

## Autonomous behaviour and the limits of human volition

Keiji Ota & Patrick Haggard

Institute of Cognitive Neuroscience, University College London

The nature of volition and autonomy is one of the longest-running scientific debates. Philosophical accounts of free action distinguish between **freedom from** external constraint and **freedom to** achieve personal goals and desires. Many studies of volition focussed on the “freedom from” aspect of voluntary action by creating conditions where action is or is not independent of an external triggering stimulus. However, such studies neglect the reasons or goals that an agent may have for acting, and omit the crucial “**why**” aspect of voluntary action. Here we propose that some competitive game scenarios can capture both the “freedom from” and “freedom to” aspects of voluntary action. For example, in the ‘rock, paper, scissors’ game, each player selects an action without first seeing the action of their opponent, while at the same time, innovating and discovering a specific action that defeats their opponent. In this study, we examined how free and how flexible such reasons-guided voluntary actions can be by identifying whether people can become free from **internal constraints** on behaviour in adaptation to competitive pressures.

Participants chose when to make a simple key press (Libet et al., 1983) while trying to avoid colliding with a virtual competitor. The timing of the events in the game was set so that participants could not simply respond to the competitor, so the participants’ actions were already free from external triggers in this sense. In addition, three competitor algorithms were designed to predict and punish stereotypical patterns of action timing, pressurising participants to become free from a corresponding exploitative strategy. This operation allowed us to quantify three distinct forms of behavioural autonomy, defined freedom from (1) repetition of the same action (such as always responding early), (2) rule-based sequential action choices (such as early, late, early, late...) and (3) outcome-dependent action choices (such as win-stay lose-shift).

We found that people are surprisingly good at becoming free from repeating the same action timing, and can even avoid sequential biases in transition from one action to the next when punished. Conversely, people had limited ability to become free from the influence of previous action outcomes on current action choices and did not achieve autonomy with respect to reinforcement history. A follow-up experiment confirms that the mechanism of behavioural adaptation seems not to require conscious awareness about constraints of being pressurised, in contrast to classical ideas about the relation between volition and conscious, rational thought. Rather, becoming free from internal constraints primarily relies on an implicit process where participants develop a model of the opponent by learning what the opponent is going to do and guide their own action choices accordingly. This suggests an interesting convergence between theory of mind, model-based belief learning, and voluntary action. Finally, we will discuss what readiness potential truly reflects by comparing stereotypical, repeated behaviour (i.e., seemingly less volitional action) with innovative, non-stereotypical behaviour (i.e., more volitional action). Overall, our work provides a novel scientific investigation of human volition and highlights reasons-responsive flexible behaviour in the view of volition and autonomy.