Navigation of an abstract discrete environment by rhesus macaques

James L Butler, Shirley Mark, Tim Behrens* & Steve Kennerley*

* Equal contribution

Since the discovery of place cells and the 'cognitive map' by O'Keefe & Dostrovsky in 1971 much research has revealed the mechanisms underlying navigation of spatial environments. However, how mammals navigate more abstract conceptual spaces, such as those that underlie complex cognitive processes, remains unclear. We therefore taught two rhesus macaques (*Macaca mulatta*) several abstract discrete environments composed of networks of associated stimuli (e.g. a 5x5 grid with 4 edges per node). After just seven days of training subjects could successfully navigate to rewarded target locations several stimuli away with high precision. When learning a new environment, due to the high number of unique pathways (> 500), some routes would be encountered for the first time near the beginning of training whereas others would be encountered later in training. Subjects performed better in their first encounter of a unique pathway if it occurred later in training rather than earlier, therefore demonstrating a global representation of the map.

When presented with two options equidistant to the target, it is advantageous to choose the one with more possible subsequent routes to the goal. This is however computationally-expensive to calculate in an environment with such a large state space. A successor representation-like strategy which assigns a value based on the potential reward of all possible future states would perform well at this task, whereas a tree search-like strategy that individually searches all possible routes would take exponentially longer to solve this task. We found both subjects preferentially chose options with multiple possible routes to the target over those with fewer possible routes, and this was consistent across multiple environments tested. We have therefore shown that non-human primates can navigate complex abstract environments, can learn knowledge about the structure of these environments, and can navigate in an optimal manner which indicates a computationally-efficient process.

Additional Details

In a novel task designed to probe navigation of abstract conceptual environments, we constructed several 2D networks of associated stimuli with up to 4 neighbours per node (Figure 1A). Given different starting and target locations, these environments had hundreds of unique paths of different lengths, which would take non-human primates (NHPs) many months to learn in traditional laboratory settings. We therefore developed a novel home cage training system which allowed NHPs to train for up to 12 hours a day thereby maximising the training process.

Subjects were taught these environments using the following behavioural task. A target location was presented and the subjects were placed at a random starting position a certain distance away (Figure 1B). The four stimuli associated with their starting stimulus were given as options and the subjects could choose which stimulus to move to. The position of the choices was randomised on each trial such that no spatial information was present in the task. Upon making a choice they would then be shown the neighbours of their new position and the process repeated until they reached their target destination whereupon they were given a reward. The subjects were never shown the global structure of the conceptual environment, and therefore had to learn to navigate the environment by sampling the different positions and their associated neighbours.

Subjects learned to navigate the environment over a period of 7 training sessions, during which time the average number of steps to reach a target when starting 3 positions away dramatically decreased (Figure 1C). To assess global versus local learning of the map we assessed subject's performance on the first exposure of each unique start-end combination as a function of when each combination was first offered during training. Subjects took a more direct route to the target location on combinations that they first experienced later in training compared to combinations first experienced earlier in training (Figure 1D). The

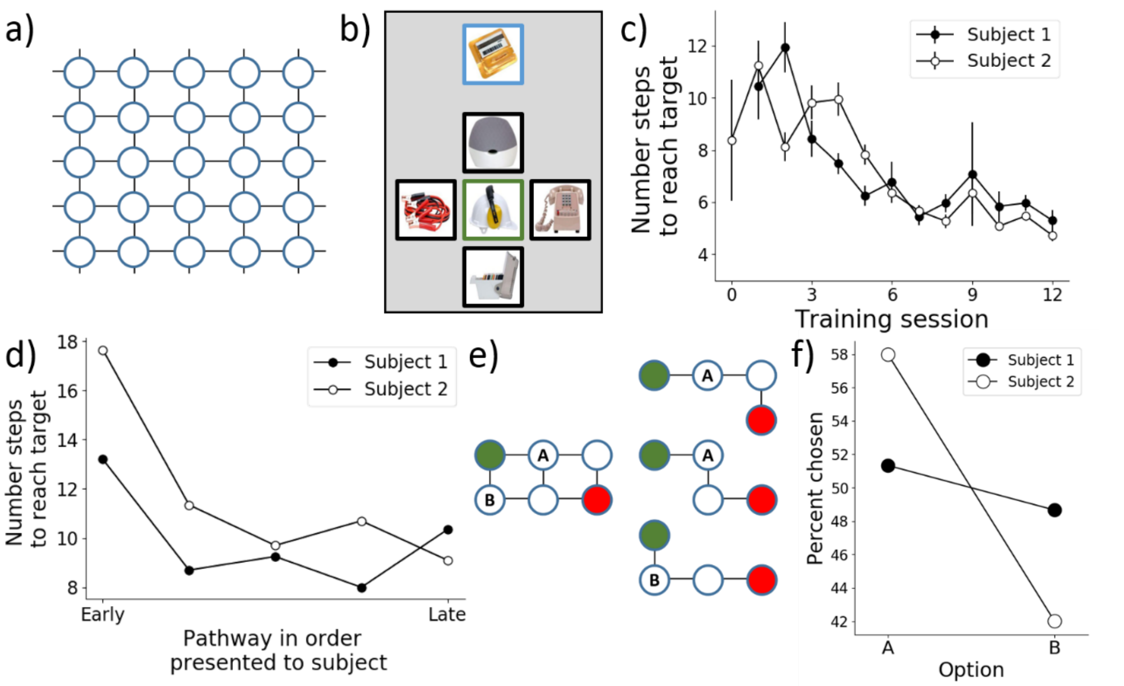

***Figure 1. a)** One of the environments taught to the subjects, made up of many stimuli (circles) connected to their neighbours (lines). **b)** The experimental task consisted of a target location (blue) which they had to reach from their current location (green) by choosing any of the neighbouring stimuli (black) until they reached the target. The location of the black stimuli was randomised so that the task provides no spatial information. **c)** Macaques learned to navigate the maps after a week of training. **d)** When presented with a novel start-target trial they had not experienced before, subjects performed better later in the training process by taking a more direct route to the target location. **e)** Start-target combinations of distance 3 could be used to assess strategy by looking at the ratio the subjects went via node A or B. **f)** Both subjects went via node A more often.*

NHPs could therefore utilize partial path information from previously experienced paths to optimize performance on novel paths, hence reflecting a global representation of the environment and allowing them to infer optimal routes without ever having directly experienced them.

Using a discrete environment affords high mathematical precision over the behavioural data collected, enabling us to analyse in depth the strategies employed by the subjects to navigate the environment. Using a 5x5 grid wrapped in a torus (Figure 1A) meant that all start-target combinations of distance 3 shared an asymmetric structure with three possible routes to the target location (Figure 1E). Two of these three routes shared the same first node (option A), whereas the third route went via a different node (option B) on the first step. An ideal strategy would be to pick option A over option B as this is more likely to lead to reward. However, with 200 unique pathways containing this structure this is cognitively demanding to do. It would therefore be difficult to compute this using a tree search-like strategy which would have to search each of the different options available. Alternatively, a successor representation (SR)-like strategy that assigns a value to each option based on the likelihood of possible future states leading to reward would be better suited to performing such a task. We therefore compared the ratio of A versus B choices to investigate the subject's ability to employ such a strategy. Both subjects preferentially chose option A over option B (Figure 1F). This was also consistent across multiple maps tested. To be capable of performing such a strategy suggests NHPs employ an efficient method of navigating, such as an SR-like strategy, that weights future states associated with each option, rather than a tree search-like strategy which would be unlikely in such a large state space.

We have therefore used a novel cognitive task to demonstrate that NHPs are capable of learning to navigate large 2D abstract environments with high precision in a short time frame. Their learning rate increases with training, even on novel start-target routes they have never experienced before, suggesting they learn information about the structure of the environment through experience. We have also shown that they employ strategies to maximise their reward which require an efficient way of evaluating different states, such as an SR-like strategy.