

In the experiment, participants were asked to learn a ranking sequence of 8 items (A<B<C<D<E<F<G<H) in the learning phase and were tested in the subsequent test phase. Item images were manually selected from online open-access resources. The same eight items were reused across participants while the ranking sequence of each participant was randomly shuffled. Fig. 1a illustrates an example sequence for one participant.

In each trial of the learning phase, two items with two ticked bars were presented to participants. Participants were instructed that each item represented a movie and that they needed to rank the movie by popularity according to their box office performance. The ticked bars showed how much the two movies sold in one single day. Participants were explicitly instructed to focus on the relative difference of the bars instead of the absolute values as box office selling is subject to day-by-day fluctuation. Participants need to rank the popularity of the movies according to the item pairs. Participants did not need to respond in the learning phase. Eight non-adjacent item pairs were presented for participants.

In the learning phase, participants learned a fixed set of eight non-adjacent item pairs, identical across all participants: (A,F), (B,C), (B,E), (C,G), (D,F), (D,G), (E,H), (A,H). Each item appeared in two pairs equally. The relative differences of a pair ranged from 1 to 7 units (Fig. 1a-b). Although relative differences were determined by item relationships in the rank, the absolute values of bars were determined randomly each trial. The maximum values. The learning phase comprised 4 continuous blocks, with each of the 8 pairs presented exactly once per block in a randomized order. This design ensured that participants received identical few-shot exposure (4 presentations per pair) while controlling for sequence effects through randomization within each

block.

Following the learning phase, the testing phase consisted of 10 continuous testing blocks. In each block, a participant completed all $\binom{8}{2} = 28$ possible pairwise comparisons in a random sequence. In each trial, a participant made a comparison by clicking on the item with the higher rank while two items were presented in parallel on the screen. The next trial starts after 2 second intertrial interval. Critically, no feedback was provided during the testing phase.

Behavioral experiments were implemented with PsychoPy and hosted on Naodao and Credamo. Participants were provided with the experiment link and completed the experiment on their own personal computer/laptop.

Behavioral analysis

Beta distribution fitting. To distinguish whether participants' accuracy patterns reflected polarized responses (clustering near perfect accuracy or complete errors) versus normally distributed performance around an intermediate value, we employed Beta distribution fitting. We fitted a Beta distribution parameterized by shape parameters (α, β) to the accuracy distribution for each pair across all participants. Based on the values of α and β, we classified accuracy distributions into four categories: unimodal, high-accuracy, low-accuracy, and bimodal. When both $\alpha > 1$ and $\beta > 1$, the distribution exhibited a unimodal bell-curve shape centered around an intermediate accuracy level. When $\alpha > 1$ and $\beta < 1$, the distribution showed a high-accuracy pattern with right-skewed distributions toward high performance, while $\alpha < 1$ and $\beta > 1$ produced low-accuracy patterns with left-skewed distributions toward poor performance. When both parameters fell below 1 ($\alpha < 1$, $\beta < 1$), the distribution became bimodal with probability mass concentrated at the extremes (0 and 1), indicating that participants predominantly exhibited either consistently correct or consistently incorrect responses for those pairs. For our analysis, we focused primarily on distinguishing between unimodal and bimodal patterns, as the high-accuracy and low-accuracy conditions could not reliably indicate whether participants exhibited bimodal or unimodal error patterns due to ceiling and floor effects.

Consistent errors. We quantified the error consistency of a participant by calculating the proportion of consistent errors among all pairs for the participant. A pair is defined as a consistent error pair if the proportion of error trials exceeds a threshold. As we have 10 trials for each pair, the threshold has 5 levels: 60%, 70%, 80%, 90%, and 100%. We first use the strictest threshold 100%, i.e., an error pair is treated as consistent only when a participant made wrong choices in all 10 trials for the pair. We calculated the proportion of pairs with consistent errors for each participant (Fig.2h). The proportions under different thresholds were also calculated. For the reference level, we use a threshold of 80%, which still represented high consistency, a conservative threshold that did not compromise our findings.

Ranking self-consistency. After categorizing the pair-wise judgement, we can symbolically

examine whether the ranking judgements of a participant are self-consistent. Any item triad can fall into one of the following three categories (Fig. 3a) by the consistency of the three judgements: correct triad (e.g., A < B, B < C, A < C; in line with A < B < C); self-consistent but incorrect triad (e.g., A < B, B > C, A < C; in line with A < C < B), and self-inconsistent and incorrect triad (e.g., A < B, B < C, A > C; violating transitivity). We quantified ranking consistency through the self-consistency coefficient, defined as:

$$1 - \frac{N_d}{N_T} \tag{1}$$

where N_d represents the number of self-inconsistent and incorrect triads; $N_T = 20$ denotes the maximum possible circular triads for 8 items. The maximum number of the circular triads in a set of *n* elements is defined by two distinct formulas, depending on whether n is odd or even: $(n^3 - n)/24$ and $(n^3 - 4n)/24$, respectively (Moran, 1947). The self-consistency coefficient ranges from 0 (complete inconsistency) to 1 (perfect consistency).

Permutation test for self-consistency coefficient is comparing the group-level self-consistency coefficient with a surrogation distribution. We broke the self-consistency by shuffling the pairwise accuracy matrix of each participant. The self-consistency coefficient was then calculated based on the shuffled accuracy matrix. After 1000 permutations, we compared the group-level self-consistency with the surrogation distribution.

The ranking of a participant was classified as globally self-consistent ranking if and only if its self-consistency coefficient equaled 1 (i.e., $N_d = 0$). This strict criterion follows from transitive logic: any single circular triad (Fig. 3a, right panel) suffices to establish global inconsistency, as larger cycles (e.g., A < B < C < D < A) necessarily contain at least one such circular triad (e.g., A < B < C < A).

Ranking reconstruction. For participants with self-consistent rankings, a single deterministic global ranking can be directly derived from their pairwise judgments. For those participants, we inferred their subjective rankings using the Hodge-rank method (Xu et al., 2012). The Hodge-rank method aims to find optimal global ranking scores *s* for each item by minimizing the following objective function:

$$\min_{s \in R^8} \sum_{i,j} \left(s_i - s_j - \hat{Y}_{ij} \right)^2 \tag{2}$$

where $\hat{Y}_{ij}$ represents the average preference score across all 10 trials for each pair (i, j). For each trial, $\hat{Y}_{ij} = 1$ if the participant chose item i over item j, and $\hat{Y}_{ij} = -1$ if the participant chose item j over item i. The individual ranking for each participant was determined by sorting the items according to their computed ranking scores s.

Inter-subject ranking similarity. To quantify inter-subject ranking similarity, we computed Kendall's tau correlation coefficient between the global ranks of every two participants of all 69 participants. To compare ranking consistency between empirical data and computational modeling,

we statistically contrasted two distributions: (1) the inter-subject similarity distribution derived from participants' reconstructed global ranks, and (2) the corresponding distribution generated by model simulations, using independent samples t-tests.

Preregistration

We pre-registered the behavioral experiment on the Open Science Framework prior to data collection (https://osf.io/dpxq4). In this pre-registration, we specified our novel behavioral paradigm, target sample size, measured variables, and statistical models. The preregistered analysis applied to the MEG experiment; while the MEG experiment itself was not preregistered, its results confirmed our pre-registered hypothesis.

Prior to formal hypothesis pre-registration, we conducted an exploratory behavioral experiment (*N*=35, Peking University participants) using the same learning phase procedure as would later be preregistered, but with a reduced testing phase assessing each pair only 5 times. This preliminary study informed our pre-registered hypothesis regarding consistent errors and guided sample size determination.

Modeling

Q-learning model. In Q-learning model (Ciranka et al., 2022; Rescorla, 1972), we assume that participants update the ranking value of each item using a simple delta rule during learning phase. Specifically, values are updated upon presentation of a learned pair. During the learning phase, when a pair (m, n) is presented, their values (Q) are updated as follows:

$$Q_m \leftarrow Q_m + \alpha[D(m,n) - (Q_m - Q_n)]/2, \tag{3}$$

$$Q_n \leftarrow Q_n + \alpha[D(n,m) - (Q_n - Q_m)]/2, \tag{4}$$

where free parameter α denotes the learning rate ($0 < \alpha \leq 1$). The presented relative difference $D(m,n)$ between the m-th-ranked item and n-th-ranked item is quantified as:

$$D(m,n) = (m-n)/7, \tag{5}$$

The initial values of all Qs were set to 0.

During the testing phase, when the pair (i, j) is presented, the item i is selected with a probability defined by a softmax function with an inverse temperature parameter $\gamma(\gamma \geq 0)$:

$$p_{i,j}(i) = \frac{1}{1+\exp\left(-\gamma\left(Q_i - Q_j\right)\right)} \tag{6}$$

where $Q(\cdot)$ denotes the final $Q_t(\cdot)$ after the learning phase.

Beta-Q model. Beta-Q model involves beta distribution updating in learning (Jensen et al., 2015). The ranking value of each item was assumed to follow a beta distribution, $\mathrm{Beta}(U,L)$, with two shape parameters, U and L for each item. The probability density function of $\mathrm{Beta}(U,L)$:

$$\mathrm{Beta}(x;U,L) = \frac{\Gamma(U+L)}{\Gamma(U)\Gamma(L)} x^{U-1}(1-x)^{L-1} \tag{7}$$

Before learning, the initial values of U_i and L_i were set to 1 as $\mathrm{Beta}(1,1)$ is a uniform distribution

between 0 and 1. When a pair (m, n) is presented at the learning phase (m>n), the ranking value distributions are updated by adjusting U and L as follows:

$$U_m \leftarrow U_m + \left[D(m,n) - \Delta V(m,n)\right], \tag{8}$$

$$L_m \leftarrow L_m + \alpha\left[D(n,m) - \Delta V(n,m)\right], \tag{9}$$

$$U_n \leftarrow U_n + \alpha\left[D(n,m) - \Delta V(n,m)\right], \tag{10}$$

$$L_n \leftarrow L_n + \left[D(m,n) - \Delta V(m,n)\right], \tag{11}$$

where α is a bias factor toward boundary ($0 < \alpha \leq 1$); $\Delta V(m,n)$ is the expected value difference between m and n,

$$\Delta V(m,n) = V_m - \mathrm{V_n}, \tag{12}$$

where $\mathrm{V} = U/(\mathrm{U} + L)$ is the expected value of an item.

During the testing phase, when the pair (i, j) is presented, the judgment is made by independently sampling values from $\mathrm{Beta}(U_i, L_i)$ and $\mathrm{Beta}\left(U_j, L_j\right)$ for item i and j, respectively, and then comparing the two drawn values. Thus, the probability of choosing item i over j is given by

$$p_{i,j}(i) = \int_0^1 \mathrm{Beta}(x|U_i, L_i) \int_0^x \mathrm{Beta}\left(y\middle|U_j, L_j\right) \mathrm{d}y\, \mathrm{d}x. \tag{13}$$

Beta-sort model. The beta-sort model is an extension of beta-Q model (Jensen et al., 2015). The additional assumption beta-sort model is that the learning, or value updating, not only takes place at the items being presented but also at those which are not presented. When a pair (m, n) is presented (m>n), the ranking value distribution of item m and n are updated as in beta-Q (Equation 5-8). For items that are not presented, their updating is as follows.

For ranking value distributions whose expected values fall between V_m and V_n, their variances shrink while their means keep the same, as the current observation confirms the values. For any item o which satisfies $V_n < V_o < V_m$,

$$U_o \leftarrow U_o + V_o, \tag{14}$$

$$L_o \leftarrow L_o + 1 - V_o. \tag{15}$$

For items with lower expected ranking value, their distributions are pushed downward. For any item o which satisfies $V_o < V_n$,

$$U_o \leftarrow U_o + \Delta V(m,n), \tag{16}$$

$$L_o \leftarrow L_o + \alpha\Delta V(m,n). \tag{17}$$

Similarly, for items with higher expected ranking value, their distributions are pushed upward. For any item o which satisfies $V_o > V_m$,

$$U_o \leftarrow U_o + \alpha\Delta V(m,n), \tag{18}$$

$$L_o \leftarrow L_o + \Delta V(m,n). \tag{19}$$

Choice policy is the same as beta-Q model.

Model fitting and simulation. We fitted the models to participants' trial-by-trial choices in the testing phase through maximum likelihood estimation for each participant. Optimization was implemented with *optimize.minimize*() function of *SciPy* package in Python. Using the best-fitting parameters and the learning sequence for each participant, we subsequently randomly generated participants' trial-by-trial choices. Behavioral analyses were conducted on these model-simulated responses the same as on participants data, enabling direct quantitative comparison between the model's predictions and the observed human behavior.

MEG Experimental Design

The MEG experiment comprised three sequential phases conducted under continuous MEG recording: a pre-learning testing phase, a few-shot learning phase, and a post-learning testing phase. We used a different set of eight images, identical across all participants.

The learning phase replicated the behavioral experiment design exactly. Participants learned artificial rankings of the eight items, presented as movie popularity rankings based on box office performance. Each participant was assigned a unique ranking sequence and learned the same set of eight non-adjacent pairs through identical few-shot exposure (four presentations per pair across four blocks).

The experiment included two testing phases that employed neural decoding paradigms. In the pre-learning testing phase, participants judged pairs based on real-world size rankings of the objects depicted in the eight images. In the post-learning testing phase, participants judged the same pairs based on the newly learned artificial rankings of these items. This design enabled direct comparison between neural representations of innate knowledge (real-world size) and newly acquired subjective rankings.

Both testing phases employed identical trial structures optimized for MEG analysis. Each trial began with a 400 ms central fixation cross, followed by a 700 ms image presentation. After a maintenance period with central fixation (800-1000 ms), participants judged whether the current image ranked higher or lower than the previous trial's image. This sequential comparison paradigm allowed examination of neural representations during image processing while maintaining task engagement.

Each testing phase consisted of five blocks containing 81 images each, yielding 80 pairwise judgments per block. To ensure balanced stimulus presentation, we generated fifty randomized sequences of the eight pictures and concatenated them to form the complete testing sequence for

each phase.

MEG Data Acquisition and Preprocessing

MEG recordings were acquired using a whole-head Elekta Neuromag system with 204 planar gradiometers and 102 magnetometers, housed in a magnetically shielded environment. Data were sampled at 1,000 Hz.

Visual stimuli were presented through a rear-projection system with a 32-inch display screen positioned 75 cm from participants. The system delivered stimuli at full HD resolution (1,920 × 1,080 pixels) with a 60 Hz refresh rate. The experiment was implemented in MATLAB using Psychtoolbox, and participant responses were recorded using a dedicated input device.

Data preprocessing was conducted using MNE-Python. Signal Space Separation (Maxfilter) was first applied for noise reduction and motion artifact correction. Subsequent preprocessing steps included bandpass filtering (1-40 Hz), downsampling to 200 Hz, and epoch extraction from -200 ms to 1500 ms relative to stimulus onset. After independent component analysis using the FastICA algorithm, Ocular and cardiac components were manually identified and removed.

Representational Similarity Analysis

We employed representational similarity analysis (RSA) to examine the neural encoding of both real-world size rankings and subject-specific subjective rankings across space and time. This approach allowed us to test whether neural activity patterns reflected shared knowledge versus individualized ranking representations. All analyses were performed using Python *RSAtoolbox* (v.3.0) and MNE-Python.

Neural RDM construction. Neural representational dissimilarity matrices (RDMs) were computed for each participant at every timepoint within epochs spanning -200 ms to 1500 ms relative to stimulus onset. For each of the eight items, we extracted spatial patterns of neural activity across all 204 gradiometer sensors. Neural dissimilarity between any two items was quantified as 1 minus the Pearson correlation coefficient between their corresponding spatial activity patterns. This yielded participant-specific 8×8 neural RDMs at each timepoint, where each element represented the neural dissimilarity between a pair of items. The resulting RDMs were symmetric matrices with diagonal elements excluded from subsequent regression analyses, since identical items cannot exhibit meaningful dissimilarity with themselves.

Model RDM construction. Two distinct model RDMs served as regression predictors. The size-rank model RDM was identical across all participants, computed from absolute differences in canonical object size rankings (1-8) that reflected shared real-world knowledge about physical object sizes. The subjective-rank model RDMs were uniquely generated for each participant based on their individually reconstructed ranking sequences. This design enabled us to distinguish between neural representations of common knowledge versus personalized subjective orderings.

Regression analysis procedure. At each timepoint, we performed multiple linear regression where each participant's neural RDM served as the dependent variable and both model RDMs (size-rank and subjective-rank) served as independent predictors. This regression yielded two β coefficients per participant per timepoint: β representing the strength of neural encoding for real-world size rankings, and β representing the strength of neural encoding for subjective rankings.

For real-world size representations, we quantified neural encoding as the average of β coefficients across both pre-learning and post-learning phases, reflecting the persistent presence of this innate knowledge throughout the experiment. For subjective ranking representations, we computed the difference between post-learning and pre-learning β coefficients (post-learning β minus pre-learning β), isolating neural changes specifically attributable to few-shot learning while using pre-learning activity as a baseline control.

Statistical analysis. Statistical significance was assessed through non-parametric cluster-based permutation testing involving 100,000 sign-flips of regression coefficients to generate null distributions for population-level mean beta values. We applied cluster-based permutation procedures with a cluster-defining threshold of $p < 0.05$ to address multiple comparisons across timepoints (Maris & Oostenveld, 2007).

Spatial distribution analysis. We conducted sensor-level searchlight RSA to examine the spatial distribution of neural representations. For each sensor, we constructed neural RDMs by incorporating activity patterns from the target sensor and its neighbors (mean 8 sensors, range 4-11), averaged across significant time windows identified in whole-brain RSA (245-265 ms for size-rank; 685-775 ms for subjective-rank). These sensor-level neural RDMs were regressed against both model RDMs to generate participant-specific spatial β-maps.

Validation of subject-specific patterns. To confirm that significant neural representations were genuinely tied to individual participants' unique subjective rankings, we performed a validation analysis. We systematically shuffled subjective-rank assignments across participants while preserving original neural data. For each of 1,000 permutation iterations, we randomly reassigned subjective-rank models among participants and regressed neural signals averaged across the significant time window (685-775 ms) against the shuffled models. This procedure generated a null distribution of β values against which we statistically compared the originally observed mean β coefficients, confirming whether identified neural representations were participant-specific rather than due to chance alignments.

References

Acuna, B. D., Eliassen, J. C., Donoghue, J. P., & Sanes, J. N. (2002). Frontal and parietal lobe activation during transitive inference in humans. *Cerebral Cortex*, *12*(12), 1312-1321. https://doi.org/10.1093/cercor/12.12.1312

Acuna, B. D., Sanes, J. N., & Donoghue, J. P. (2002). Cognitive mechanisms of transitive inference. *Experimental Brain Research*, *146*(1), 1-10. https://doi.org/10.1007/s00221-002-1092-y

Brunamonti, E., Mione, V., Di Bello, F., Pani, P., Genovesio, A., & Ferraina, S. (2016). Neuronal Modulation in the Prefrontal Cortex in a Transitive Inference Task: Evidence of Neuronal Correlates of Mental Schema Management. *J Neurosci*, *36*(4), 1223-1236. https://doi.org/10.1523/jneurosci.1473-15.2016

Bryant, P. E., & Trabasso, T. (1971). Transitive Inferences and Memory in Young Children. *Nature*, *232*(5311), 456-&. https://doi.org/DOI 10.1038/232456a0

Buschman, T. J., Siegel, M., Roy, J. E., & Miller, E. K. (2011). Neural substrates of cognitive capacity limitations. *Proceedings of the National Academy of Sciences*, *108*(27), 11252-11255. https://doi.org/10.1073/pnas.1104666108

Ciranka, S., Linde-Domingo, J., Padezhki, I., Wicharz, C., Wu, C. M., & Spitzer, B. (2022). Asymmetric reinforcement learning facilitates human inference of transitive relations. *Nature Human Behaviour*, *6*(4), 555-564. https://doi.org/10.1038/s41562-021-01263-w

Cowan, N. (2010). The Magical Mystery Four. *Current Directions in Psychological Science*, *19*(1), 51-57. https://doi.org/doi: 10.1177/0963721409359277

Daisley, J. N., Vallortigara, G., & Regolin, L. (2021). Low-rank chicks are better at transitive inference reasoning. *Communications Biology*, *4*(1). https://doi.org/ARTN 1344 10.1038/s42003-021-02855-y

Davis, H. (1992). Transitive inference in rats (Rattus norvegicus). *Journal of Comparative Psychology*, *106*(4), 342-349. https://doi.org/10.1037/0735-7036.106.4.342

Dehaene, S., Al Roumi, F., Lakretz, Y., Planton, S., & Sablé-Meyer, M. (2022). Symbols and mental programs: a hypothesis about human singularity. *Trends in Cognitive Sciences*, *26*(9), 751-766. https://doi.org/10.1016/j.tics.2022.06.010

Delius, J. D., & Siemann, M. (1998). Transitive responding in animals and humans: Exaptation rather than adaptation? *Behavioural Processes*, *42*(2), 107-137. https://doi.org/https://doi.org/10.1016/S0376-6357(97)00072-7

Di Antonio, G., Raglio, S., & Mattia, M. (2024). A geometrical solution underlies general neural principle for serial ordering. *Nature Communications*, *15*(1). https://doi.org/10.1038/s41467-024-52240-6

Di Bello, F., Mione, V., Pani, P., Brunamonti, E., & Ferraina, S. (2024). Prefrontal cortex contribution in transitive inference task through the interplay of beta and gamma oscillations. *Communications Biology*, *7*(1). https://doi.org/10.1038/s42003-024-07418-5

Gillan, D. J. (1981). Reasoning in the chimpanzee: II. Transitive inference. *Journal of Experimental Psychology: Animal Behavior Processes*, *7*(2), 150-164. https://doi.org/10.1037/0097-7403.7.2.150

Hinton, E. C., Dymond, S., von Hecker, U., & Evans, C. J. (2010). Neural correlates of relational reasoning and the symbolic distance effect: involvement of parietal cortex. *Neuroscience*, *168*(1), 138-148. https://doi.org/https://doi.org/10.1016/j.neuroscience.2010.03.052

Iñiguez, G., Pineda, C., Gershenson, C., & Barabási, A.-L. (2022). Dynamics of ranking. *Nature Communications*, *13*(1), 1646. https://doi.org/10.1038/s41467-022-29256-x

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. https://doi.org/10.1371/journal.pcbi.1004523

Jensen, G., Munoz, F., Meaney, A., Terrace, H. S., & Ferrera, V. P. (2021). Transitive inference after minimal training in rhesus macaques (Macaca mulatta). *J Exp Psychol Anim Learn Cogn*, *47*(4), 464-475. https://doi.org/10.1037/xan0000298

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

Jin, Y., Jensen, G., Gottlieb, J., & Ferrera, V. (2022). Superstitious learning of abstract order from random reinforcement. *Proceedings of the National Academy of Sciences*, *119*(35), e2202789119. https://doi.org/doi:10.1073/pnas.2202789119

Kay, K., Biderman, N., Khajeh, R., Beiran, M., Cueva, C. J., Shohamy, D., Jensen, G., Wei, X.-X., Ferrera, V. P., & Abbott, L. F. (2024). Emergent neural dynamics and geometry for generalization in a transitive inference task. *PLOS Computational Biology*, *20*(4), e1011954. https://doi.org/10.1371/journal.pcbi.1011954

Lake, B. M., Salakhutdinov, R., & Tenenbaum, J. B. (2015). Human-level concept learning through probabilistic program induction. *Science*, *350*(6266), 1332-1338. https://doi.org/10.1126/science.aab3050

Li, V., Herce Castañón, S., Solomon, J. A., Vandormael, H., & Summerfield, C. (2017). Robust averaging protects decisions from noise in neural computations. *PLOS Computational Biology*, *13*(8), e1005723. https://doi.org/10.1371/journal.pcbi.1005723

Lippl, S., Kay, K., Jensen, G., Ferrera, V. P., & Abbott, L. F. (2024). A mathematical theory of relational generalization in transitive inference. *Proceedings of the National Academy of Sciences*, *121*(28), e2314511121. https://doi.org/10.1073/pnas.2314511121

Mannella, F., & Pezzulo, G. (2024). Transitive inference as probabilistic preference learning. *Psychonomic Bulletin & Review*, *32*(2), 674-689. https://doi.org/10.3758/s13423-024-02600-6

Maris, E., & Oostenveld, R. (2007). Nonparametric statistical testing of EEG- and MEG-data. *Journal of Neuroscience Methods*, *164*(1), 177-190. https://doi.org/https://doi.org/10.1016/j.jneumeth.2007.03.024

McGonigle, B. O., & Chalmers, M. (1977). Are monkeys logical? *Nature*, *267*(5613), 694-696. https://doi.org/10.1038/267694a0

Miconi, T., & Kay, K. (2025). Neural mechanisms of relational learning and fast knowledge reassembly in plastic neural networks. *Nature Neuroscience*, *28*(2), 406-414. https://doi.org/10.1038/s41593-024-01852-8

Mitchell, T. M. (2007). The Need for Biases in Learning Generalizations.

Moran, P. A. (1947). On the method of paired comparisons. *Biometrika*, *34*(Pt 3-4), 363-365.

Munoz, F., Jensen, G., Kennedy, B. C., Alkan, Y., Terrace, H. S., & Ferrera, V. P. (2020a). Learned Representation of Implied Serial Order in Posterior Parietal Cortex. *Scientific Reports*, *10*(1). https://doi.org/10.1038/s41598-020-65838-9

Munoz, F., Jensen, G., Kennedy, B. C., Alkan, Y., Terrace, H. S., & Ferrera, V. P. (2020b). Learned Representation of Implied Serial Order in Posterior Parietal Cortex. *Scientific Reports*, *10*(1), 9386. https://doi.org/10.1038/s41598-020-65838-9

Nelli, S., Braun, L., Dumbalska, T., Saxe, A., & Summerfield, C. (2023). Neural knowledge assembly in humans and neural networks. *Neuron*, *111*(9), 1504-1516.e1509. https://doi.org/10.1016/j.neuron.2023.02.014

Palminteri, S., Lefebvre, G., Kilford, E. J., & Blakemore, S.-J. (2017). Confirmation bias in human reinforcement learning: Evidence from counterfactual feedback processing. *PLOS Computational Biology*, *13*(8), e1005684. https://doi.org/10.1371/journal.pcbi.1005684

Piaget, J. (1923). La pensée symbolique et le pensée de l'enfant. [Symbolic thought and the thought of the child.]. *Archives de Psychologie*, *18*, 273-304.

Ramawat, S., Marc, I. B., Ceccarelli, F., Ferrucci, L., Bardella, G., Ferraina, S., Pani, P., & Brunamonti, E. (2023). The transitive inference task to study the neuronal correlates of memory-driven decision making: A monkey neurophysiology perspective. *Neuroscience & Biobehavioral Reviews*, *152*, 105258. https://doi.org/https://doi.org/10.1016/j.neubiorev.2023.105258

Ramawat, S., Mione, V., Di Bello, F., Bardella, G., Genovesio, A., Pani, P., Ferraina, S., & Brunamonti, E. (2022). Different Contribution of the Monkey Prefrontal and Premotor Dorsal Cortex in Decision Making During a Transitive Inference Task. *Neuroscience*, *485*, 147-162. https://doi.org/https://doi.org/10.1016/j.neuroscience.2022.01.013

Rescorla, R. (1972). A theory of Pavlovian conditioning : Variations in the effectiveness of reinforcement and nonreinforcement.

Sorscher, B., Ganguli, S., & Sompolinsky, H. (2022). Neural representational geometry underlies few-shot concept learning. *Proceedings of the National Academy of Sciences*, *119*(43), e2200800119. https://doi.org/10.1073/pnas.2200800119

Tenenbaum, J. B., Kemp, C., Griffiths, T. L., & Goodman, N. D. (2011). How to Grow a Mind: Statistics, Structure, and Abstraction. *Science*, *331*(6022), 1279-1285. https://doi.org/10.1126/science.1192788

Treichler, F. R., & Van Tilburg, D. (1996). Concurrent conditional discrimination tests of transitive inference by macaque monkeys: List linking. *Journal of Experimental Psychology: Animal Behavior Processes*, *22*(1), 105-117. https://doi.org/10.1037/0097-7403.22.1.105

Van Opstal, F., Fias, W., Peigneux, P., & Verguts, T. (2009). The neural representation of extensively trained ordered sequences. *NeuroImage*, *47*(1), 367-375. https://doi.org/https://doi.org/10.1016/j.neuroimage.2009.04.035

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

Wei, C. A., Kamil, A. C., & Bond, A. B. (2014). Direct and Relational Representation During

Transitive List Linking in Pinyon Jays (

). *Journal of Comparative Psychology*, *128*(1), 1-10. https://doi.org/10.1037/a0034627

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

Xu, B., Wu, J., Xiao, H., Münte, T. F., & Ye, Z. (2024). Inferior parietal cortex represents relational structures for explicit transitive inference. *Cerebral Cortex*, *34*(4). https://doi.org/10.1093/cercor/bhae137

Xu, Q., Huang, Q., Jiang, T., Yan, B., Lin, W., & Yao, Y. (2012). HodgeRank on Random Graphs for Subjective Video Quality Assessment. *IEEE Transactions on Multimedia*, *14*(3), 844-857. https://doi.org/10.1109/TMM.2012.2190924