Isolating Motor Learning Mechanisms in Embodied Virtual Reality

Federico Nardi (f.nardi21@imperial.ac.uk)
Department of Computing, Imperial College London,
London, SW7 2AZ United Kingdom

Mabel Ziman (m.ziman21@imperial.ac.uk)
Department of Bioengineering, Imperial College London,
London, SW7 2AZ United Kingdom

Shlomi Haar (s.haar@imperial.ac.uk)
Department of Brain Sciences, Imperial College London,
London, W12 0BZ United Kingdom

A. Aldo Faisal (a.faisal@imperial.ac.uk)
Department of Computing & Department of Bioengineering, Imperial College London,
London, SW7 2AZ United Kingdom
Institute of Artificial & Human Intelligence, University of Bayreuth,
Bayreuth, 95447 Germany

Abstract

We developed an embodied Virtual Reality (eVR) environment and paradigm for studying sensorimotor learning mechanisms in real-world tasks. By using a real-world setting, i.e. playing pool billiard, we showed previously that subjects could learn by error-based and reward-based learning mechanisms, yet actually split into two distinct groups using one or the other mechanism primarily. Now, we isolated *Error-based* and *Reward-based* reinforcement learning by providing different partial visual feedback to two groups of subjects (that clamped either error or reward feedback), to understand how and whether the learning processes were different between groups. Visual feedback was manipulated in form of a rotation of the target ball's trajectory, that requires from learners a cognitive abstraction that acts only on an object, in contrast to traditional visuomotor learning experiments which rotate the whole visual field. We found that subjects' compensation for the perturbation and the learning curve differed between groups, with the *error-based* subjects showing exponential learning curve and achieving a higher improvement than *reward* subjects who presented a linear learning curve. Yet, no significant difference was found in inter-trial variability and success rate.

Keywords: real-world neuroscience; embodied virtual reality; motor learning; visuomotor rotation; error-based learning; reward-based learning

Background

Motor learning is driven by two main learning processes categorized broadly as *error-based* learning and *reward-based* learning (Doyon, Penhune, & Ungerleider, 2003; Krakauer, Hadjiosif, Xu, Wong, & Haith, 2019), which can be isolated by giving selective feedback (Therrien, Wolpert, & Bastian, 2016; Uehara, Mawase, & Celnik, 2018; Palidis, Cashaback, & Gribble, 2019). While those are often isolated in simple lab based tasks, it is not trivial to dissociate between them in the real-world. Here we address the challenge to test the wider applicability of classic results originally obtained in restrictive laboratory settings to more complex tasks (Krakauer, Pine, Ghilardi, & Ghez, 2000; Ingram & Wolpert, 2011).

In previous works we established the game of pool billiards as a real-world paradigm to study motor learning, tracking the balls on the table to measure task performance, and using both motion tracking and electroencephalography (EEG) to capture full-body movement and brain activity in the task (Haar, van Assel, & Faisal, 2020; Haar & Faisal, 2020). We then incorporated it in an embodied Virtual Reality (eVR) environment (Haar, Sundar, & Faisal, 2021) to enable perturbations and visual manipulations. In the eVR setup, the subject is physically playing pool while a visual feedback is provided by the VR headset. One possible way to manipulate a task is to use visuomotor rotations, process that has often been used to initiate new learning processes (Krakauer et al., 2000). We previously demonstrated that during real-world learning in the pool task different individuals use varying contributions of error-based and reward-based learning mechanism (Haar & Faisal, 2020). Here, we use the eVR environment of a real-world pool task to introduce visuomotor rotations and selectively provide error-based or reinforcement learning feedback, as was previously done in lab tasks, to verify if and how subjects can learn in the real-world using single mechanisms.

Methods

Experimental Setup The experimental setup (fig. 1a) is an extension of that used in Haar et al. (2021), with a real environment represented simultaneously in VR. A pool table is positioned below four *Optitrack* cameras that stream the real cue stick in the virtual environment. The cue ball is both real and virtual, in order to provide natural haptic feedback of the shot, whereas the target ball is only present in VR.

Experimental Design 16 healthy right-handed volunteers with normal or corrected-to-normal vision (4 males and 12 females, aged 20-25) were randomly divided into two equal groups. The experiment included 250 trials divided in 10 blocks of 25 shots: a baseline phase (75 trials), a perturbation phase (150 trials) and a final washout block (25 trials). The same shot was performed repeatedly, aiming at pocketing the target ball positioned close to the far-left corner of the pool table. During the perturbation phase, the trajectory of the cue ball was rotated by 5° counterclockwise (fig. 1b). The participants then had to correct, to compensate for the directional bias, aiming to the right in order to pocket successfully. During the rotation phase a group-specific feedback was provided in VR. In the error-based feedback group, the trajectories of the balls were hidden after their collision, so that subjects could improve only based on the trajectory of the cue ball towards the target ball. In the reward feedback group, in contrast, the balls were hidden once the cue ball was hit. In the case of a successful shot, a *reward* was provided by showing the subject an artificial successful trajectory. This constitutes reinforcement independent of any possible error correction. Fol-

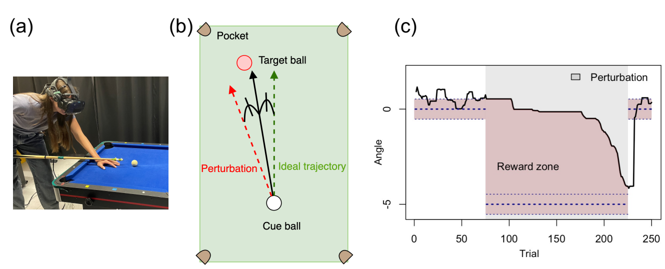

Figure 1: Experimental setup: (a) the Pool eVR setup; (b) 5° perturbation applied on the cue ball; (c) a dynamic reward zone (pink) is defined for the perturbation phase in the reward task. The solid black line represents the median of past ten rewarded trials. The dashed lines represent the ideal target and corresponding funnel (0° for baseline, -5° for perturbation).

lowing Therrien et al. (2016), the reward zone was defined by two criteria: (i) when shooting inside the success funnel around the pocket (just as in baseline blocks or during the error task) and (ii) when a shot was better than the median of the past 10 rewarded trials. A threshold for outliers removal was set on the baseline ($\mu \pm 3\sigma$), and applied to the perturbation learning zone, i.e. 3σ higher than 0° or lower than -5°.

Results

Figure 2 shows the trial-by-trial average error across subjects for the two different groups, as well as fits of the points to a decay curve. In the baseline phase, the two conditions lead to similar behaviours, which start changing during perturbation.

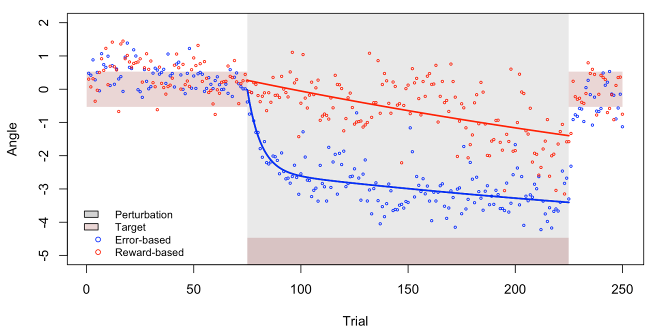

Figure 2: Mean cue ball angle of error-based subjects (blue) and reward-based subjects (red). Gray area represents the presence of perturbation, while pink area indicates the successful angles. Double exponential fit is in bold.

Despite an overall similar trend for the two groups, the correction for the perturbation is lower for the reward subjects than for error ones. The error-based group showed a classic double exponential learning curve with a fast component with time constant of $\tau_f = 6.25$ trials and a slow with $\tau_s = 325$. The reward-based group, presumably due to the reward regime, showed no fast learning and the slow component was linear.

Over blocks (of 25 trials) there is a clear difference between groups in the average angle (figure 3a) but not difference in the between subject variability. The washout block demonstrates slightly higher retention for the error group, in line with the higher compensation for the perturbation. There was no group difference in success rate over blocks (fig. 3b) between the groups (mixed-effects ANOVA during rotation: $F(1,14) = 0.214, p = 0.65$). Importantly, during the rotation phase, error group successes were counted based on minimal error from the ideal rotated trajectory, whereas for the reward group the success count was based on reward feedback, hence covering the full reward zone. The reward participants, had significantly higher variability in the success rate between blocks and subjects, due to the dynamic reward zone.

Finally, in order to understand the variability between trials at subject-level, we considered the corrected standard deviation within blocks, which accounts for and isolates the linear trends to derive a clearer measure of change (for methods see Haar et al. (2020)). The variability (fig. 3c) showed a rough decreasing trend over trials (though not significant) for the reward group but not for the error group (mixed-effects ANOVA interaction (group x block): $F(5,70) = 1.754, p = 0.134$).

Discussion

In this paper, we present an embodied Virtual Reality environment which allows the study of motor learning in a real-world framework. The sense of embodiment given by the setup favours the preservation of complex natural movements, and the complex nature of the real world-task, involving multiple objects, increases its cognitive demand. Hence, this can better generalise to the real-world as cognitive effort has critical impact on learning process (Lee, Swinnen, & Serrien, 1994).

Due the different feedback regimes, different computational learning mechanisms with different cognitive demands were used by the different groups. Presumably, the unknown success criteria as well as the disappearance of the trajectories could have made the cognitive demands of the reinforcement task higher. Yet, although in the real-world humans use a combination of *error-based* and *reward-based* learning (Haar & Faisal, 2020), here we demonstrate that with restricted feedback subjects can learn to correct for a perturbation in a real-world task with either of these computational mechanisms.

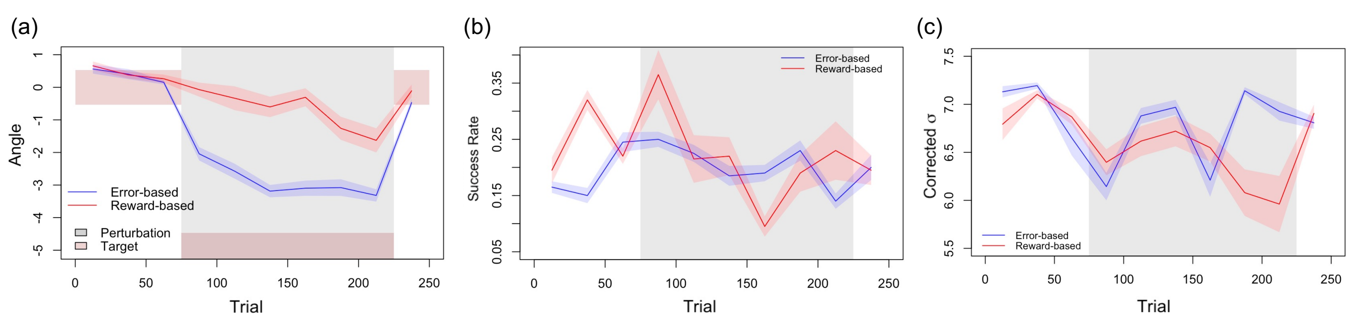

Figure 3: (a) Mean cue ball error over blocks; (b) Success rate over blocks (includes the entire reward zone during perturbation for the reward group). (c) Corrected standard deviation (after removal of linear trends within blocks). All calculated over blocks of 25 trials and averaged over subjects within group. Shaded areas represent inter-subjects variability. Blue: *error-based* group; Red: *reward-based* group.

Acknowledgments

F.N. is supported by UK Research and Innovation [UKRI Centre for Doctoral Training in AI for Healthcare grant number EP/S023283/1]. S.H. is supported by the Edmond and Lily Safra Fellowship Program and by the UK Dementia Research Institute Care Research & Technology Centre.

References

Doyon, J., Penhune, V., & Ungerleider, L. G. (2003). Distinct contribution of the cortico-striatal and cortico-cerebellar systems to motor skill learning. *Neuropsychologia*, *41*(3), 252–262.

Haar, S., & Faisal, A. A. (2020). Brain activity reveals multiple motor-learning mechanisms in a real-world task. *Frontiers in human neuroscience*, 354.

Haar, S., Sundar, G., & Faisal, A. A. (2021). Embodied virtual reality for the study of real-world motor learning. *Plos one*, *16*(1), e0245717.

Haar, S., van Assel, C. M., & Faisal, A. A. (2020). Motor learning in real-world pool billiards. *Scientific reports*, *10*(1), 1–13.

Ingram, J. N., & Wolpert, D. M. (2011). Naturalistic approaches to sensorimotor control. *Progress in Brain Research*, *191*, 3–29.

Krakauer, J. W., Hadjiosif, A. M., Xu, J., Wong, A. L., & Haith, A. M. (2019). Motor learning. *Compr Physiol*, *9*(2), 613–663.

Krakauer, J. W., Pine, Z. M., Ghilardi, M.-F., & Ghez, C. (2000). Learning of visuomotor transformations for vectorial planning of reaching trajectories. *Journal of neuroscience*, *20*(23), 8916–8924.

Lee, T. D., Swinnen, S. P., & Serrien, D. J. (1994). Cognitive effort and motor learning. *Quest*, *46*(3), 328–344.

Palidis, D. J., Cashaback, J. G., & Gribble, P. L. (2019). Neural signatures of reward and sensory error feedback processing in motor learning. *Journal of neurophysiology*, *121*(4), 1561–1574.

Therrien, A. S., Wolpert, D. M., & Bastian, A. J. (2016). Effective reinforcement learning following cerebellar damage requires a balance between exploration and motor noise. *Brain*, *139*(1), 101–114.

Uehara, S., Mawase, F., & Celnik, P. (2018). Learning similar actions by reinforcement or sensory-prediction errors rely on distinct physiological mechanisms. *Cerebral Cortex*, *28*(10), 3478–3490.