Multi-task Learning in the Human Brain

Tobias Ludwig (tobias.ludwig@uni-tuebingen.de)
Dept. Neural Dynamics and MEG, Hertie Institute for Clinical Brain Research
Centre of Integrative Neuroscience
University of Tübingen
Tübingen, Germany

Markus Siegel (markus.siegel@uni-tuebingen.de) †
Dept. Neural Dynamics and MEG, Hertie Institute for Clinical Brain Research
Centre of Integrative Neuroscience
Tübingen, Germany

Eric Schulz (eric.schulz@helmholtz-munich.de) †
Human-centered AI, Helmholtz Munich
München, Germany

† shared senior authorship

This research is supported by the European Research Council (ERC CoG 864491).

Abstract

As humans we are naturals at learning multiple related tasks together. Which mechanisms do we use to transfer knowledge between tasks and generalize to new ones? Here, we conducted an MEG study in which N=42 participants virtually fed animals with fruits. The animals had different preferences, serving as tasks, while the position of the fruits could be transferred. Behaviorally, we found that people are using a hybrid model with transferable features and task-specific values. In MEG data, we found enhanced activation of the parietal cortex for successful transfer learning compared to within-task learning.

Keywords: multi-task learning; knowledge transfer; magnetoencephalography (MEG)

Introduction

Multi-task Learning is a part of our everyday behavior. We act in relatively stable environments with multiple goals or tasks, e.g. we go shopping in the same supermarket with different shopping lists. However, while Multi-task Reinforcement Learning is an important topic in robotics (Barreto et al., 2019; Andrychowicz et al., 2017), there is little research in human behaviour (Tomov, Schulz, & Gershman, 2021; Flesch, Saxe, & Summerfield, 2023), and nearly none of it uses brain imaging (Hall-McMaster, 2025).

Formally, multi-task learning refers to jointly learning a set of M related tasks by using a shared representation (Caruana, 1997). Here, we run an MEG study to see if people use these shared representations, and how performance depends on M. We use behavioral modelling to categorize trials as model-based and model-free (with and without knowledge transfer) and we find differences in the neural activation of the two kinds of trials.

Experiment

Before the experiment, participants were instructed on how to feed 6 animals with 8 fruits. Specifically, animals had certain needs corresponding to their size: small animals needed small fruits, big animals needed big fruits, and so on. On top of this need (boxes in Fig. 1A), which was rewarded by 1 point if satisfied, animals had a preference (thick lines in Fig. 1A), which resulted in two points, if satisfied. This animal-fruit mapping stayed constant throughout the experiment and participants were trained to know it well.

The experiment itself required them to collect fruits for the current animal from a market (Fig. 1B). At first, they did not know the arrangement and had to explore different paths, as they could not see the whole tree. An optimal strategy would be to memorize the positions of all the fruits, and to use this mental map for planning the best path even for new animals. Fruits stayed at the same positions within blocks of 9 trials (Fig. 1C), but after each block they were rearranged, such that participants had to relearn. Each block was subdivided into 6 training trials, within which $M = 1, 2,$ or 3 different animals could occur, and 3 test trials with 3 new animals. We instructed the number of training tasks M at the start of each block. Participants completed 30 blocks in the MEG scanner (Fig. 1D). They saw the animal, defining the task, decided for an action by pressing the left or right button, reached a fruit on the first level, took a second action, and arrived at the second fruit. At the end of a trial, the cumulative reward was shown in the form of smileys (rewards per fruit are not shown). The reward along a path ranged from 0 to 3, however, crucially, the maximum reward in any given block was not known a priori, as it depended on the specific arrangement of the fruits (i.e. 3 points were not always achievable for all animals). Moreover, there were 8 fruits in total, but only 6 states, ensuring that the fruits in unvisited states could not be inferred by exclusion.

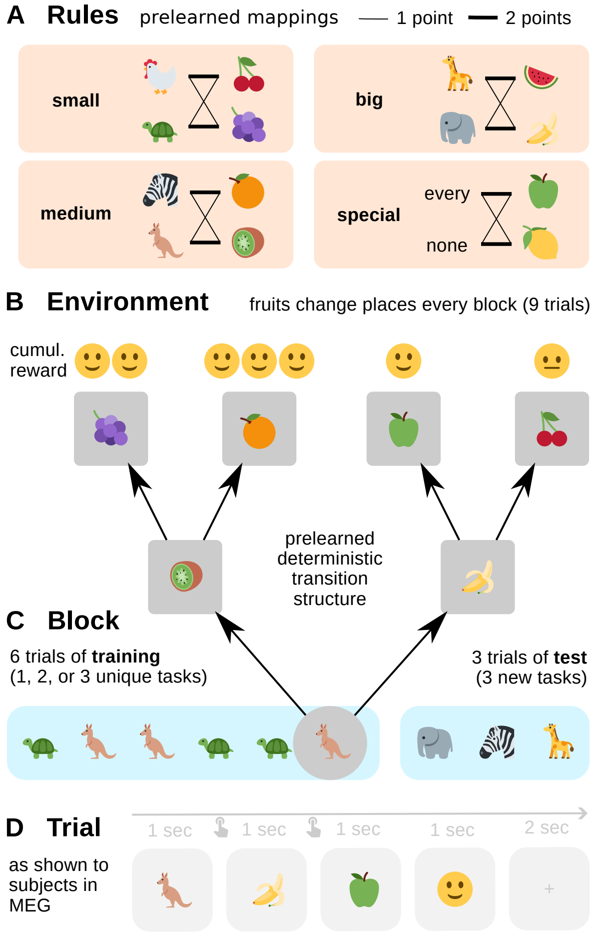

Figure 1: A) Participants are pretrained to know how rewarding the fruits are to the animals. B) They accumulate reward by collecting fruits along paths in a binary tree. C) In every block, there are 6 training trials with $M = 1, 2, 3$ different animals, and 3 test trials with new animals. After each block, the fruit arrangement changes. D) Participants only see the current animal, fruits, and cumulative reward, not the whole tree.

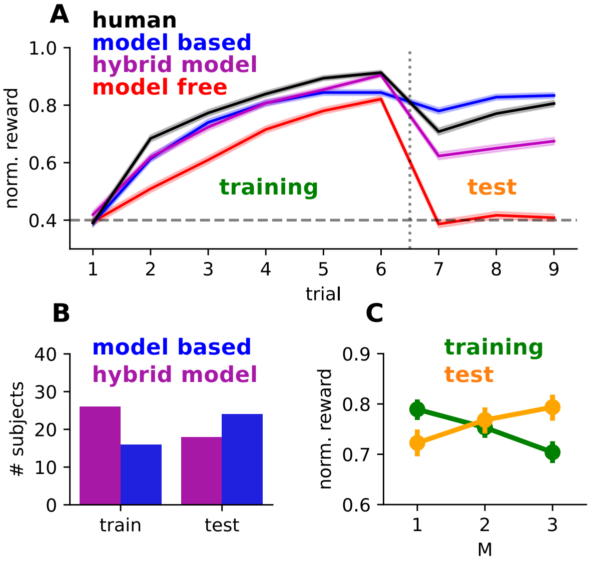

Figure 2: A) Performance of humans and models over trials. B) Model frequency in training and test. C) Human performance depending on the number of training tasks M.

Behavioural results

N=42 participants (20 female, mean age: 27.5) took part in the MEG study. The session took 90min and they were paid a bonus relative to their overall performance. Fig. 2A shows their performance over trials (black). Reward is normalised to be 1 for the maximal possible reward given the current environment and animal. It rises over the 6 training trials, but drops with the start of the test trials, which had new animals.

We modelled this behaviour with three agents. Firstly, the model-based agent (MB; in blue) acquires a mental map of the fruit positions and uses planning to predict the best path for a given animal. Thereby, it can generalise even to new animals, and shows the least performance drop in test. On the other extreme, the model-free agent (MF; in red) does not represent the fruits at all, but only stores the rewards per fruit separately for each animal. For each new task, it has to guess, leading to random performance in test. Finally, the hybrid model (HM; in magenta) combines the values from MB and MF agent with a weighting factor: $V_{HM} = w \cdot V_{MB} + (1-w) \cdot V_{MF}$. Here we fixed $w = 0.5$. HM shows a moderate performance drop in test. All models had a temperature parameter τ for random exploration as well as an uncertainty bonus β for directed exploration towards unvisited states U. We tuned τ and β per subject and computed model likelihoods $p \propto \exp(-(V+\beta U)/\tau)$. In the simulation, models were conditioned on human observations.

Fig. 2B shows the model frequency split by training and test. In training, the hybrid model was the more frequent model, fitting 26 participants best, while in test, MB won with 24 subjects. Interestingly, no participant behaved in a purely model-free way, neither in training nor in test. Furthermore, we see an effect of manipulating the number of training tasks M on performance (Fig. 2C). In training, there is a significant decrease of performance with increasing M ($\beta = -0.0427, p < 0.001$), due to increased switch costs. During test, on the other hand, participants benefited from having seen more training tasks ($\beta = 0.0355, p < 0.001$), as they had a more robust representation of the market. This may also be mediated by more exploration, measured as the number of visited paths during training depending on M ($\beta = 0.3774, p < 0.001$).

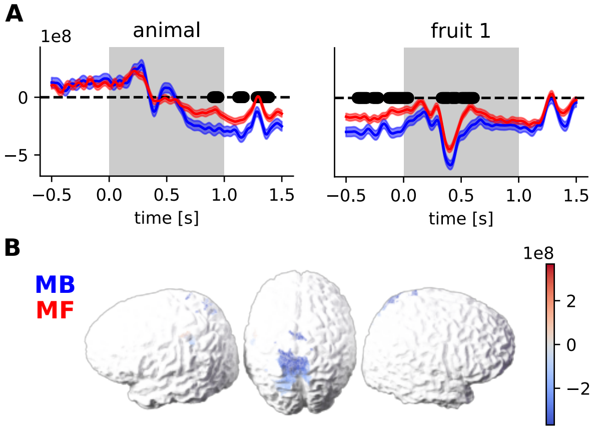

Figure 3: A) Source-localised evoked potentials in MB (blue) and MF (red) trials. Grey windows mark stimulus presentation. Black dots mark clusters of significant difference ($p < 0.05$). B) Significant difference on cortex (cluster permutation test).

MEG results

To understand the neural mechanism of knowledge transfer across tasks, we looked at trials with choices that were either more likely under the MB or under the MF model. By design, MB transfers values perfectly, whereas MF can only learn within task. We found significant differences of neural activity between MF and MB trials when planning the path for the current task (Fig. 3A), i.e. after seeing the animal, leading up to the button press (at fruit onset), and during the first fruit. The cortical distribution (Fig. 3B) revealed that model-based reasoning in our task was accompanied by enhanced activity in the central parietal cortex. The effect could not be driven by reaction times, as those did not differ between the two types of trials (two-sided t-test, $t = 0.0578, p = 0.954$).

Discussion

In this MEG study, we set out to study how people transfer knowledge while learning related tasks jointly. We find that they use a hybrid model of transferable representations (in this case the fruit arrangement) as well as task specific rewards. While multi-task learning can be impeded by task switching, generalisation to new tasks benefits from more robust representations. Lastly, we find enhanced activation in parietal cortex for trials that were more in line with model based choices, specifically during planning. Future work should have a closer look at alternative transfer mechanisms, for instance the use of task similarity for generalisation or counterfactual reasoning for updating other tasks (Andrychowicz et al., 2017).

References

Andrychowicz, M., Wolski, F., Ray, A., Schneider, J., Fong, R., Welinder, P., . . . Zaremba, W. (2017). Hindsight Experience Replay. In *Advances in Neural Information Processing Systems* (Vol. 30). Curran Associates, Inc.

Barreto, A., Borsa, D., Quan, J., Schaul, T., Silver, D., Hessel, M., . . . Munos, R. (2019, January). Transfer in Deep Reinforcement Learning Using Successor Features and Generalised Policy Improvement. *arXiv:1901.10964 [cs]*. (arXiv: 1901.10964)

Caruana, R. (1997). Multitask Learning. *Machine Learning*, *28*.

Flesch, T., Saxe, A., & Summerfield, C. (2023, March). Continual task learning in natural and artificial agents. *Trends in Neurosciences*, *46*(3), 199–210. (Publisher: Elsevier) doi: 10.1016/j.tins.2022.12.006

Hall-McMaster, T. M. G. S. . S. N., Sam. (2025, June). Neural evidence that humans reuse strategies to solve new tasks. *PLOS Biology*. doi: 10.1371/journal.pbio.3003174

Tomov, M. S., Schulz, E., & Gershman, S. J. (2021, June). Multi-task reinforcement learning in humans. *Nature Human Behaviour*, *5*(6), 764–773. doi: 10.1038/s41562-020-01035-y