Decomposing dynamical subprocesses for compositional generalization

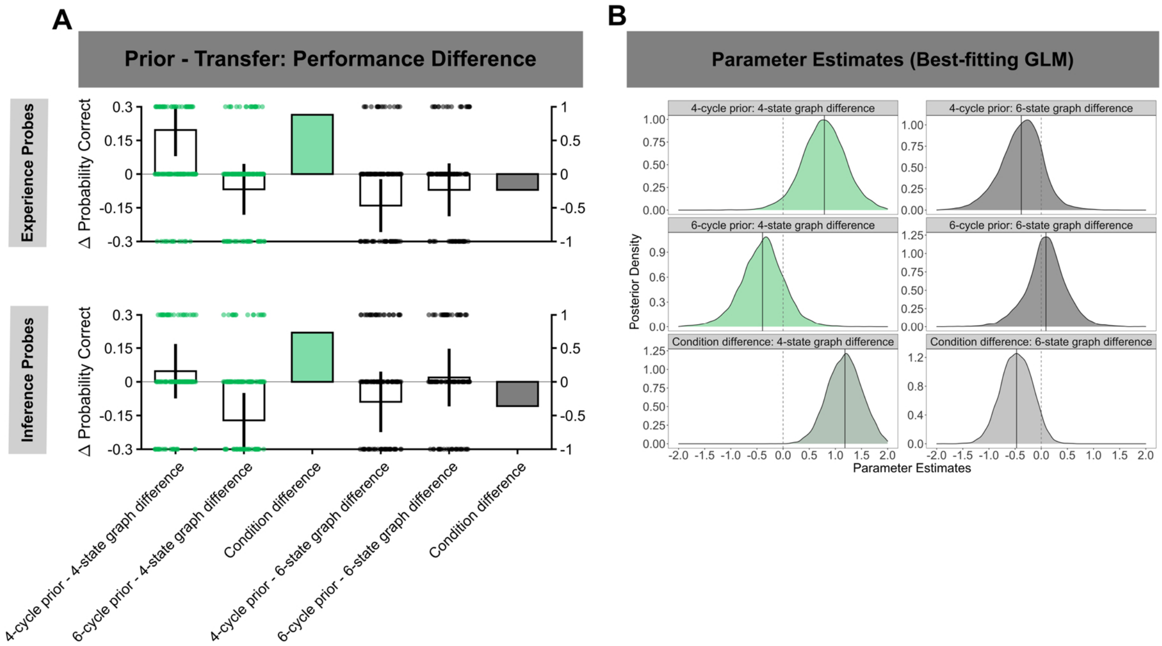
Supplementary Material

Supplementary Results

First trial analyses

Building on the results in the main manuscript, we reasoned that there should be a within-subjects performance benefit in the initial trials of transfer learning – compared to the first trials of the prior learning phase – specifically for graph factors consistent across task phases. This is because structure consistent across prior and transfer learning should be available immediately and exclusively during the initial trials of transfer learning. To test this idea, we compared performance on the first occurrence of each respective trial type during prior and transfer learning, for both conditions separately. We observed greater performance change from prior to transfer learning on the 4-state graph factor for participants in the 4-cycle prior condition relative to the 6-cycle prior condition on both experience (Supplementary Fig. 1A, top panel, green filled bar) and inference probes (Supplementary Fig. 1A, bottom panel, green filled bar). The opposite was the case for the 6-state graph factor: Here, participants in the 4-cycle prior condition exhibited a more negative performance change from prior to transfer learning than participants in the 6-cycle prior condition on both experience (Supplementary Fig. 1A, top panel, black filled bar) and inference probes (Supplementary Fig. 1A, bottom panel, black filled bar). To formally test this, we employed two different multilevel Bayesian GLM (GLM7-8, Supp. Eq. 1-2). The most parsimonious model (GLM7) featured a non-zero interaction effect between trial type, condition and phase ($\mu_{TTxCONDxPHASE} = .91$, CI = [.55; 1.27]). In post-hoc tests, we found that the interaction effect was driven by higher prior to transfer learning accuracy change in 4-state cycle probes in the 4-cycle prior condition (vs 6-cycle prior condition, mean difference of 10,000 posterior samples: 1.18, CI = [.66; .1.72], Supplementary Fig. 1B, bottom left panel) and by greater performance change from prior to transfer learning in 6-state cycle probes in the 6-state cycle prior condition (vs 4-state cycle prior condition, mean difference of 10,000 posterior samples: -.47, CI = [-.95; -.01], Supplementary Fig. 1B, bottom right panel). The credible interval around the posterior mean difference of the difference: 1.65 [.95; 2.38] did not include 0.

The better fit of a model considering condition and trial type specific changes in accuracy across task phases, coupled with the observed non-zero interaction effect and post-hoc contrasts of parameter estimates, suggests that participants had specific performance advantages for the graph factor when this was consistent across prior and transfer learning even in the very first trials of transfer learning. Together with the analysis of the first trial during transfer learning, these results provide convergent support for a proposal that participants



Supplementary Figure 1. Performance difference in initial trials across learning phases

A) Comparison of performance in the first occurrence of each respective trial type during both prior and transfer learning phase (difference transfer – prior). Each dot represents one participant (right y-axis), bars represent the arithmetic mean of the distribution, and error bars depict the 95% confidence interval around the mean (left y-axis).
B) Posterior density plots for each parameter estimate from each individual level of the interaction effect. The dashed vertical line represents a zero effect. The credible interval around the posterior mean difference of the difference (bottom row): 1.65 [.95; 2.38] did not include 0.

decomposed and abstracted the structure underlying their experience during prior learning and reused this acquired structural knowledge during transfer learning. This efficient form of knowledge representation supports faster (re)discovery of graph factors where these are consistent across contexts, even with very limited exposure to novel environments.

Analysis of transitions unexpected under structural priors

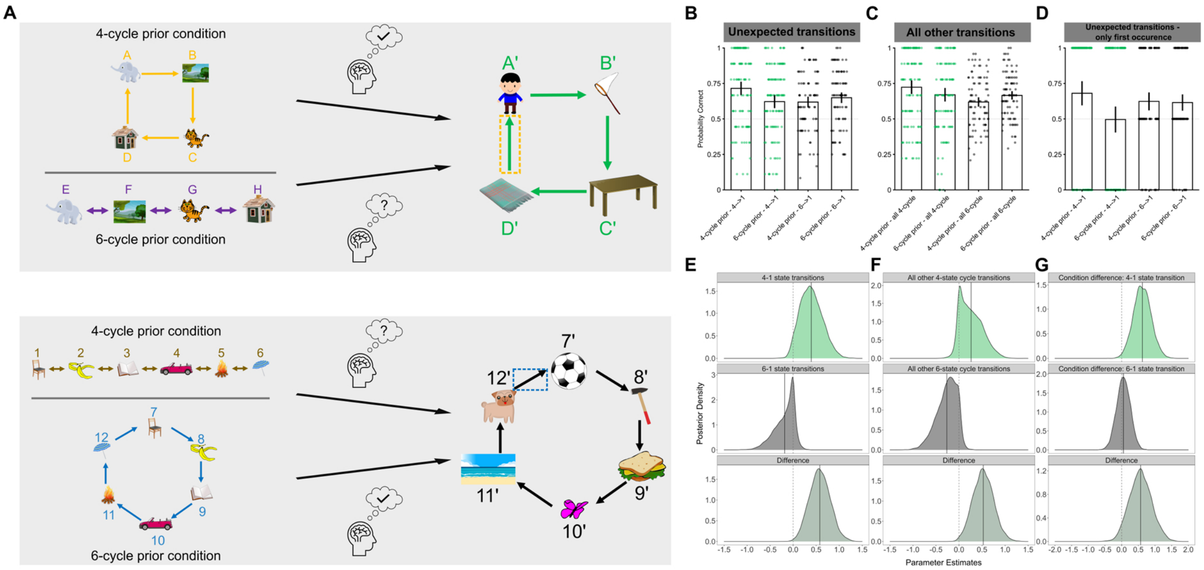
If participants abstract and compositionally reuse prior knowledge, performance between conditions and graph factors should not only differ across all trials of transfer learning. More specifically, we would predict performance differences across conditions in probe trials testing transitions that are not expected under the respective structural form experienced during prior learning. For example, in probe trials testing the 6-state graph factor, participants in the 4-state cycle prior condition would not expect to see transitions from the 6th state to the 1st state on the 6-state cycle during transfer learning, considering that they had only ever experienced a transition back to the 5th state in the 6-state graph factor during prior learning (Supplementary Fig. 2A, bottom). Hence, for these specific transitions, accuracy should be lower than for the 6-state cycle prior condition – that has accrued experience with transitions from the 6th state to the 1st state of this structural form. Contrarily, participants in the 6-state cycle prior condition would not expect to see transitions from the 4th state to the 1st state on the 4-state cycle during transfer learning, given that they had only experienced transitions back to the 3rd state in the 4-state graph factor during prior learning (Supplementary Fig. 2A, top). Therefore, accuracy for these specific transitions should be reduced.

We used two sets of analyses to test these hypotheses. First, we modelled the probability of correct answers using three different multilevel GLMs (GLM9-11, Supp. Eq. 3-5). Model comparison indicated that the data was best explained by GLM11, featuring a parameter with 2x2x2 (8 total) levels, estimating a posterior distribution for each possible combination of transition type and transition expectation, specific for both conditions. In post-hoc tests, we found that in unexpected transitions, the probability of correct answers on 4->1 transitions was enhanced in the 4-state cycle prior condition (vs 6-state cycle prior condition, mean difference of 10,000 posterior samples: .39, CI = [.03; .80], Supplementary Fig. 2B, green and 2E, top panel). Contrarily, the probability of correct answers on 6->1 transitions was only descriptively larger in the 6-state cycle prior condition (vs 4-state cycle prior condition, mean difference of 10,000 posterior samples: -.18, CI = [-.61; .07], Supplementary Fig. 2B, black and 2E, middle panel). Importantly, the credible interval around the posterior mean difference of the difference: .58 [.22; .94] did not include 0 (Supplementary Fig. 2E, bottom panel).

When investigating the same contrasts for expected transitions, the probability of correct answers on all other 4-cycle transitions (but 4->1 transitions) was only descriptively larger in the 4-state cycle prior condition (vs 6-state cycle prior condition, mean difference of 10,000 posterior samples: .26, CI = [-.04; .69], Supplementary Fig. 2C and F, top panel). The probability of correct answers on all other 6-state cycle predictions (but 6->1 transitions) was also only descriptively enhanced in the 6-state cycle prior condition (vs 4-state cycle prior condition, mean difference of 10,000 posterior samples: -.26, CI = [-.67; .02], Supplementary Fig. 2C and F, middle panel). Also, in the case of expected transitions, the credible interval around the posterior mean difference of the difference: .53 [.17; .88] did not include 0 (Supplementary Fig. 2F, lower panel). The transition type specific condition difference contrasts between unexpected and expected transitions (i.e., Supplementary Fig. 2E and F, bottom panels) did not show differences. The credible interval for the difference of expected and unexpected transitions between the condition difference for 4-state cycle probe trials: .13 [-.11; .38], as well as for the difference of expected and unexpected transitions between the condition difference for 6-state cycle probe trials: .08 [-.12; .28], included 0. These results suggest that the observed condition differences across trial types were present in the case of both unexpected and expected transitions.

We next reasoned that the transition type specific differences on unexpected transitions across conditions should be present already during the initial probing of each transition type during the transfer learning phase – under the hypothesis that participants abstracted and reused decomposed prior knowledge of graph factors during transfer learning. We examined performance during the trials at which each unexpected transition was probed for the first time. We observed a reduction of performance on probe questions testing transitions from the 4th state to the 1st state in participants in the 6-state cycle prior condition, relative to participants in the 4-state cycle prior condition (Supplementary Fig. 2D, green). However, there was no such effect for probe questions testing transitions from the 6th state to the 1st state. Participants in both prior learning conditions performed equally well (Supplementary Fig. 2D, black). GLM13, the best-fitting model in this analysis, featured a non-zero interaction effect between transition type and condition ($\mu_{TTxCOND}$ = .38, CI = [.08; .67]). Post-hoc tests indicated a higher probability of correct answers on 4->1 transitions in the 4-state cycle prior condition (mean difference of 10,000 posterior samples: .62, CI = [.20; 1.05], Supplementary Fig. 4D and G, top panel), but no difference between conditions for the probability of correct answers on 6->1 transitions (mean difference of 10,000 posterior samples: .05, CI = [-.27; .38], Supplementary Fig. 2D and G, middle panel). The credible interval around the posterior mean difference of the condition differences: 0.57 [.04; 1.11] did not include 0 (Supplementary Fig. 2G, bottom panel).

Taken together, we observed better fit and non-zero interaction effects in models considering condition specific disadvantages for transitions that are not expected under consideration of a structural form included in prior learning. These patterns already emerged at first occurrence of probe questions involving the unexpected transitions and strengthen the notion that participants transferred specific decomposed and abstracted structural knowledge, to be reused in an entirely novel situation.



Supplementary Figure 2. Performance in transitions unexpected under a specific prior

A) Schematic of the predictions for performance differences in unexpected transitions: The probability of giving a correct answer should be reduced in a condition-specific manner for transitions that are not expected under a structural prior acquired during prior learning.

B) Probabilities of giving correct answers in unexpected transition probe trials, broken up for transition trial type (4->1 state transitions (green) and 6->1 state transitions (black)) and condition, across all probe trials during transfer learning.

C) Probabilities of giving correct answers in all other transition probe trials (excluding unexpected transitions), broken up for transition trial type (4-state cycle probe trials (green) and 6-state cycle probe trials (black)) and condition, across all probe trials during transfer learning.

D) Probabilities of giving correct answers in unexpected transitions, broken up for transition trial type (4->1 state transitions (green) and 6->1 state transitions (black)) and condition, only during the first occurrence of probing the respective transition. Each dot in B), C), and D) represents the arithmetic mean performance of one participant across experience and inference probe trials for the respective transition, bars represent the arithmetic mean of the distribution, error bars depict the 95% confidence interval around the mean and the dashed gray line represented chance level (probability correct = 0.5).

E) Posterior density plots for parameters of the best-fitting GLM for unexpected transitions predictions. The dashed vertical line represents a zero effect. The credible interval around the posterior mean difference of the difference: .58 [.22; .94] did not include 0 (bottom panel).

F) Posterior density plots for parameters of the best-fitting GLM for all other transitions (excluding unexpected transitions). The dashed vertical line represents a zero effect. The credible interval around the posterior mean difference of the difference: .53 [.17; .88] did not include 0 (lower panel). The transition type specific condition difference contrasts between unexpected and expected transitions (i.e., E and F, bottom panels) did not show differences. The credible interval for the difference between expected and unexpected transitions between conditions difference for 4-state cycle probe trials: .13 [-.11; .38], as well as for the difference between expected and unexpected transitions between the condition difference for 6-state cycle probe trials: .08 [-.12; .28], included 0.

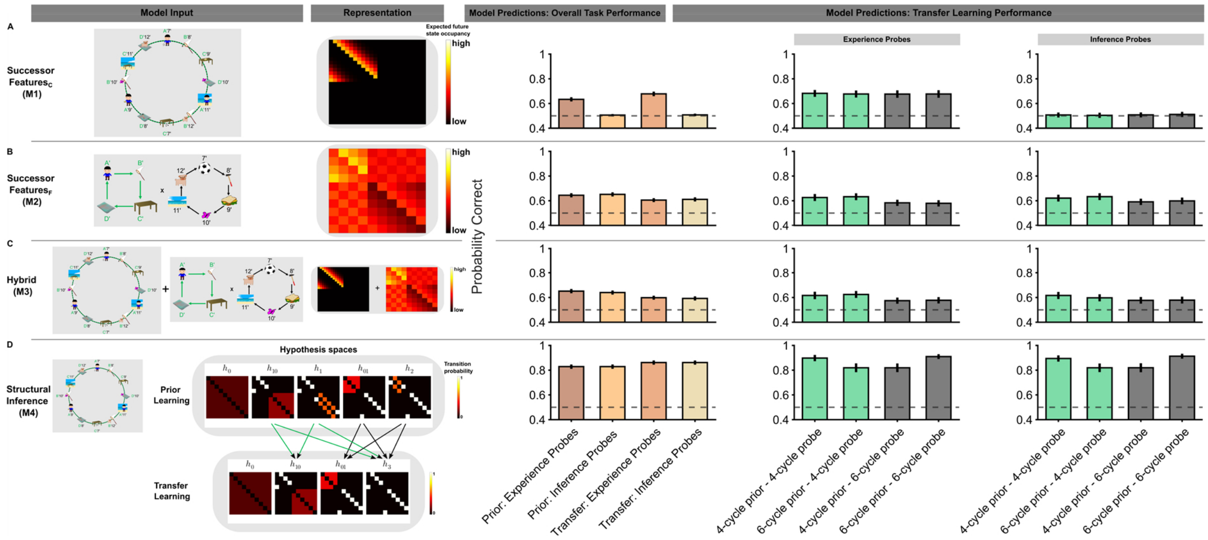
G) Posterior density plots for parameters of the best-fitting GLM for unexpected transition predictions during the initial occurrence of probing each transition type. The dashed vertical line represents a zero effect. The credible interval around the posterior mean difference of the condition differences: 0.57 [.04; 1.11] did not include 0 (bottom panel).

Computational modeling: Apriori simulations

We leveraged apriori simulations using four classes of agents performing the experiment to further detail cognitive strategies that could be used to solve the task (Supplementary Fig. 3). To this end, we simulated 250 model instantiations per computational model class with varying parameters, randomly sampled from beta or exponential distributions (see Methods for details), ensuring a reasonable range of parameter values and matching sampled parameters across prior learning conditions. Models were presented with identical sequences of numbers corresponding to the compound images (or decomposed sequences of numbers representing the features in each compound image) presented to human participants and made choices in relation to both experience and inference probes.

While all successor predictive models (**M1-M3**) performed well on experience probes (Supplementary Fig. 3, Overall Task Performance and Transfer Learning Performance), only **M2** and **M3** showed above-chance performance for inferring transitions between unobserved compositions (inference probes). Crucially, none of the models **M1-3** predicted an overall performance increase from prior learning to transfer learning (Supplementary Fig. 3, Overall Task Performance), nor enhanced transfer learning performance for the graph factor that was consistent across prior and transfer learning (Supplementary Fig. 3, Transfer Learning Performance).

M4 showed performance increases in overall performance from prior to transfer learning (Supplementary Fig. 3D, Overall Task Performance), as well as enhanced transfer learning performance for the graph factor that was consistent across prior and transfer learning in both conditions (Supplementary Fig. 3D, Transfer Learning Performance). These apriori simulation results suggest that a structural inference model was capable of discovering a factorized state space underlying the sequence of compound images and could reuse this prior knowledge compositionally.



Supplementary Figure 3. Representations and apriori computational model simulations

A) A successor feature model (**M1**) learned an expected future state occupancy from the sequence of unique compound states and stored this quantity in a successor feature matrix (Representation column). Each row and column represent one compound state. Depicted is the expected, discounted future compound state occupancy after the completion of all trials in transfer learning. Top rows and leftward columns correspond to experience compounds whereas bottom rows/rightward columns correspond to inference compounds. Note that because inference states are never experienced during sequence learning they contribute 0 successor entries.

B) A successor feature model (**M2**) learned expectations about the sequences of features in each graph factor separately (same updating mechanism as **M1**, but operated over two separate features, for the 4-state and 6-state factor separately). Depicted is the expected, discounted future feature occupancy after the completion of all trials in transfer learning. Each row and column represent one feature. The 4 topmost rows and leftmost columns represent features associated with the 4-cycle factor whereas remaining rows/columns represent features of the 6-cycle factor. Note that because this model does not realize that features are arranged into two separate sequences, non-diagonal blocks contribute non-zero successor entries.

C) The third successor feature model (**M3**) was a hybrid between **M1** and **M2**, involving a trial-to-trial varying weighted average of the degree to which expectations about upcoming states/features in **M1** and **M2** are used to generate choices. **M1**'s and **M2**'s Softmax weights in predicting unobserved compounds ("model confidence") were used to determine trial-wise changes in model weighting.

D) Finally, we examined a structural inference model (**M4**) that uses the experienced sequence of compound images to perform trial-to-trial inferences over a hypothesis space of graph structures (Model Input, right column). **M4** maintains a current structural hypothesis. After each sequence of 12 compound images, this hypothesis is updated and compared to other hypotheses from a hypothesis space of graph structures. The posterior belief about the underlying structure in prior learning is used to update the prior probability of hypotheses in the transfer learning phase if they include a subgraph that has been encountered during prior learning, as denoted by arrows. **M4** predicts increases in overall performance from prior to transfer learning, as well as enhanced transfer learning performance for the graph factor that was consistent across prior and transfer learning in both conditions.

Supplementary Methods

Multilevel models comparing prior and transfer learning performance

To further support the idea that participants reused decomposed prior structural knowledge during transfer learning, we computed performance changes from prior to transfer learning for the first occurrence of each respective trial type, for both conditions. We employed two multilevel Bayesian GLMs (GLM7 - GLM8, Supp. Eq. 1-2) – as in previous analyses jointly analyzing both experience and inference probe trial correct responses – to test if participants showed such advantage during transfer learning relative to prior learning.

GLM7: Multilevel model representing the null hypothesis of no trial type and condition specific performance changes from initial trials of prior to initial trials of transfer learning. The model features individually varying intercepts, main effects of trial type, task phase (*PHASE*) and condition, and covarying intercepts for trial type, task phase effects

$$Correct_{i,k,m,n} \sim Binomial(number\ of\ trials_{i,k,m,n}, p_{i,k,m,n})$$

$$logit(p_{i,k,m,n}) = \mu_{\alpha} + \mu_{TT[i]} + \mu_{PHASE[m]} + \mu_{COND[k]} + \alpha_{subject[n],TT[i]} + \gamma_{subject[n],PHASE[m]}$$

$$\begin{bmatrix} \alpha_{n,1} \\ \alpha_{n,2} \end{bmatrix} \sim MVNormal\left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, S_{TT}\right)$$

$$\begin{bmatrix} \gamma_{n,1} \\ \gamma_{n,2} \end{bmatrix} \sim MVNormal\left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, S_{PHASE}\right)$$

$$S_{TT} = \begin{pmatrix} \sigma_{subject} & 0 \\ 0 & \sigma_{TT} \end{pmatrix} R_{TT} \begin{pmatrix} \sigma_{subject} & 0 \\ 0 & \sigma_{TT} \end{pmatrix}$$

$$S_{PHASE} = \begin{pmatrix} \sigma_{subject} & 0 \\ 0 & \sigma_{PHASE} \end{pmatrix} R_{PHASE} \begin{pmatrix} \sigma_{subject} & 0 \\ 0 & \sigma_{PHASE} \end{pmatrix}$$

(Supp. Eq. 1)

$$\mu_{\alpha} \sim Normal(0,1)$$

$$\mu_{TT[i]} \sim Normal(0,1)$$

$$\mu_{PHASE[m]} \sim Normal(0,1)$$

$$\mu_{COND[k]} \sim Normal(0,1)$$

$$\sigma_{subject} \sim Half-Normal(0,1)$$

$$\sigma_{TT} \sim Half-Normal(0,1)$$

$$\sigma_{PHASE} \sim Half - Normal(0,1)$$

$$R_{TT} \sim LKJcorr(2)$$

$$R_{PHASE} \sim LKJcorr(2)$$

where *m* denotes the parameter estimate for the m-th phase (prior learning or transfer learning)

GLM8: Multilevel model representing the hypothesis that there are trial type and condition specific performance changes from initial trials of prior to initial trials of transfer learning. The model features individually varying intercepts, main effects of trial type, task phase and condition, an interaction effect of trial type x condition x task phase, and covarying intercepts for trial type, task phase effects

$$Correct_{i,k,m,n} \sim Binomial\left(number\ of\ trials_{i,k,m,n}, p_{i,k,m,n}\right)$$

$$logit\left(p_{i,k,m,n}\right) = \mu_{\alpha} + \mu_{TT[i]} + \mu_{PHASE[m]} + \mu_{COND[k]} + \alpha_{subject[n],TT[i]} +$$

$$\gamma_{subject[n],PHASE[m]} + \mu_{TTx\,CONDxPHASE}$$

$$\begin{bmatrix} \alpha_{n,1} \\ \alpha_{n,2} \end{bmatrix} \sim MVNormal\left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, S_{TT}\right)$$

$$\begin{bmatrix} \gamma_{n,1} \\ \gamma_{n,2} \end{bmatrix} \sim MVNormal\left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, S_{PHASE}\right)$$

$$S_{TT} = \begin{pmatrix} \sigma_{subject} & 0 \\ 0 & \sigma_{TT} \end{pmatrix} R_{TT} \begin{pmatrix} \sigma_{subject} & 0 \\ 0 & \sigma_{TT} \end{pmatrix}$$

$$S_{PHASE} = \begin{pmatrix} \sigma_{subject} & 0 \\ 0 & \sigma_{PHASE} \end{pmatrix} R_{PHASE} \begin{pmatrix} \sigma_{subject} & 0 \\ 0 & \sigma_{PHASE} \end{pmatrix} \qquad \textit{(Supp. Eq. 2)}$$

$$\mu_{\alpha} \sim Normal(0,1)$$

$$\mu_{TT[i]} \sim Normal(0,1)$$

$$\mu_{PHASE[m]} \sim Normal(0,1)$$

$$\mu_{COND[k]} \sim Normal(0,1)$$

$$\mu_{TTx\,CONDxPHASE} \sim Normal(0,1)$$

$$\sigma_{subject} \sim Half - Normal(0,1)$$

$$\sigma_{TT} \sim Half - Normal(0,1)$$

$$\sigma_{PHASE} \sim Half - Normal(0,1)$$

$$R_{TT} \sim LKJcorr(2)$$

$$R_{PHASE} \sim LKJcorr(2)$$

GLM8 explained the data better than a model not including the trial type x condition x phase interaction effect (GLM7) (PSIS values (± standard error): GLM8 = 1913.5 (±23.44), GLM7 = 1929.9 (±22.22); WAIC values (± standard error): GLM8 = 1912.8 (±23.42), GLM7= 1929.2 (±22.20)). GLM7 represents the null hypothesis of no condition and trial type specific differences across phases.

Multilevel models testing specific unexpected (and expected) transitions

We used different multilevel GLMs to test the most specific prediction of the state space factorization account. If participants abstract and compositionally reuse prior knowledge, performance should be different across conditions in probe trials testing transitions that are not expected under the respective structural forms experienced during prior learning.

GLM9: Multilevel model featuring individually varying intercepts, main effects of transition type (4->1 state transition vs 6->1 state transition) and transition expectation (*EXP*, unexpected vs expected transitions) as well as covariation with the individually estimated intercept parameters and a main effect of condition.

$$Correct_{i,k,n,o} \sim Binomial\left(number\ of\ trials_{i,k,n,o}, p_{i,k,n,o}\right)$$

$$logit\left(p_{i,k,n,o}\right) = \mu_{\alpha} + \mu_{TT[i]} + \mu_{EXP[o]} + \mu_{COND[k]} +$$

$$\alpha_{subject[n],TT[i]} + \gamma_{subject[n],EXP[o]}$$

$$\begin{bmatrix}\alpha_{n,1} \\ \alpha_{n,2}\end{bmatrix} \sim MVNormal\left(\begin{bmatrix}0 \\ 0\end{bmatrix}, S_{TT}\right)$$

$$\begin{bmatrix}\gamma_{n,1} \\ \gamma_{n,2}\end{bmatrix} \sim MVNormal\left(\begin{bmatrix}0 \\ 0\end{bmatrix}, S_{EXP}\right)$$

$$S_{TT} = \begin{pmatrix}\sigma_{subject} & 0 \\ 0 & \sigma_{TT}\end{pmatrix} R_{TT} \begin{pmatrix}\sigma_{subject} & 0 \\ 0 & \sigma_{TT}\end{pmatrix} \qquad \textit{(Supp. Eq. 3)}$$

$$S_{EXP} = \begin{pmatrix}\sigma_{subject} & 0 \\ 0 & \sigma_{EXP}\end{pmatrix} R_{EXP} \begin{pmatrix}\sigma_{subject} & 0 \\ 0 & \sigma_{EXP}\end{pmatrix}$$

$$\mu_{\alpha} \sim Normal(0, 1)$$

$$\mu_{TT[i]} \sim Normal(0, 1)$$

$$\mu_{EXP[o]} \sim Normal(0, 1)$$

$$\mu_{COND[k]} \sim Normal(0, 1)$$

$$\sigma_{subject} \sim Half - Normal(0,1)$$
$$\sigma_{TT} \sim Half - Normal(0,1)$$
$$\sigma_{EXP} \sim Half - Normal(0,1)$$
$$R_{TT} \sim LKJcorr(2)$$
$$R_{EXP} \sim LKJcorr(2)$$

where *o* denotes the parameter estimate for the o-th level of expectation (unexpected vs expected transitions)

GLM10: Multilevel model featuring individually varying intercepts, main effects of transition type (4->1 state transition vs 6->1 state transition) and transition expectation (unexpected vs expected transitions) as well as covariation with the individually estimated intercept parameters, a main effect of condition, and an interaction effect of transition type x condition), but not including a transition type x expectation x condition interaction effect.

$$Correct_{i,k,n,o} \sim Binomial(number\ of\ trials_{i,k,n,o}, p_{i,k,n,o})$$
$$logit(p_{i,k,n,o}) = \mu_{\alpha} + \mu_{TT[i]} + \mu_{EXP[o]} + \mu_{COND[k]} +$$
$$\alpha_{subject[n],TT[i]} + \gamma_{subject[n],EXP[o]} + \mu_{TTx\ COND}$$
$$\begin{bmatrix}\alpha_{n,1}\\ \alpha_{n,2}\end{bmatrix} \sim MVNormal\left(\begin{bmatrix}0\\0\end{bmatrix}, S_{TT}\right)$$
$$\begin{bmatrix}\gamma_{n,1}\\ \gamma_{n,2}\end{bmatrix} \sim MVNormal\left(\begin{bmatrix}0\\0\end{bmatrix}, S_{EXP}\right)$$
$$S_{TT} = \begin{pmatrix}\sigma_{subject} & 0\\ 0 & \sigma_{TT}\end{pmatrix} R_{TT} \begin{pmatrix}\sigma_{subject} & 0\\ 0 & \sigma_{TT}\end{pmatrix}$$
$$S_{EXP} = \begin{pmatrix}\sigma_{subject} & 0\\ 0 & \sigma_{EXP}\end{pmatrix} R_{EXP} \begin{pmatrix}\sigma_{subject} & 0\\ 0 & \sigma_{EXP}\end{pmatrix} \qquad \textit{(Supp. Eq. 4)}$$
$$\mu_{\alpha} \sim Normal(0,1)$$
$$\mu_{TT[i]} \sim Normal(0,1)$$
$$\mu_{EXP[o]} \sim Normal(0,1)$$
$$\mu_{COND[k]} \sim Normal(0,1)$$
$$\mu_{TTxCOND} \sim Normal(0,1)$$
$$\sigma_{subject} \sim Half - Normal(0,1)$$
$$\sigma_{TT} \sim Half - Normal(0,1)$$
$$\sigma_{EXP} \sim Half - Normal(0,1)$$

$$R_{TT} \sim LKJcorr(2)$$

$$R_{EXP} \sim LKJcorr(2)$$

GLM11: Multilevel model featuring individually varying intercepts, main effects of transition type (4->1 state transition vs 6->1 state transition) and transition expectation (unexpected vs expected transitions) as well as covariation with the individually estimated intercept parameters, a main effect of condition, and an parameter estimating all individual levels of transition type and expectation, specific for both conditions (*COND/TT/EXP*, akin to an interaction effect of transition type x expectation x condition)

$$Correct_{i,k,n,o} \sim Binomial\left(number\ of\ trials_{i,k,n,o}, p_{i,k,n,o}\right)$$

$$logit\left(p_{i,k,n,o}\right) = \mu_{\alpha} + \mu_{TT[i]} + \mu_{EXP[o]} + \mu_{COND[k]} +$$

$$\alpha_{subject[n],TT[i]} + \gamma_{subject[n],EXP[o]} + \beta_{COND[k],TT[i]/EXP[o]}$$

$$\begin{bmatrix} \alpha_{n,1} \\ \alpha_{n,2} \end{bmatrix} \sim MVNormal\left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, S_{TT}\right)$$

$$\begin{bmatrix} \gamma_{n,1} \\ \gamma_{n,2} \end{bmatrix} \sim MVNormal\left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, S_{EXP}\right)$$

$$\begin{bmatrix} \beta_{k,1,1} \\ \beta_{k,1,2} \\ \beta_{k,2,1} \\ \beta_{k,2,2} \end{bmatrix} \sim MVNormal\left(\begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}, S_{COND/TT/EXP}\right)$$

$$S_{TT} = \begin{pmatrix} \sigma_{subject} & 0 \\ 0 & \sigma_{TT} \end{pmatrix} R_{TT} \begin{pmatrix} \sigma_{subject} & 0 \\ 0 & \sigma_{TT} \end{pmatrix}$$

$$S_{EXP} = \begin{pmatrix} \sigma_{subject} & 0 \\ 0 & \sigma_{EXP} \end{pmatrix} R_{EXP} \begin{pmatrix} \sigma_{subject} & 0 \\ 0 & \sigma_{EXP} \end{pmatrix} \qquad \text{(Supp. Eq. 5)}$$

$$S_{COND/TT/EXP} = \begin{pmatrix} \sigma_{COND} & 0 \\ 0 & \sigma_{TT/EXP]} \end{pmatrix} R_{TT/EXP} \begin{pmatrix} \sigma_{COND} & 0 \\ 0 & \sigma_{TT/EXP} \end{pmatrix}$$

$$\mu_{\alpha} \sim Normal(0, 1)$$

$$\mu_{TT[i]} \sim Normal(0, 1)$$

$$\mu_{EXP[o]} \sim Normal(0, 1)$$

$$\mu_{COND[k]} \sim Normal(0, 1)$$

$$\sigma_{subject[n]} \sim Half - Normal(0, 1)$$

$$\sigma_{TT} \sim Half - Normal(0, 1)$$

$$\sigma_{EXP} \sim Half - Normal(0, 1)$$

$$\sigma_{COND} \sim Half - Normal(0, .15)$$

$$\sigma_{TT/EXP} \sim Half - Normal(0, .15)$$

$$R_{TT} \sim LKJcorr(2)$$

$$R_{EXP} \sim LKJcorr(2)$$

$$R_{COND/TT/EXP} \sim LKJcorr(2)$$

The data in the analysis of unexpected and expected transitions across all trials was best explained by GLM11: PSIS values ($\pm$ standard error): GLM11 = 4695.7 ($\pm$66.26), GLM10 = 4702.0 ($\pm$67.97), GLM9 = 4715.1 ($\pm$67.59); WAIC values ($\pm$ standard error): GLM11 = 4585.0 ($\pm$62.28), GLM10 = 4586.7 ($\pm$63.08), GLM9 = 4596.4 ($\pm$62.85).

To test for transition type specific differences on unexpected transitions across conditions during the initial occurrence of probing each transition type during the transfer learning phase, we used two different multilevel Bayesian GLMs.

GLM12: GLM only featuring a single predictor with four different levels, capturing 4->1 or 6->1 transitions in both conditions separately (equivalent to an interaction effect of transition type x condition).

$$Correct_{i,k} \sim Binomial(number\ of\ trials_{i,k}, p_{i,k})$$

$$logit(p_{i,k}) = \mu_{COND[k]/TT[i]}$$

$$\begin{bmatrix} \mu_{k,1} \\ \mu_{k,2} \end{bmatrix} \sim MVNormal\left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, S\right)$$

$$S = \begin{pmatrix} \sigma_{COND} & 0 \\ 0 & \sigma_{TT} \end{pmatrix} R_{TT/COND} \begin{pmatrix} \sigma_{COND} & 0 \\ 0 & \sigma_{TT} \end{pmatrix} \quad \textit{(Supp. Eq. 6)}$$

$$\sigma_{COND} \sim Half - Normal(0, .5)$$

$$\sigma_{TT} \sim Half - Normal(0, .5)$$

$$R_{TT/COND} \sim LKJcorr(2)$$

GLM13: Multilevel model featuring individually varying intercepts, main effects of transition type (4->1 state transition vs 6->1 state transition), as well as covariation with the individually estimated intercept parameters, a main effect of condition, and an interaction effect of transition type x condition.

$$Correct_{i,k,n} \sim Binomial(number\ of\ trials_{i,k,n}, p_{i,k,n})$$

$$logit(p_{i,k,n}) = \mu_{\alpha} + \mu_{TT[i]} + \mu_{COND[k]} + \alpha_{subject[n],TT[i]} + \mu_{TTx\,COND}$$

$$\begin{bmatrix} \alpha_{n1} \\ \alpha_{n,2} \end{bmatrix} \sim MVNormal\left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, S_{TT}\right)$$

$$S_{TT} = \begin{pmatrix} \sigma_{subject} & 0 \\ 0 & \sigma_{TT} \end{pmatrix} R_{TT} \begin{pmatrix} \sigma_{subject} & 0 \\ 0 & \sigma_{TT} \end{pmatrix}$$

$$\begin{gathered} \mu_{\alpha} \sim Normal(0,1) \\ \mu_{TT[i]} \sim Normal(0,1) \\ \mu_{COND[k]} \sim Normal(0,1) \\ \mu_{TTx\,COND} \sim Normal(0,1) \\ \sigma_{subject} \sim Half-Normal(0,1) \\ \sigma_{TT} \sim Half-Normal(0,1) \\ R_{TT} \sim LKJcorr(2) \end{gathered}$$

(Supp. Eq. 7)

In the analysis focusing on the first occurrence of the unexpected transitions, we found that GLM12 explained the data better than a more complex model also considering individual varying intercepts, trial type and condition main effects (GLM13) (PSIS values (± standard error): GLM12 = 800.9 (±10.32), GLM13 = 803.6 (±12.04); WAIC values (± standard error): GLM12 = 800.9 (±10.31), GLM13 = 802.9 (±12.02)).