Neural signatures of coping with multiple tasks in mouse visual cortex

Márton A. Hajnal[1], **Duy T. Tran**[2], **Michael C. Einstein**[2], **Gergő Orbán**[1]*, **Pierre-Olivier Polack**[3], **Peyman Golshani**[2]**

[1] Computational Systems Neuroscience Lab, MTA Wigner Research Centre for Physics, 1121 Budapest, Hungary

[2] Department of Neurology and Psychiatry, David Geffen School of Medicine, UCLA, Los Angeles, California 90095

[3] Center for Molecular and Behavioral Neuroscience, Rutgers, State University of New Jersey, Newark, New Jersey 07102

corresponding authors: *orban.gergo@wigner.mta.hu, **pgolshani@mednet.ucla.edu

Abstract

Flexibly adapting to different task requirements is a key challenge of the visual system. In particular, depending on a potentially unobserved context the same stimuli might require different behavior. While task-related activity has been identified as early as the primary visual cortical (V1) activity in mice, it remained unknown if and how the visual cortex primarily dealing with feed-forward input contributes to efficient arbitration between tasks. Mice were trained to perform a multimodal task-switching paradigm where animals were required to make decisions either based on the identity of visual or that of auditory stimuli. Neural activity was recorded from all layers of V1 on 128 channels with extracellular electrodes. Our analysis has identified task-related variables in population responses. Importantly, while task-related variables could be identified mostly during stimulus presentation, the variable that could identify the specific task being performed could be reliably decoded from intertrial intervals, indicating a representation which is aware of the across trial contingency of task context. These results provide insights into how continual learning, the major challenge concerning the acquisition of multiple tasks relying on the same neural circuitry, can be achieved in biological agents.

Keywords: mouse; vision; multiple task; multiple stimulus; context

Adaptation is key to survival. Adaptation encompasses learning both the stimulus statistics and task contingencies. Critically, in a rich environment animals need to adapt to multiple tasks. Continual learning is a central challenge in machine learning and it concerns the question how multiple tasks can be learned using the same circuitry. Elastic weight consolidation (Kirkpatrick et al., 2017) and learning latent task variables (Flesch, Balaguer, Dekker, Nili, & Summerfield, 2018) are two contrasting methods to avoid catastrophic forgetting, the effect when more novel tasks tend to overwrite learned representations. The challenge of multiple task learning is particularly prominent when animals need to make different choices based on the same stimulus depending on task context. While being challenging, it provides an opportunity to study task effects in more detail.

Successful execution of a task implies that task-relevant quantities are represented in some form in the neural circuitry. In the primary visual cortex (V1) of rodents, a large number of studies have demonstrated task-induced variance in neuronal responses (Poort et al., 2015; Allen et al., 2017; Caras & Sanes, 2017). Besides the feed forward input carrying sensory information through the thalamus, top-down inputs are associated with such task-related activity (Caras & Sanes, 2017). Earlier studies have discovered an intricate set of variables, including locomotion (Polack, Friedman, & Golshani, 2013), attentional signals that change between task-engaged and passive conditions (Allen et al., 2017), and a signal that has been interpreted as a potential reward expectation signal, which emerges prior to the presentation of stimuli directly relevant to task-execution (Jaramillo & Zador, 2011; Poort et al., 2015). It remains unclear however, how these signals can contribute to a flexible arbitration between task demands.

Investigation of how a V1 neuron population deals with the execution of two tasks provides an opportunity to understand how the neural system can deal with continual learning and to get an insight how top-down influences contribute to task execution.

Results

Isolated units were extracellularly recorded with 2x64 electrode shanks from all layers of V1 of 14 mice performing a multimodal (visual and auditory) perceptual discrimination task. Animals were rewarded for licking upon perceiving a go signal and avoided delay of the next trial for withholding licking at the no-go signal. The modality of the go/no-go pair was not constant in a recording session but the animals were required to switch the relevant modality. Cuing epochs, featuring trials where stimuli with a single modality were present, were followed by dual-modality epochs in which both modalities were present but animals were required to make decisions based on the modality presented in the preceding epoch while ignoring the other (Fig. 1). This setting permits the study of task-related signals since the same pair of stimuli could appear in different task contexts.

Representation of behavioral choice. First, following earlier indications (Reynolds & Heeger, 2009) that stimuli relevant to a particular task are represented with higher precision, we trained a linear decoder for decoding the orientation of the visual stimulus on two subsets of trials: first, when the animal was required to make its decision based on this stimulus modality; second when it was required to ignore it. Decoding visual information was more efficient at most, albeit not all animals when attended to, compared to when ignored (Fig. 2). Similarly, we repeated the analysis with a focus on the representation of the auditory stimulus. Interestingly, the

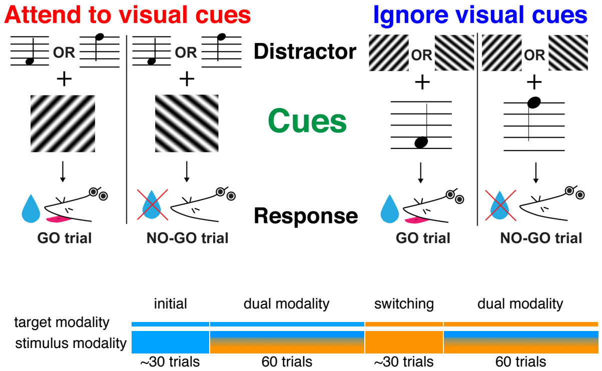

Figure 1: Task switching between contexts. Top: Stimuli and role of stimulus modalities in dual-modality tasks. Bottom: Block layout of an experimental session.

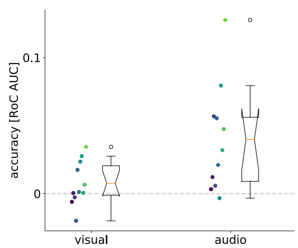

Figure 2: Stimuli decodability difference between attended to and ignored blocks, for visual and audio respectively; integrals of accuracy timecourse 95% significant differences during stimulus; colors: individual mice, boxplots: mean, 2 s.e.m. for 11 mice

analysis for the auditory modality yielded similar results, with a more pronounced improvement when attended. These results confirm that attended task-relevant stimulus can be more efficiently read out from the neuron population.

Next, our goal was to elaborate on population phenomena leading to a more efficient readout of information when a particular modality is relevant. We investigated population responses with a linear unsupervised method, tensor component analysis (Williams et al., 2018, TCA), which indicated that variability is induced in population responses by a number of easily interpretable features related to task execution (data not shown). To explicitly test the contribution of different task-related variables, we constructed decoders for these variables, including visual stimulus identity, auditory stimulus identity, the choice the animal made (Fig. 3A,B,D). While most reliable decoding can be achieved for the visual stimulus identity, both auditory stimulus identity and choice identity could be obtained with different time courses. Importantly, since animals were required to be able to perform multiple tasks, the task identity (referred later as context) is also a relevant factor, and we constructed a specific decoder for this variable (Fig. 3C). Importantly, high decoding performance of task identity indicates that context contributes to population activity by driving it to separate subspaces when context changes.

Using the insight obtained from the decoding analysis, we wanted to investigate the source of performance improvement

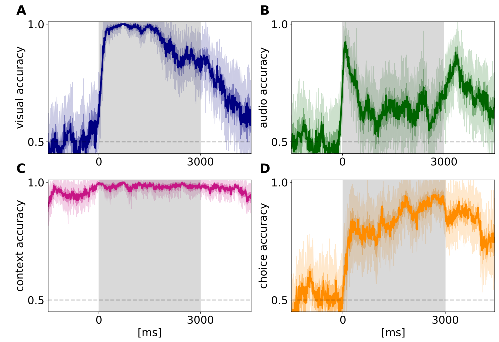

Figure 3: Time course of linear decoding of task variables in multimodal trials for a well-trained mouse. Decoder performance was measured at different time windows along the trial (10 ms resolution sliding 50 ms wide windows of instantaneous firing rate). Grey: stimulus present. Panels correspond to visual and audio stimulus identity, context and choice for panels A-D, respectively.

when the decoded variable is relevant for decision. In particular, we wanted to test if the more reliable representation of stimulus identity is a result of increased separability of stimuli along the axis where changes in stimulus identity introduce variance, or changes in other task-related variables. We used the decision boundaries determined by the linear decoders of (Fig. 3) to establish the directions in the population activity space along which population responses can be classified. We used the normal vectors of the decision boundary hyperplanes as basis vectors spanning particular subspaces, and projected instantaneous firing rate vectors of the population onto these subspaces. In particular, to investigate performance improvement in visual decoding, we used the decision normal vectors of visual and choice decoders as a basis to construct a two-dimensional subspace, on which the projection of individual trials are represented as individual points (Fig. 4). The bases obtained from the multimodal blocks are not orthogonal, meaning that these variables are correlated, but are consistently linearly highly independent (*blue and yellow lines*). Trial-by trial analysis of the planar subspace under the ignore visual condition (*purple and red symbols*) reveals that in this task condition the decision normal alone would be closer to orthogonal to the choice condition. However, upon performing the attend to visual task a shift is introduced in the population response vectors along the choice direction (*blue and turquoise symbols*). This systematic shift introduces larger separation which accounts for decoding performance improvement when the visual stimulus is relevant for decisions. Importantly, the separation is more pronounced if we take into account the performance of the animal: in incorrect trials separation is less pronounced thus decreasing the separation of the responses according to stimulus identity. In summary, these results indicate that the attention-related im-

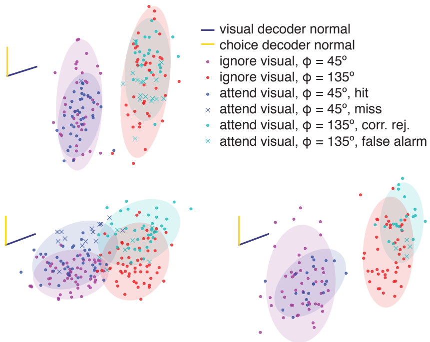

Figure 4: Trial-by-trial instantaneous firing rate responses of the neuron population in the subspace defined by the normal vectors of the choice and visual stimulus identity decoders. Dots represent the projection of rates calculated in the first second of the trials on the normal vectors of the decoders in multimodal trials. Three panels correspond to data from three mice. Note that decision boundaries are not orthogonal. Shaded areas correspond to mean + 2SD in a given condition. Attending to the visual stimulus results in consistent shifts along the choice dimension. Note that incorrect trials (crosses) tend to be closer to the mean of the unattended trials along the choice axis.

provement can be traced back to the introduction of a choice variable that operates in a subspace linearly independent form the subspace where the stimulus identity introduces variability.

Representation of context. Earlier accounts of learning identified a task-specific signal in a period preceding the appearance of the target stimulus, which has been speculated to be linked to reward expectation (Jaramillo & Zador, 2011; Poort et al., 2015). Our task design permits a more detailed investigation of prestimulus task-related activity, more specifically, we address the question if prestimulus task-related activity conveys information solely about the expected reward or about the actual task context as well. Linear decoding analysis of the activity during stimulus presentation (ON-stimulus activity) reveals that stimulus and task variables can be identified from ON-stimulus activity (Fig. 5A, left). Critically, our analysis on the identification of task context based on the V1 population activity shows a relatively stable representation of spontaneous activity not only during stimulus presentation, but task identity can be identified with equally high reliability beyond stimulus presentation (Fig. 3C). Indeed, while other variables cannot be identified in the population activity preceding stimulus presentation, the contextual variable identifying the task can (Fig. 5A, center). Decodability of task context both from the prestimulus and ON-stimulus activities does not ensure that the prestimulus and ON-stimulus activities would be decoded along the same activity subspace. We therefore designed an analysis which tests if the prestimulus activity distribution represents the context variable identically to that in the ON-stimulus activity distribution. We used a decoder trained on a short period during the prestimulus activities of the multimodal tasks (the same period as the one shown on Fig. 5B) but tested on activity distributions obtained from a different time window in the prestimulus period and on a time window during ON-stimulus activity. This across-time test revealed equally efficient task context decoding both during the prestimulus period and during ON-stimulus period (Fig. 5A, right). Thus, in the period preceding the stimulus not simply reward expectation but also a signal specific to the task is present. Furthermore, this task-specific signal is invariant across on-stimulus and off-stimulus periods.

Task learning is a delicate problem and studies of classical conditioning indicate that animals are able to build parsimonious models of environmental variables during learning (Courville, Daw, & Touretzky, 2006). It is unclear if the task variable we identified in the prestimulus activity reflects such a parsimonious representation of the task. In particular a parsimonious representation would indicate that the task is identical if the set of stimuli used for making decisions are not changing irrespective if other distractor stimuli are present (in our case these are the non-attended stimuli). We tested if the context variable that we identified in the V1 population activity in multimodal trials is invariant to the presence or absence of the unattended stimuli. We used the decoder trained on the prestimulus activity of multimodal trials to assess context in single-modality trials (Fig. 5B). Decoding context from the first block of the experiment revealed that the behaviorally relevant modality reliably predicted population activity along the dimension defined by the context variable both in the prestimulus (Fig. 5B) and in the ON-stimulus periods (Fig. 5B). Interestingly, this block being at the very beginning of the session the animal had no exposure to multimodal stimuli. Still, the context variable reliably identified the whole task and data indicates that animals recognize the task almost immediately after the start of the block. A similar analysis of the single-modality block separating the two multimodal blocks reveals that the population activity gradually drifts from one context to the other context (Fig. 5B), with considerable variability in speed of adaptation across animals. These results indicate that task context variable reflects a parsimonious representation of the task.

The robust presence of the task context variable in both prestimulus and on-stimulus activities indicates that the activity of the neuron population reflects the characteristics of the specific task. In particular, even if bottom-up signals have overlapping representations in different tasks, the context variable distinguishes the two representations. We wanted to test if the representation of the context variable is related to other measures that reflect task acquisition. We used the task-specificity of the prestimulus activity as a measure for the efficiency of task representation. The prestimulus task-specificity of the activity across the whole population of recorded animals is well reflected in the task specificity in ON-stimulus periods (Fig. 5C). Interestingly, the decodability of choice from ON-stimulus activity was also highly predictable based on the

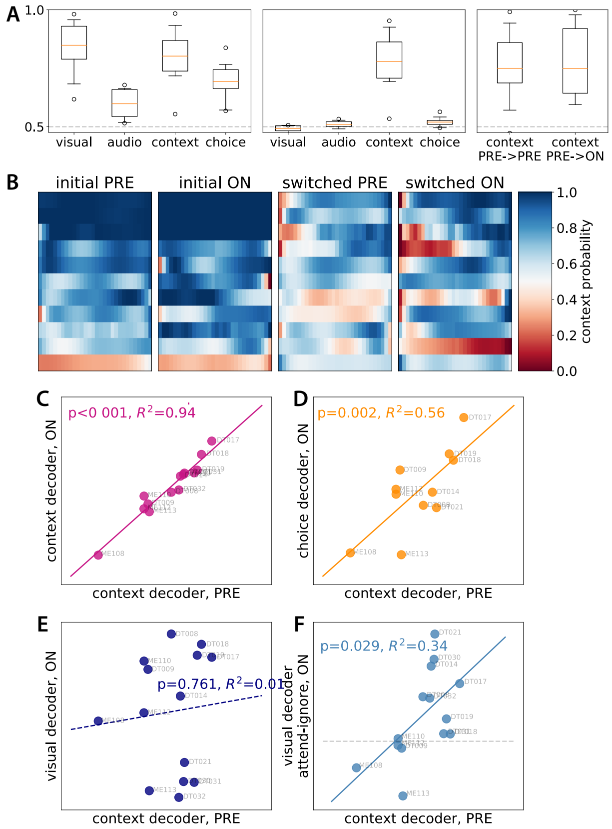

Figure 5: **A**, Discrimination between identities of various task relevant variables in the multimodal blocks during stimulus presentation (left), and prior to stimulus presentation (center). Right, Across-time prediction of context: a context decoder trained in a short time window prior to stimulus presentation is used to predict the context variable in a different time window prior to the stimulus and during stimulus presentation. **B** Decoding task context in individual single modality trials (horizontal axis) based on decoders trained on the prestimulus periods of multimodal trials. Colors: smoothed single trial probabilities from incorrect context (red) through chance (white) to correct context (blue). Decoding is performed on both initial and switching block trials and both from prestimulus and on-stimulus activities. **C-F** Predictability of decoder accuracy from the performance of prestimulus decoder for context during stimulus presentation (**C**), choice (**D**) and visual stimulus identity (**E**). **F**, Visual decoder accuracy difference between attended and ignored context, dashed line denotes no difference.

reliability of the context representation (Fig. 5D). To test if the representation of context corresponds to a more efficient representation of the stimulus, we compared context decodability with stimulus decodability. We could not find significant dependence between the two (Fig. 5E). Finally, we contrasted context decodability to the level of improvement we measured for decoding the attended stimuli over that of unattended stimuli. We found that the advantage of the attended stimulus decoding was highly predictable with the quality of context representation (Fig. 5F).

Conclusions

Using a multimodal decision making task we have demonstrated that signatures of a sophisticated representation can be found which can be an underpinning of arbitrating between multiple tasks using the same sensory cortical circuitry. Key to our study is that identical stimuli are used in multiple contexts, which made it possible to identify a task-specific variable for context and a choice variable that was invariant across tasks. Our results provide insights into how continual learning can be achieved without catastrophic forgetting.

Acknowledgments.

This work has been supported by the Hungarian Brain Research Program (2017-1.2.1-NKP-2017-00002), and a Human Frontier Science Program grant (GO).

References

Allen, W. E., Kauvar, I. V., Chen, M. Z., Richman, E. B., Yang, S. J., Chan, K., . . . Deisseroth, K. (2017). Global Representations of Goal-Directed Behavior in Distinct Cell Types of Mouse Neocortex. *Neuron*, *94*, 891-907.

Caras, M. L., & Sanes, D. H. (2017). Top-down modulation of sensory cortex gates perceptual learning. *Proc Natl Acad Sci*, *37*, 9972-9977.

Courville, A. C., Daw, N. D., & Touretzky, D. S. (2006). Bayesian theories of conditioning in a changing world. *Trends Cogn Sci*, *10*, 294-300.

Flesch, T., Balaguer, J., Dekker, R., Nili, H., & Summerfield, C. (2018). Comparing continual task learning in minds and machines. *Proc Natl Acad Sci*, *115*, E10313-E10322.

Jaramillo, S., & Zador, A. M. (2011). The auditory cortex mediates the perceptual effects of acoustic temporal expectation. *Nat Neurosci*, *14*, 246-251.

Kirkpatrick, J., Pascanu, R., Rabinowitz, N., Veness, J., Desjardins, G., Rusu, A. A., . . . Hadsell, R. (2017). Overcoming catastrophic forgetting in neural networks. *Proc Natl Acad Sci*, *114*, 3521-3526.

Polack, P.-O., Friedman, J., & Golshani, P. (2013). Cellular mechanisms of brain state–dependent gain modulation in visual cortex. *Nat Neurosci*, *16*, 1331-1339.

Poort, J., Khan, A. G., Pachitariu, M., Nemri, A., Orsolic, I., Krupic, J., . . . Hofer, S. B. (2015). Learning Enhances Sensory and Multiple Non-sensory Representations in Primary Visual Cortex. *Neuron*, *86*, 1478-1490.

Reynolds, J. H., & Heeger, D. J. (2009). The normalization model of attention. *Neuron*, *61*, 168-185.

Williams, A. H., Kim, T. H., Wang, F., Vyas, S., Ryu, S. I., Shenoy, K. V., . . . Ganguli, S. (2018). Unsupervised discovery of demixed, low-dimensional neural dynamics across multiple timescales through tensor component analysis. *Neuron*, *98*, 1099–1115.e8.