1 # Getting stuck in a rut as an emergent feature of a

2 # dynamic decision-making system

3 Matthew Warburton[1], Jack Brookes[1, 2], Mohamed Hasan[3], Matteo Leonetti[3], Mehmet

4 Dogar[3], He Wang[3,4], Anthony G. Cohn[3], Faisal Mushtaq*[1,4] & Mark A. Mon-

5 Williams[1,4,5,6]

6 [1]School of Psychology, University of Leeds, Leeds, West Yorkshire, United Kingdom

7 [2]Max Planck UCL Centre for Computational Psychiatry and Ageing Research, University College
8 London, London, United Kingdom

9 [3]School of Computing, University of Leeds, Leeds, West Yorkshire, United Kingdom

10 [4]Centre for Immersive Technologies, University of Leeds, Leeds, West Yorkshire, United Kingdom

11 [5]Centre for Applied Education Research, Wolfson Centre for Applied Health Research, Bradford
12 Teaching Hospitals NHS Foundation Trust, Bradford, West Yorkshire, UK

13 [6]National Centre for Optics, Vision and Eye Care, University of South-Eastern Norway, Kongsberg,
14 Hasbergs vei 36, 3616, Norway

15

16 *Correspondence should be addressed to Faisal Mushtaq (f.mushtaq@leeds.ac.uk), School of

17 Psychology, University of Leeds, Leeds, United Kingdom.

Abstract

25

26   Human sensorimotor decision-making has a tendency to get 'stuck in a rut', being biased towards

27   selecting a previously implemented action structure ('hysteresis'). Existing explanations cannot

28   provide a principled account of when hysteresis will occur. We propose that hysteresis is an

29   emergent property of a dynamical system learning from the consequences of its actions. To

30   examine this, 152 participants moved a cursor to a target on a tablet device whilst avoiding an

31   obstacle. Hysteresis was observed when the obstacle moved sequentially across the screen between

32   trials, but not with random obstacle placement. Two further experiments (n = 20) showed an

33   attenuation when time and resource constraints were eased. We created a simple computational

34   model capturing dynamic probabilistic estimate updating that showed the same patterns of results.

35   This provides the first computational demonstration of how sensorimotor decision-making can get

36   'stuck in a rut' through the dynamic updating of its probability estimates.

37 **Significance Statement**

38 Humans show a bias to select the organisational structure of a recently carried out action, even

39 when an alternative option is available with lower costs. This 'hysteresis' is said to be more efficient

40 than creating a new plan and it has been interpreted as a 'design feature' within decision-making

41 systems. We suggest such teleological arguments are redundant, with hysteresis being a naturally

42 emergent property of a dynamic control system that evolved to operate effectively in an uncertain

43 and partially observable world. Empirical experimentation and simulations from a 'first principle'

44 computational model of decision-making were consistent with our hypothesis. The identification of

45 such a mechanism can inform robotics research, suggesting how robotic agents can show human-

46 like flexibility in complex dynamic environments.

Introduction

47

48   Humans are creatures of habit and often repeat behaviours - despite the selected action having

49   a higher cost than an available alternative. This propensity can be seen when humans continue to

50   use the road well-travelled when moving between two buildings even after construction work has

51   created a shorter route. The phenomenon is particularly remarkable because adult humans are

52   generally so adept at selecting optimal movement patterns (Trommershäuser et al., 2008). Indeed,

53   the ability of humans to rapidly and efficiently execute actions far exceeds the capabilities of even

54   the most sophisticated robotic systems (Dogar & Srinivasa, 2012). The incredible repertoire of

55   skilled behaviour in humans reflects the presence of learning processes that have been trained over

56   the countless occasions when adults have interacted with the external world. These myriad

57   interactions allow the human nervous system to accurately estimate the costs associated with

58   various behaviours and thereby select an optimal (or close to optimal) action when presented with

59   a goal directed task. The issue of relevance within this manuscript relates to the observation that

60   adult humans will select different options on different occasions as a function of whether the choice

61   is made *de novo* or following a previous successful action – despite the choices having the same

62   relative costs on both occasions.

63   The tendency to show a bias towards a previously selected action plan can be described as

64   'hysteresis' (or the sequential effect) and is well-studied. Hysteresis effects have been found in

65   grasp selection (Cohen & Rosenbaum, 2004, 2011; Dixon et al., 2012; Kelso et al., 1994; Kent et al.,

66   2009; Rosenbaum & Jorgensen, 1992; Schütz et al., 2011; Short & Cauraugh, 1997; Weigelt et al.,

67   2009), hand selection (Rostoft et al., 2002; Schweighofer et al., 2015; Weiss & Wark, 2009), and

68   hand path priming experiments (Jax & Rosenbaum, 2007, 2009; van der Wel et al., 2007). However,

69   there are no satisfactory explanations to account for this phenomenon. In fact, most explanations

70   are teleological in nature: it is proposed that modifying a previously used action is more cognitively

71   efficient than planning from scratch, so hysteresis exists to increase planning efficiency

72 (Meulenbroek et al., 1993; Rosenbaum et al., 2007; Schütz & Schack, 2019; Weiss & Wark, 2009), as

73 indexed by reduced reaction times (RTs) when using the same action as previously (Valyear et al.,

74 2018). The problem is that such interpretations do not provide a principled account that can

75 explain when hysteresis will occur, why its magnitude differs under different task constraints, or

76 why its presence is a function of the costs of the available choices.

77   We propose that the process of 'getting stuck in a rut' is an emergent property of a decision-

78 making system that dynamically learns from the consequences of its actions. In order to deal with

79 unpredicted changes in the world (and adapt to novel environmental states), an efficient system

80 must frequently update its estimates of the success probabilities associated with a given action (in

81 Bayesian terms, the system must continually update its priors). To update these estimates, humans

82 must use feedback about the outcomes of their actions – actions which cause the environment to

83 transition to a new state. We suggest that this principle – the updating of success probabilities – will

84 naturally result in a system that shows hysteresis. In fact, it is common practice in computer science

85 to model the environment as a POMDP (POMDP; Kaelbling et al., 1998) when designing agents that

86 need to act under uncertainty. In a POMDP, the agent does not directly observe the environment's

87 state but receives an observation which is a function of the state of the environment following an

88 action executed by the agent. POMDPs reflect well the challenges faced by the human nervous

89 system which must infer the hidden states of the environment from the sensory inputs that follow

90 an action (as a Markov blanket separates the nervous system from the external world; Friston,

91 2010). The important point from the perspective of this manuscript is that the external world is not

92 static and this means that a human agent must frequently update its internal representation (i.e. the

93 approximate conditional density on the causes of sensory input) in order to act optimally in a noisy

94 and changing world. The dynamical updating of the internal representation enables the human to

95 predict the sensory input that will result from a generated action. The ability to make accurate

96 predictions allows a human to generate an action that will produce a desired change in the sensory

97 input (i.e. achieve a goal-directed change in the external environment). It follows that efficient

98 action selection requires frequent updating of the success probabilities associated with a given

99 action. We hypothesised that this updating would produce hysteresis as an emergent property of

100 the dynamical learning system.

101 In order to explore hysteresis, we needed to design a canonical task that would allow us to

102 parametrically vary critical task parameters, reflect a naturalistic action, and produce data

103 amenable to computational modelling. We also needed a task that would allow us to examine

104 behaviour on a trial-by-trial basis so that we could address issues relating to the frequency of

105 updating (e.g. whether we would observe hysteresis on a trial-by-trial basis or whether it was only

106 manifest after a number of iterations of a given action structure). We decided to use aiming

107 movements (moving an end effector from a start point to a target location) to meet our task

108 requirements. We therefore created a simple multi-trial sensorimotor decision-making task in

109 which participants needed to move around an obstacle (left or right) to hit a target (where the

110 reward was the same for each choice). We manipulated obstacle position on each trial across blocks

111 such that it either moved systematically across the screen or was randomly positioned across trials

112 (Figure 1). We predicted that participants would show a strong bias towards repeating previously

113 selected actions in the sequential condition, even when the obstacle position indicated an

114 alternative route would be preferable. We expected that this effect would be diminished in blocks

115 where obstacle position moved randomly across the screen.

116 We wished to explore whether the empirical data generated through our empirical

117 investigations could be captured by a model that incorporated dynamic probabilistic estimate

118 updating (i.e. whether hysteresis could be captured through a POMDP type model). We therefore

119 created a model of human decision-making (Figure 2) that included a trial-by-trial update of the

120 success probabilities associated with one action versus another. The goal of the model was to

121   simulate how an agent would respond to a choice between two options that both allow a given goal

122   to be achieved but have different costs. The output of the model was an action that would cause the

123   environment to transition (with a given probability) to a new state. The model was arranged such

124   that after an action is executed the agent receives an observation which is a function of the new

125   environmental state. This input was then used to update the success probabilities associated with

126   the action. In many choice tasks, there is also a difference in the reward associated with the options

127   (Dreher & Tremblay, 2009; Gold & Shadlen, 2007; Mushtaq et al., 2016), so our model incorporated

128   an estimate of the reward to allow future studies to explore behaviour in such tasks (but in the

129   reported experiments the reward was identical across options).

130   We were also interested in exploring the dynamical aspects of decision-making under

131   temporal constraint. Converging evidence suggests that the decision-making process is governed by

132   neural circuits that accumulate noisy evidence for possible options over time, with a decision

133   triggered when sufficient evidence is accumulated to cross an action threshold (Bogacz et al., 2006;

134   Brody & Hanks, 2016; Gold & Shadlen, 2007). It seems reasonable to suppose that the action

135   threshold will be a function of the available time period within which a decision must be made (i.e.

136   temporal constraints will push the system towards making a choice that might be different were

137   more time available to weigh up the respective costs of the different options). Notably, the existence

138   of evidence accumulation processes predicts that actions will be selected more rapidly (i.e. RTs will

139   decrease) when there is a bias towards one action versus another. This suggests that RTs will be

140   faster in the presence of hysteresis. Once more, it is important to emphasise that we are proposing

141   that hysteresis is an emergent property of a dynamic learning system where faster RTs are a useful

142   by-product of the system's organisation rather than the planned product of a system designed with

143   an inbuilt function to produce hysteresis.

144   On the assumption that evidence accumulation processes are a core component of

145   sensorimotor decision-making, we anticipated finding RT differences as a function of the magnitude

146   of hysteresis associated with a given task. We further hypothesised that relaxing the temporal

147   constraints of the task would attenuate the size of the hysteresis effect (as the available time can be

148   used to more fully evaluate the costs of either action, reducing the reliance of the decision-making

149   system on previous successes and failures to inform current action selection). In Experiment 2, we

150   directly manipulated the temporal constraints of the task by creating a 'waiting period' before

151   which an action could be executed. In Experiment 3, we indirectly manipulated the temporal

152   constraints by decreasing the 'higher order' cognitive demands of the task. We reasoned that

153   decreasing the cognitive constraints would allow the task goal to be identified more rapidly and

154   thereby create a longer period in which the respective costs of the alternative actions could be

155   computed.

156

157   **Methods**

158   **Participants**

159   In Experiment 1, 152 adults (41 males, 100 females; mean age 22.51 years, range 18– 39 years; 139

160   self-reported right-handed; eleven participants did not report age or gender) were recruited as part

161   of a larger motor control project. Participants for Experiment 2 (n = 20, 1 male, 19 females; mean

162   age 19.09, range 18-20 years; 20 self-reported right-handed) and Experiment 3 (n = 20, 1 male, 19

163   females; mean age 18.86 years, range 18-20 years; 18 self-reported right-handed) were recruited

164   through word of mouth from the University of Leeds undergraduate population. All had normal or

165   corrected-to-normal vision and provided informed consent to participate. Participation in these

166   studies was incentivised through remuneration of £2 on completion of the experiment. Ethical

167   approval was obtained from the University of Leeds ethics committee.

168   **Procedure**

169   Participants sat at a desk with a touchscreen computer tablet (Lenovo ThinkPad Helix 2,

170   1920x1080 pixels, 11.6" screen, 60Hz refresh rate) placed directly in front of them and interacted

171   with the screen using a stylus (sampled from the screen digitiser at a rate of 100Hz) in their chosen

172   hand. Participants were shown a pictorial instruction sheet prior to starting the experiment that

173   explained how to complete a single trial. Prior to starting the experiment, participants were

174   instructed that each trial should be completed as quickly and accurately as possible. The core trial

175   structure was the same across all three experiments and key differences for each study are detailed

176   below and presented in Figure 1.

177

178   *Experiment 1: Biases in Sensorimotor Decision-making*

179   In Experiment 1, we introduced a novel sensorimotor decision task in which participants were

180   asked to select one of two possible routes around an obstacle to reach a target using a stylus on a

181   tablet display. During each trial, participants had to stay within a 200mm high by 106mm wide

182   workspace, displayed as a rectangle on the screen. Participants began each trial by placing the

183   stylus on the screen and moving the cursor (5mm diameter circle) to a start-point (10mm diameter

184   circle, horizontally central) at the bottom centre of the screen. Events in the scene (e.g. entering the

185   start-point) were triggered when the perimeter of the cursor intersected the perimeter of another

186   object. After a randomly sampled delay (generated by a random number generator) between 300-

187   600ms (across all participants and experiments, there was a grand mean delay of 448ms (SD =

188   9ms), with the mean average delay across participants ranging from 423– 475ms), the start-point

189   changed to a colour and shape combination, randomly selected from a list of three of each (cyan,

190   magenta, yellow; circle, square, triangle); a check-point (10mm diameter circle, horizontally

191   central) appeared in the top centre of the screen; and an obstacle (30mm wide x 10mm high

192   rectangle) appeared equidistant between the start- and check-points, vertically central to the

193   screen. The distance between the start and check-points was 140mm. Participants were instructed

194   to remember the colour they were shown and move as quickly and accurately after stimulus display

195   to the checkpoint. Participants were allowed up to 2 seconds 'preparation time' in the start-point.

196   Upon leaving the start-point the coloured shape disappeared, and the participant had up to 500ms

197   to reach the check-point.

198   Upon entering the check-point, the obstacle disappeared. The participant then had to wait in

199   the check-point. After a randomly sampled delay (generated by a random number generator)

200   between 300-600ms (across all participants and experiments, grand mean delay = 450ms, SD =

201   11ms, range = 418–476ms), three targets at the bottom of the screen appeared, spaced equally in

202   the horizontal axis. Each target had an invisible 10mm diameter circle used to detect the cursor

203   hitting. The vertical distance between the check-point and targets was 150mm, with 30mm spacing

204   between targets horizontally. Each target had a randomly assigned colour and shape combination,

205   selected by randomly shuffling the list of three colours and three shapes and allocating the

206   combinations to each target. Participants were instructed to move as quickly and accurately to the

207   target of the same colour that was shown at the start-point. Participants were allowed up to 2

208   seconds in the check-point. After leaving the check-point the participant had up to 1 second to

209   reach the target.

210   A trial was successfully completed when the participant moved to the correct target. There

211   were several ways to fail a trial: hitting the obstacle or task boundaries; spending too long in the

212   start or checkpoint; moving too slowly between the start- and check-points, or moving too slowly

213   between the check-point and target; leaving the checkpoint before the targets were revealed; and

214   moving to the wrong coloured target. If a failure was triggered the trial was immediately

215   terminated. Once a trial finished, visual feedback was presented for 1 second to indicate the

216   outcome of a trial, where the target turning green indicated a successful trial, and an object turning

217   red indicated a failed trial. For failed trials, the object that turned red indicated the type of failure

218   (e.g. if the participant hit the obstacle, it would turn red). Across all trials a running score was

219   shown in the top left of the screen, which increased by one after each successful trial. After visual

220   feedback had been presented for 1 second, the start-point was shown on the screen, and

221   participants were able to begin a new trial. During the experimental trials, the mean within-subject

222   time between the start of successful trials was 4.17 seconds, and 3.67 seconds for unsuccessful

223   trials.

224   The experiment comprised a total of 106 trials and took approximately 10 minutes to

225   complete. This included 4 example trials, 6 practice trials, 9 baseline trials, and 87 experimental

226   trials made up in 3 blocks of 29 trials, shown in Figure 1d. During the example trials, the instructor

227   showed the participant a set of 4 standard trials including two successful trials and two failed trials.

228   Participants were provided 6 practice trials to make sure they understood the task mechanics,

229   which included text feedback after every trial to indicate the outcome, in addition to the regular

230   visual feedback. A baseline block followed that aimed to make participants move as quickly as they

231   could (while maintaining accuracy) by giving text feedback telling them they needed to move more

232   quickly if their movement time between the start- and check-point was slower than their previous

233   fastest time (following the baselining block, across all experiments the within-subject mean

234   movement time between the start and check-point = 369ms, SD = 47ms, range = 249 – 452ms).

235   During each of these three blocks, the obstacle was located horizontally central to the screen, and

236   after the completion of each block, text was displayed for 10 seconds on the screen to indicate the

237   participant was starting a new phase of the experiment.

238   The experimental trials were then organised into three blocks, presented to the participant as

239   one uninterrupted block. The three conditions were where the obstacle's horizontal position moved

240   sequentially between trials from the left of the screen to the right (Rightward), from the right of the

241   screen to the left (Leftward), and where the obstacle's positions were randomly shuffled (Random).

242   The order of the three conditions was randomly allocated per participant. Each obstacle position

243   was presented once per block. Twenty-nine obstacle positions were used with extreme positions of

244   -34.2mm and 34.2mm, with equally spaced jumps between each position.

245   The experimental task was developed using Unity (Unity Technologies, 2018; version 2018.1)

246   and the Unity Experiment Framework (Brookes et al., 2019).

247

248   ***Experiment 2: Decreasing Temporal Constraints***

249   Experiment 2 was conducted to explore whether easing the temporal constraints placed on the

250   decision-making system could attenuate hysteresis. In Experiment 2, participants were forced to

251   wait while the stimuli were shown before executing the movement. Upon entering the start-point a

252   red box appeared surrounding the start point and the participant's cursor. While the red box was

253   visible the participant was not allowed to leave the start-point or the trial would terminate in

254   failure. In common with Experiment 1, there was a randomly sampled delay of between 300-600ms

255   before the stimuli were presented. However, in Experiment 2 the red box remained on the screen

256   after stimulus display. After 1.5 seconds the red box disappeared, and the participant completed the

257   rest of the trial as described in Experiment 1. Differences between these Experiments are

258   illustrated in Figure 1b. In Experiment 2, the mean within-subject time between the start of

259   successful trials was 5.50 seconds, and 4.22 seconds for unsuccessful trials in the experimental

260   block.

261

262   ***Experiment 3: Reducing Task Cognitive Demands***

263   Experiment 3 was conducted to explore whether reducing the cognitive demands associated

264   with the task could attenuate hysteresis. Participants were presented with only one target after

265   waiting in the check-point, with the colour shown at the start of the trial always matching that of

266   the target. The remainder of the trial followed the same structure as Experiment 1 (differences

267   illustrated in Figure 1c). In Experiment 3, the mean within-subject time between the start of

268   successful trials was 3.96 seconds, and 3.30 seconds for unsuccessful trials in the experimental
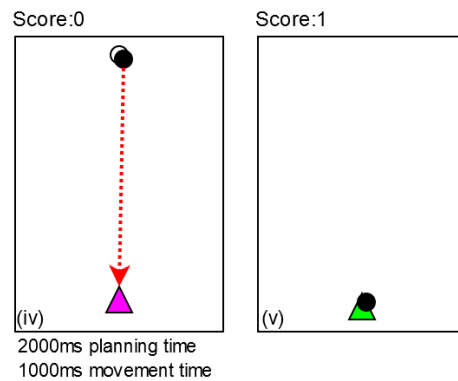
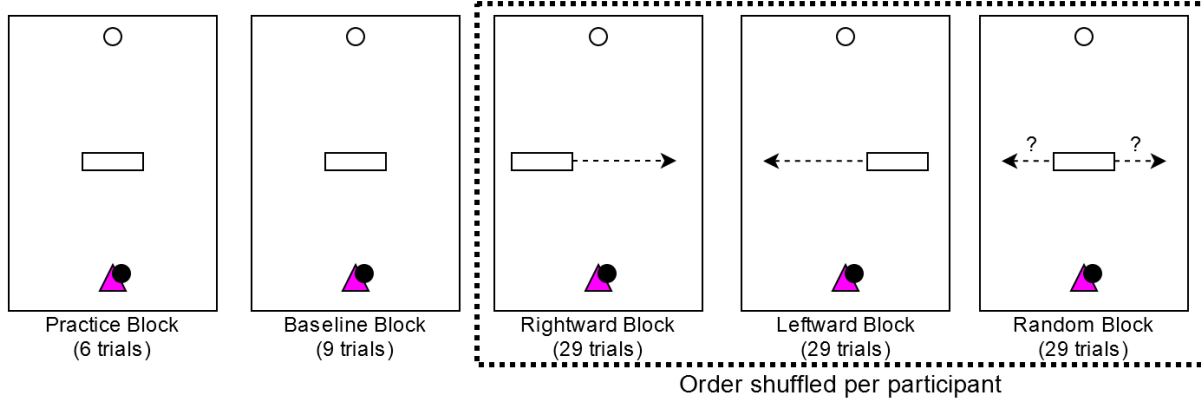269   block.
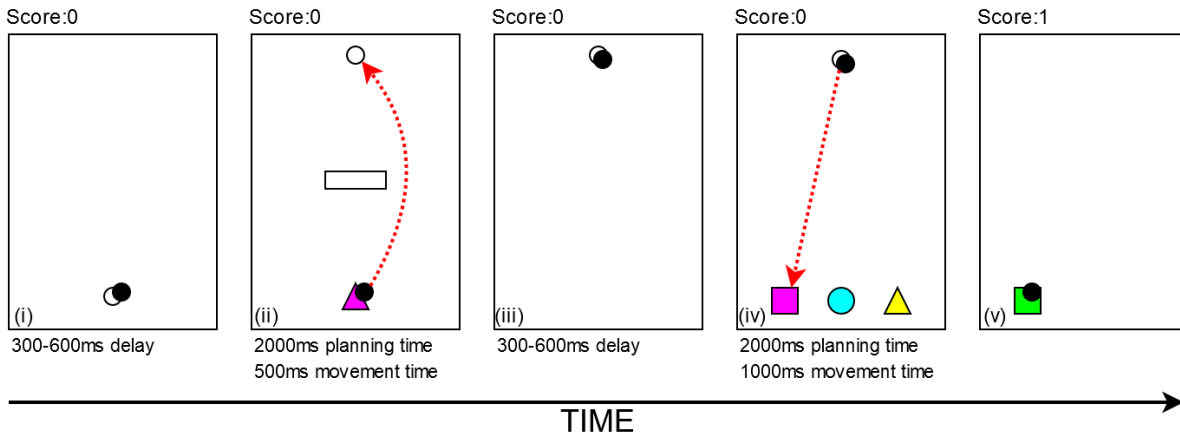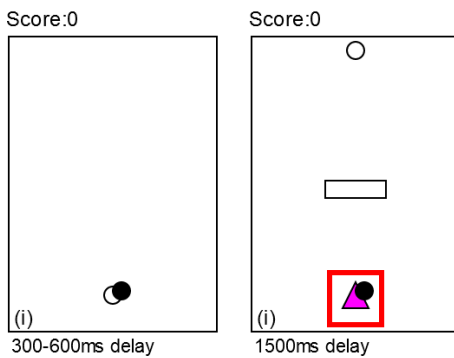
Figure 1. Trial and block structure of the experiments. (a) A complete trial for Experiment 1. Red dashed lines indicate potential movement trajectories and filled circle indicates movement endpoint. In Step (i), participants moved to a start-point and waited 300-600ms until it changed to a colour and a shape indicating the target shape colour. Simultaneously, an obstacle and a checkpoint appeared. Participants were allowed 2000ms planning time at the start-point before moving around the obstacle to the checkpoint (<500ms) and waited 300-600ms until 3 targets appeared. Participants were allowed up to 2000ms in the checkpoint before moving to the target that matched the colour shown at the start-point (<1000ms). (b) For Experiment 2, step (i) of Experiment 1 was replaced by two steps. Participants moved to a start-point and were immediately shown a red box around the start-point, indicating they could not leave. After a random 300-600ms delay, the stimuli were revealed but the red box remained on screen for a further 1500ms. (c) For Experiment 3, steps (iv) and

281 (v) of experiment 1 were changed so only one target was revealed, of the same colour and shape as the start-
282 point. (d) The block structure of the experiments. Participants completed a practice and baselining block, where
283 the obstacle was always central to the screen, before completing a shuffled order of the Rightwards block
284 (obstacle moves from the left to the right between trials), Leftwards block (the obstacle moves from the right
285 to the left between trials), and Random block (obstacle moves randomly between trials).

286

A Computational Model of Action Selection

288   Figure 2 shows a computational model where the selection of an action can be influenced by

289 successful completion of a previous action (because the estimate of the probability of success is

290 dynamically updated on a trial-by-trial basis).

291
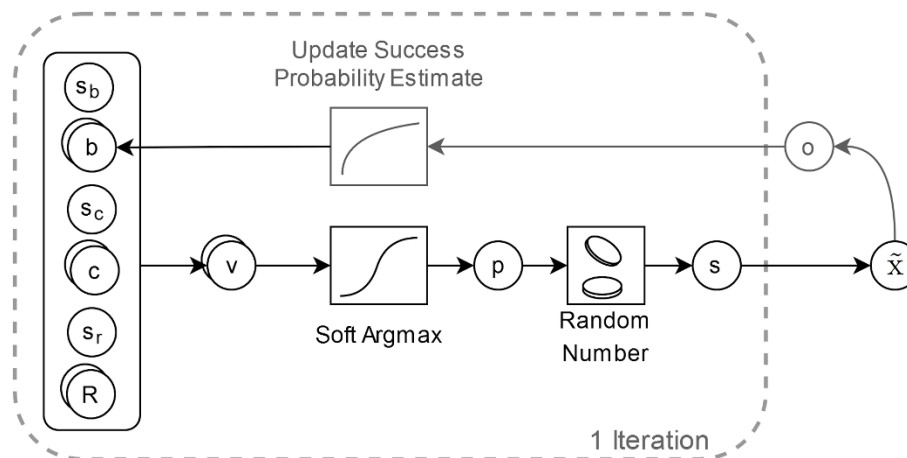
292 Figure 2. A probabilistic choice model for action selection. In a trial, the 'value' for the two actions (going left or
293 right around the obstacle) is calculated from the current costs, rewards and biases built up over previous trials.
294 The values are input to the soft argmax function which gives the probability of selecting the left action. A
295 random number is uniformly sampled and if it is below the probability of the left action then left is selected,
296 otherwise right is selected. The outcome of executing the associated action is observed and the selection and
297 outcome are used to update the biases for each action according to a reinforcement learning rule.

298

299   The model weights the probability of action selection (going left or right around the obstacle)

300 by the 'value' of each action, where we define value as a combination of the costs, rewards, and a

301 bias term reflecting the increased probability of success from repeating the previous action. The

302 model comprises four free parameters – a scaling parameter for each of the cost, bias, and reward

303 terms, s_c, s_b, s_r, respectively, which are used to bring the terms to the same scale so the relative

304   importance of each can be compared, and the rate at which biases are accumulated, r, which has a

305   value in the range $[0, 1]$. Formally, the 'value' of the action i is defined as:

$$V_i = \frac{s_c}{c_i} + s_b \times b_i + s_r \times R_i \tag{1}$$

306   where c is the expected cost, R is the reward and b is the action bias accumulated over the previous

307   trials. R is normalised about the minimum reward, so a total reward of 1 vs 3 becomes a normalised

308   reward of 0 vs 2, so that the effect of an additional reward can be isolated. In this formulation, the

309   expected cost of an action is evaluated perfectly but the model could be modified to include a

310   distribution of the possible expected costs for an action. The reciprocal of the costs was used as a

311   high cost should give a lower valuation of the action, whereas a high reward should give a higher

312   valuation (thus, no transformation was used). The value was converted to selection probabilities

313   using the soft argmax function. A random number was sampled from a uniform distribution

314   between 0 to 1, and the left action selected if the random number was below the output probability

315   and vice versa.

316   Action selection and outcome were used to update the biases associated with each action,

317   according to the following reinforcement learning formula:

$$b_{t+1,i} = \begin{cases} b_{t,i} + r(1 - b_{t,i}) & \text{if } i = s \\ b_{t,i} + r(0 - b_{t,i}) & \begin{array}{l}\text{if } i \neq s \\ \text{or } outcome_t = \text{fail}\end{array} \end{cases} \tag{2}$$

318   where t is the current trial number, r is the bias rate, and s is the selected action. It is assumed

319   participants start with no bias and thus, biases are set to zero on the first trial.

320   In the experiments, the reward for successfully completing either action was the same, so the

321   reward term in the model was not included. For simplicity, this model assumed the cost function

322   was the path length of the movement trajectory. In fact, there is considerable debate (Todorov,

323   2004) within the sensorimotor research literature over the measure used for optimisation

324   (candidates include path length, movement duration, normalised jerk, end-point variability, torque

325   etc). We note that most of these factors co-vary and emphasise that there is a strong tendency for

326   participants to select the shortest possible movement trajectory in unconstrained task settings

327   (Tresilian, 2012). The observed paths were smooth and roughly symmetrical about the centre (see

328   Figure S1 in Supplementary Materials) so path length was approximated by fitting the shortest

329   parabola capable of connecting the start point to the target whilst passing the obstacle on either

330   side. To aid model convergence, path lengths were divided by the minimum possible path length

331   (140mm).

332   The model was fit to the choice data for the three experiments using Bayesian estimation via

333   Stan (Carpenter et al., 2017; version 2.18.2). Each model was fitted with eight chains of 5,000

334   warmup samples and 5,000 iteration samples, giving 40,000 samples per posterior distribution.

335   Convergence was assessed by visually inspecting chain behaviour and confirming the Gelman-

336   Rubin statistic, $\hat{R}$, was below 1.1 (maximum 1.01) for all parameters (Gelman, 2004; Gelman &

337   Rubin, 1992). Posterior distributions for each parameter were summarised using the 95% highest

338   density interval (HDI), the 95% of most credible parameter estimates. The empirical priors used for

339   model fitting are shown below. The scale of the model parameters was assessed by adjusting them

340   until data containing hysteresis was observed. Aside from informing parameter scaling, the priors

341   were then selected to be uninformative. Note that increasing the width of the priors doesn't affect

342   the results of the modelling.

343
$$s_c, s_b \sim N(0, 25)$$
$$r \sim N(0.5, 0.2)$$

344   To check the model fit, datasets were simulated using each experimental list of obstacle

345   positions from the real data. For each experiment, 10,000 samples of the posterior distribution

346   were drawn, and each combination was used to simulate a new data set. For each combination of

347   experiment and obstacle position, the probability of committing an error, extracted from the

348   collected data, was used to simulate error trials. If a sampled number from a uniform distribution

349   between 0 and 1 was below the error rate for the current experiment and obstacle position, the

350   trial was classed as a failed trial and no action was selected. Each new dataset was summarised

351   using a logistic fit for each of the three conditions, and at each obstacle position the minimum and

352   maximum predicted probability of going right around the obstacle from all the logistic fits was

353   taken, representing the credible range of possible data given the posterior distribution. The real

354   data were then overlaid with the credible range to visualise the model fit.

355   **Data Analysis**

356   The output from Unity included the experimental condition; the cursor position, sampled at

357   100Hz and output in millimetres; the timestamp when Unity's physics engine detected participants

358   had left the start-point; the position of the obstacle, output on the scale -1 to 1, where -1 indicated

359   the obstacle touched the left wall and 1 indicated the obstacle touched the right wall; the direction

360   participants moved around the obstacle, detected by Unity's physics engine when participants

361   moved past the leading edge of the obstacle; and the outcome of each trial. The extreme obstacle

362   positions used were -0.9 and 0.9 (-34.2mm and 34.2mm respectively), which ensured participants

363   could only go around the obstacle in one direction at these positions. The average error rate across

364   the three experiments was 14.8%, 95% CI = [14.3, 15.4].

365   Stylus position data were filtered using a dual-pass Butterworth second-order filter with a cut-

366   off frequency of 10Hz. To detect movement onset, the time where movement speed rose above

367   50mm/s closest to the Unity's timestamp of the participant leaving the start-point was classed as

368   movement start. For Experiment 1 and Experiment 3, reaction time (RT) was calculated as the

369   difference in time between the obstacle being shown and movement start, whereas for Experiment

370   2 it was the difference between the red box disappearing and movement starting.

371 RT data were pre-processed by removing trials where no RT was present (117 trials, 0.7%),

372 where RTs were lower than 100ms (36 trials, 0.2%, to account for participants anticipating stimuli

373 presentation), then grouping within participant and condition and removing trials outside 2SD of

374 the mean (826 trials, 4.9%), and then grouping by condition and experiment and removing

375 participant's conditions outside 2SD of the mean (240 trials, 1.4%). This RT data cleaning process

376 was necessary to reduce heteroscedasticity and ensure normal residuals from models.

377 This process removed one participant (Experiment 1, mean RT = 608ms), six participant

378 conditions (Experiment 1, mean RT = 587ms), and 1,219 trials in total (7.3% of observations) from

379 the RT analysis. The remaining trials had a mean RT of 403ms. Of the trials removed, 7.2% were

380 trials 1 and 2 in the experiment, likely because participants had only seen the obstacle presented

381 central to the screen up to that point. Analysis performed on RT data was done on the inverse of RT

382 to increase normality, and back-transformed when reporting. Choice data was pre-processed by

383 removing trials where no movement past the obstacle was detected, removing 283 trials (1.7% of

384 observations).

385 Analysis of the choice and RT data was performed using mixed-effect modelling, utilising the

386 `lme4` package in R (Bates et al., 2015; version 1.1-21). Following Barr et al. (2013), when the

387 maximal random structure did not converge, the optimal random-effects structure was identified

388 using forward model selection, with each mixed-effect model having a random intercept for

389 participant. The effect of each variable was found using likelihood ratio testing, using the `afex`

390 package (Singmann et al., 2019; version 0.23.0). Post-hoc comparisons were performed using the

391 `multcomp` package (Hothorn et al., 2008; version 1.4.8), and corrections for multiple comparisons

392 were made using the Bonferroni-Holm method. The `MuMIn` package (Barton, 2020; version 1.43.6)

393 was used to report marginal R^2 (variance explained by fixed effects), R_m^2, and conditional R^2

394   (variance explained by fixed and random effects), R_c^2, for the models (Nakagawa et al., 2017). The

395   95% confidence intervals for values are reported in square brackets throughout.

396   To examine changes in action selection, a mixed-effect logistic regression was performed. The

397   fixed effects were the obstacle's position, the condition, the experiment, and all combinations of the

398   interactions between these variables. The model had a random intercept for each participant. While

399   a model with a random slope for obstacle position converged, the single repetition of each condition

400   led to an artificially steep main effect slope of obstacle position, so was not included in the model.

401   The default condition was Random, the default experiment was Experiment 1, and the default

402   obstacle position was 0. To understand whether hysteresis changed with experiment, the log-odds

403   (LO) of going right around the obstacle at the central obstacle position was compared between

404   conditions. Hysteresis was quantified as the increased LO of going right at the central obstacle

405   position in the Rightwards condition compared to the Leftwards condition. We compared this

406   across experiments.

407   To investigate changes in RT, a mixed-effect linear regression was performed. The fixed effects

408   were the trial number in the block, the condition, the experiment, and all combinations of the

409   interactions between these variables. The model had a random intercept for participant, with

410   random slopes of condition. The default condition was Random and the default experiment was

411   Experiment 1. The trial number in block was centred about the middle trial. To understand how

412   RTs were affected, the estimated marginal mean (EMM) RT at the central obstacle position was

413   compared between conditions. The difference in RT between the sequential conditions and Random

414   was then compared between experiments.

415   The large sample size in Experiment 1 presented opportunity for more detailed analysis of

416   hysteresis. We expected that, while there would be no global hysteresis in the Random condition,

417   participants might exhibit biases within this block on a trial-by-trial basis. To explore this, a mixed

418   effect logistic regression was performed to understand how much the previous trial biased the

419   current trial inside the Random block. The fixed effects were the obstacle's position, the prime

420   condition – whether the participant went left (Left Previous) or right (Right Previous) on the

421   previous trial, and the interaction between the two. The model had a random intercept for

422   participant. As with the analysis of choice with condition, a random slope of obstacle position was

423   omitted to avoid artificially inflating the main effect slope of obstacle position. The default prime

424   condition was Left Previous. The default obstacle position was 0. To understand how selection was

425   influenced by the previous trial, the LO of going right at the central obstacle position was compared

426   between prime conditions.

427   As well as choices being biased by the previous trial, we found RTs were shorter when

428   participants repeated their previous action when compared to switching action. To investigate the

429   relationship between RT and hysteresis, a mixed-effect linear regression was performed on the

430   Random condition, splitting the data by whether the participant switched or repeated the previous

431   trial's direction. The model had fixed effects of trial number, switch condition, and the interaction of

432   the two, a random intercept for participant and a random slope of switch condition per participant.

433   The default switch condition was repeated. The trial number in block was centred about the middle

434   trial. The estimated mean RT at the central obstacle position was compared between switch

435   conditions.

436   All statistical analyses and data processing were performed using custom-written scripts in R

437   (R Core Team, 2018; version 3.5.2). Upon publication, all analyses code and model fits will be

438   available through https://github.com/immersivecognition, and the complete dataset will be made

439   available through the University of Leeds Data Repository.

440 # Results

441 **Choice Analysis**

442   We first examined whether our group-level manipulations resulted in action selection

443 biases. Hysteresis would result in participants going right around the obstacle more often in the

444 Rightwards condition (where the obstacle moved from the left of the screen to the right between

445 trials) and going left around the obstacle more often in the Leftwards condition (where the obstacle

446 moved from the right of the screen to the left between trials), whereas the Random block (where

447 the obstacle moved randomly between trials) should show no overall bias. We predicted that the

448 degree of this bias would diminish when participants were provided with more planning time

449 (Experiment 2) and when action execution was performed under a reduced cognitive task load

450 (Experiment 3).

451   The experimental task was successful in revealing hysteresis (Figure 3a), with the effect

452 was diminished in Experiment 2 (Figure 3b) and Experiment 3 (Figure 3c).
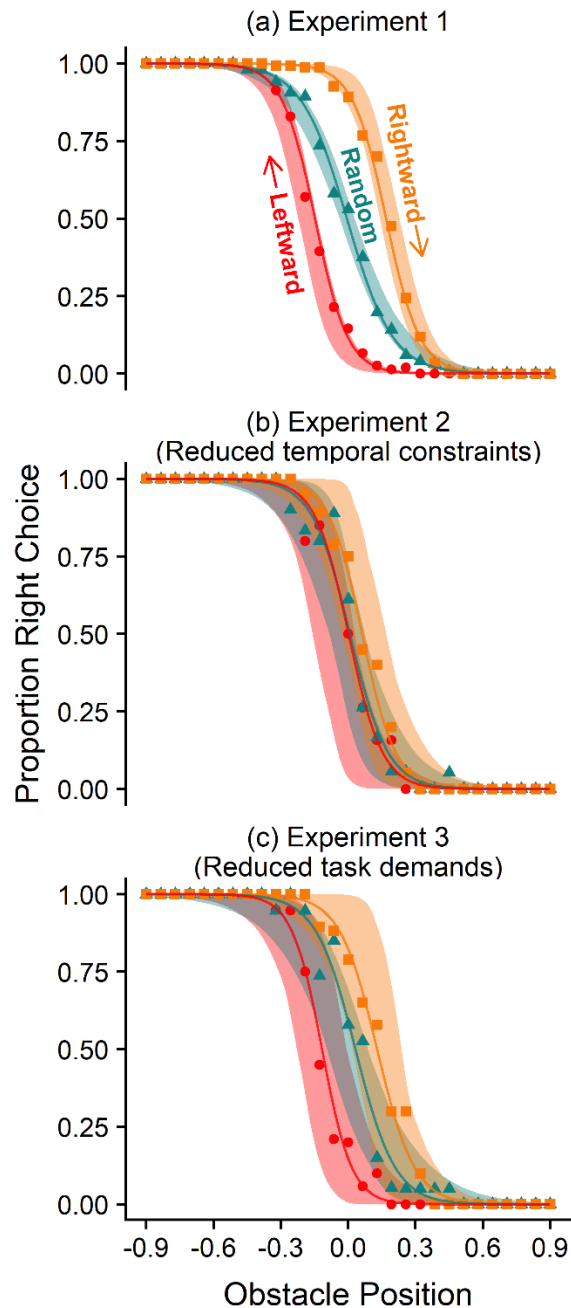
Figure 3. Comparison of experimental and simulated data for choices between experiments. The points and solid lines represent experimental data, and the ribbons represent simulated data. The points indicate mean proportion of participants who passed the obstacle on the right for each obstacle position. The solid lines represent the fit of a logistic regression for the experimental condition. Data was simulated using the decision-making model, with 10,000 samples of the posterior distribution used to simulate choices, and each new data set summarised with a logistic regression. The ribbon represents the minimum and maximum predicted probability of going right from the regressions of the simulated data. The conditions are Rightwards (where the obstacle moves from the left of the screen to right between trials), Random (where the obstacle moves randomly between trials), and Leftwards (where the obstacle moves from the right of the screen to left between trials).

464   We performed a mixed-effect logistic regression to predict the direction participants chose

465   on a given trial. The model ($\chi^2(17)$ = 18,206.60, p < 0.001, R^2_m = 0.94, R^2_c = 0.95) revealed a

466   significant main effect of position ($\chi^2(1)$ =4,627.52, p < 0.001) and condition ($\chi^2(2)$ = 1,330.53, p <

467   0.001), but no significant effect of experiment ($\chi^2(2)$ = 2.61, p = 0.272). There were significant

468   interactions between position and condition ($\chi^2(2)$ = 54.01, p < 0.001), and condition and

469   experiment ($\chi^2(4)$ = 94.53, p < 0.001), but no significant interaction between position and

470   experiment ($\chi^2(2)$ = 4.60, p = 0.100), or between position, condition and experiment ($\chi^2(4)$ = 4.36, p

471   = 0.359).

472   Bonferroni-Holm corrected comparisons were performed to see how the log-odds of passing

473   the obstacle on the right changed with condition and experiment at the central obstacle position. In

474   **Experiment 1**, participants were significantly more likely to go right in Rightwards compared to

475   Random (LO = 2.52 [2.18, 2.86], p < 0.001), and significantly less likely to go right in Leftwards

476   compared to Random (LO = -2.18 [-2.51, -1.85], p < 0.001), indicating participants were more likely

477   to continue using the previous direction in the sequential conditions. Further, participants were

478   more likely to go right in Rightwards compared to Leftwards (LO = 4.70 [4.28, 5.12], p < 0.001).

479   This comparison gives the increased log odds of passing the obstacle on the right at the central

480   obstacle positions between the sequential conditions, and is the measure of hysteresis used

481   throughout.

482   In **Experiment 2**, where participants were forced to wait in the start-point for 1.5 seconds

483   while the obstacle was shown before being allowed to move, they were significantly more likely to

484   go right in Rightwards compared to Random (LO = 0.82 [0.05, 1.59], p = 0.026), but not in

485   Leftwards compared to Random (LO = -0.07 [-0.80, 0.65], p = 0.811). Furthermore, participants

486   were more likely to go right in the Rightwards trials compared to Leftwards (LO = 0.89 [0.11, 1.68],

487   p = 0.023), indicating hysteresis was present, but to a lesser extent compared to Experiment 1.

488   In **Experiment 3**, where participants were presented with a reduced cognitive task load, they

489   were significantly more likely to go right in Rightwards compared to Random (LO = 1.24 [0.44,

490   2.03], p < 0.001), and significantly less likely to go right in Leftwards compared to Random (LO = -

491   2.30 [-3.21, -1.39], p < 0.001). Participants were more also likely to go right in Rightwards

492   compared to Leftwards (LO = 3.54 [2.54, 4.54], p < 0.001) indicating the presence of some

493   hysteresis but attenuated relative to Experiment 1.

494   To investigate how the magnitude of hysteresis changed between the experiments, the

495   increased log odds of going right in the Rightwards condition compared to the Leftwards condition

496   were compared across experiments. Participants in Experiment 2 showed significantly less

497   hysteresis than participants in Experiment 1 (LO = -3.81 [-4.70, -2.93], p < 0.001), and participants

498   in Experiment 3 also showed significantly less hysteresis than participants in Experiment 1 (LO = -

499   1.16 [-0.09, -2.24], p = 0.012).

500   These results indicate that the experiment interventions designed to: (i) decrease temporal

501   constraints and (ii) decrease cognitive task load reduced the magnitude of hysteresis in

502   Experiments 2 and 3 respectively (relative to Experiment 1). We note that the impact seems to be

503   more pronounced for planning time increase (Experiment 2) relative to task load reduction

504   (Experiment 3).

505   Simulations from the action selection model were fit with the experimental data. The 95% HDI

506   of the posterior distributions for each parameter are summarised in Table 1 (the full posterior

507   distributions for each parameter are visualised in Figure S2 in Supplementary Materials). To

508   understand whether the bias changed between experiments, the posterior distribution for the bias

509   scaler parameter in Experiments 2 and 3 were subtracted from that of Experiment 1. This showed

510   that the bias scaler estimate in Experiment 2 was lower than in Experiment 1 (mean difference = -

511   1.60, 95% HDI = [-1.18, -2.02]), and the estimate in Experiment 3 was also lower than in

512   Experiment 1 (mean difference = -0.82, 95% HDI = [-0.38, -1.25]), indicating hysteresis was

513   attenuated by reducing the contribution of biases, built up through previous successes and failures,

514   towards action selection under reduced temporal constraints and cognitive load.

515

Table 1

*Mean and 95% highest density interval (HDI) estimates of the decision-making model parameters.*

Experiment	Mean [95% HDI]		
	Cost scaler	Bias Scaler	Bias Rate
Experiment 1	39.26 [37.36, 41.24]	2.69 [2.53, 2.86]	0.32 [0.28, 0.36]
Experiment 2 (Reduced temporal constraints)	41.96 [36.59, 47.84]	1.09 [0.71, 1.48]	0.43 [0.27, 0.62]
Experiment 3 (Reduced task demands)	37.63 [32.94, 42.82]	1.87 [1.48, 2.29]	0.35 [0.21, 0.50]

516

517   The posterior distributions were then used to simulate new data so that predictions from the

518   model could be compared to the experimental data. Ten thousand samples of the posterior

519   distribution were taken, and each used to simulate new responses to the experiments. Each new

520   dataset was summarised with a logistic regression for each condition, and the upper and lower limit

521   for these predictions used to visualise the model's predictions. These predictions are shown as

522   coloured ribbons in Figure 3 per experiment. Note that the observed selection probabilities lie

523   within the range of the model's predictions, with distinct separations between the two sequential

524   conditions for Experiments 1 and 3 around the central obstacle positions, but for Experiment 2 the

525   two sequential conditions share considerable overlap, consistent with the experimental data.

526   The results thus far indicate participants are biased towards repeating previously used action

527   structures when the obstacle moves between trials in a sequential manner. While most hysteresis

528   studies employ similar sequential trial designs, some have also found hysteresis when stimuli are

529   varied randomly across trials (Jax & Rosenbaum, 2007; Valyear et al., 2018). Thus, we explored

530   whether the participant's selection on a current trial was biased by the direction they passed the

531   obstacle on the previous trial in the Random block. Analysis of the data for repeated and switched

532   trials from the Random block in Experiment 1 provided support for this idea (Figure 4).
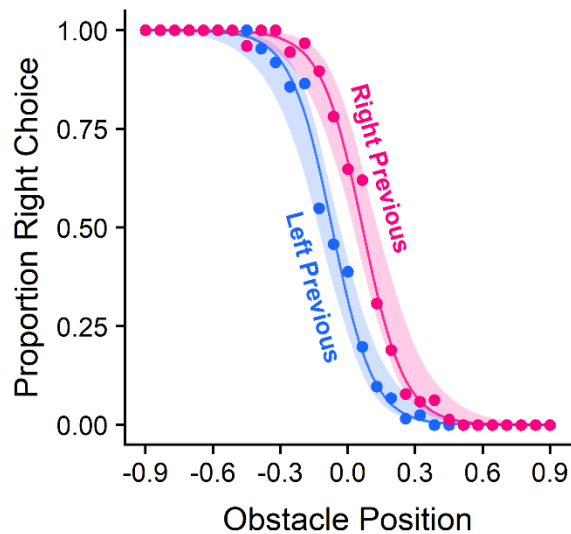
533

534   Figure 4. Empirical and simulated data for choices in Experiment 1 Random condition. The points and solid
535   lines represent experimental data, and the ribbons represent simulated data. The points indicate mean
536   proportion of participants who passed the obstacle on the left for each obstacle position. The solid lines
537   represent the fit of a logistic regression for the experimental condition. Data were simulated using the decision-
538   making model, with 10,000 samples of the posterior distribution used to simulate choices, and each new data
539   set summarised with a logistic regression. The ribbon represents the minimum and maximum predicted
540   probability of going right from the regressions of the simulated data. The conditions are Left Previous (where
541   the participant passed the obstacle on the left on the previous trial), and Right Previous (where participants
542   passed the obstacle on the right on the previous trial).

543

544   A mixed-effect logistic regression was performed on data from the Random block from

545   Experiment 1 to predict the direction participants went on the current trial. The model ($\chi^2(3)$ =

546   4,406.80, p < 0.001, R_m^2 = 0.91, R_c^2 = 0.93) showed a significant main effect of position ($\chi^2(1)$ =

547   2,215.09, p < 0.001) and prime condition ($\chi^2(1)$ = 96.62, p < 0.001), but there was no significant

548   interaction between position and prime condition ($\chi^2(1)$ = 0.16, p = 0.688). A comparison was

549   performed to see how the log-odds of passing the obstacle on the right changed with prime

550   condition at the central obstacle position. Participants were significantly more likely to go right in

551   the Right Previous condition compared to the Left Previous condition (LO = 1.51 [1.19, 1.83], p <

552   0.001). The increased log odds of going right were smaller for the two prime conditions when

553   compared to the two sequential conditions from the analysis reported above - indicating biases

554   accumulate over longer action sequences than just the previous trial.

555   To understand whether similar trial-to-trial biases emerged in the decision-making model, the

556   parameter posterior distributions (10,000 samples) from Experiment 1 were used to simulate new

557   datasets. The responses to the Random condition were then extracted, and each new dataset was

558   summarised using a logistic regression per prime condition. The minimum and maximum predicted

559   probability of going right from these regressions was used to visualise the model's predictions and

560   are represented as ribbons in Figure 4. Consistent with the observed experiment data, the model

561   shows a distinct separation between the two prime conditions, with Right Previous cases being

562   more likely to go right at the central obstacle positions.

563   **Reaction Times Analysis**

564   Participants in Experiment 1 showed a reduction in RT in the sequential conditions compared

565   to the Random condition (Figure 5a). Collapsing across all trials, the mean RT in the Random

566   condition was 417ms [414, 419], compared to 387ms [385, 389] in the Rightwards condition and

567   388ms [386, 390] in the Leftwards condition. In Experiment 2, RTs were lower than in Experiment

568   1 and there seemed to be no RT benefits in the sequential conditions (Figure 5b), with a mean RT in

569   the Random condition of 368ms [363, 373] compared to 381ms [375, 387] in the Rightwards

570   condition and 380ms [374, 386] in the Leftwards condition. In Experiment 3, RTs were again lower

571   than in Experiment 1 with a large RT benefit in the sequential conditions (Figure 5c). The mean RT

572   in the Random condition was 379ms [374, 385] compared to 331ms [327, 336] in the Rightwards

573   condition and 342ms [337, 347] in the Leftwards condition.

575   Figure 5. Comparison of RTs for each condition between experiments. The open circles show the mean RTs
576   for each participant. The solid circles show the mean RT for each combination of condition and experiment
577   across all participants, and the error bars show the 95% confidence intervals around the estimate of the
578   mean. The conditions are Rightwards (where the obstacle moves from the left of the screen to the right
579   between trials), Random (where the obstacle moves randomly between trials), and Leftwards (where the
580   obstacle moves from the right of the screen to the left between trials).

581

582   To understand how RTs were affected by biases, a linear mixed-effect model was conducted to

583   predict RTs on a given trial. The model [$\chi^2(22) = 2,337.69$, p < 0.001, $R_m^2 = 0.10$, $R_c^2 = 0.49$] showed a

584   significant main effect of trial ($\chi^2(1) = 7.53$, p = 0.006), condition ($\chi^2(2) = 104.22$, p < 0.001), and

585   experiment ($\chi^2(2) = 30.30$, p < 0.001). There were significant interactions of trial and condition

586   ($\chi^2(2) = 28.45$, p < 0.001), and condition and experiment ($\chi^2(4) = 50.58$, p < 0.001), but no

587   significant interaction between trial and experiment ($\chi^2(2) = 2.47$, p = 0.291). There was a

588   significant interaction between trial, condition and experiment ($\chi^2(4) = 19.72$, p < 0.001).

589   Bonferroni-Holm corrected comparisons were performed to see how RTs changed with

590   condition and experiment at the middle trial in the block, where the benefits of hysteresis would be

591   expected. In **Experiment 1**, the Random condition (EMM = 417ms [410, 425]) was significantly

592   slower than the Leftwards (EMM = 389ms [381, 397], p < 0.001) and Rightwards conditions (EMM

593   = 387ms [380, 396], p < 0.001), but there was no significant difference between the Leftwards and

594   Rightwards conditions (p = 0.659). Participants were ~30ms faster in the sequential conditions

595   than in the Random condition, indicating RT savings from choice perseveration.

596   In **Experiment 2**, there were no significant differences between the Random (EMM = 368ms

597   [352, 385]) and Leftwards condition (EMM = 380ms [361, 402], p = 0.151), the Random and

598   Rightwards condition (EMM = 382ms [361, 404], p = 0.151), or between the Leftwards and

599   Rightwards condition (p = 0.877).

600   In **Experiment 3**, the Random condition (EMM = 380ms [363, 398]) was significantly slower

601   than the Leftwards (EMM = 342ms [326, 359, p < 0.001) and Rightwards conditions (EMM = 331ms

602   [316, 348], p < 0.001), but there was no significant difference between the Leftwards and

603   Rightwards conditions (p = 0.093). Participants were 40-50ms quicker in the sequential conditions

604   than in the Random condition, showing savings that were marginally larger than those observed in

605   Experiment 1.

606   To investigate whether the magnitude of RT reduction changed between experiments, we

607   compared the difference in RT of the Random condition to the sequential conditions between

608   experiments. The difference in RT between the Random and Leftwards conditions was significantly

609   lower in Experiment 2 compared to Experiment 1 (p < 0.001), but significantly higher in

610   Experiment 3 compared to Experiment 1 (p = 0.005). The difference in RT between the Random

611   and Leftwards conditions was significantly lower in Experiment 2 compared to Experiment 1 (p <

612   0.001) but there was no difference between Experiment 3 and 1 (p = 0.074).

613   In the Random block of Experiment 1 participants showed a reduction in RT for repeated

614   choices compared to switched choices (Figure 6). Collapsing across all trials and participants, the

615   mean RT in the Repeated condition was 407ms [404, 410] compared to 430ms [427, 433] in the

616   Switched condition.
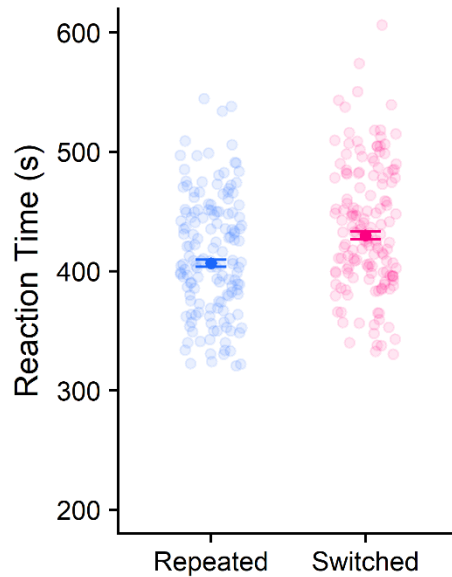
617

618   Figure 6. Comparison of RTs inside Experiment 1's Random condition. The open circles show the mean RTs
619   for each participant. The solid circles show the mean RT for each switch condition across all participants, and
620   the error bars show the 95% confidence intervals around the estimate of the mean. The conditions are
621   Repeated (where participants made the same choice on the current trial as on the last trial), and Switched
622   (where participants switched choice from the previous trial).

623

624   To understand how RTs changed with switch condition on a given trial, a linear mixed-effect

625   model was performed. The model [$\chi^2(5)$ = 197.71, p < 0.001, R_m^2 = 0.03, R_c^2 = 0.48] showed a

626   significant main effect of trial ($\chi^2(1)$ = 14.23, p < 0.001) and switch condition ($\chi^2(1)$ = 97.79, p <

627   0.001), but no interaction between trial and switch condition ($\chi^2(1)$ = 0.19, p = 0.663). A

628   comparison was performed to see how RTs changed with switch condition at the middle trial in the

629   block. The repeated condition (EMM = 407ms [400, 415]) was significantly faster than the switched

630   condition (EMM = 429ms [421, 438], p < 0.001), indicating participants had an RT benefit from

631   repeating only a single previous choice. The difference between the Switched and Repeated

632   conditions was lower than the difference between the Random and the sequential conditions from

633   the earlier analysis, indicating RT savings may be a cumulative process, building up across trials.

634 **Discussion**

635 Our goal was to examine the bias shown by skilled adult humans towards selecting a

636 previously selected action structure when alternative options that would be selected in *de novo*

637 conditions were available. To this end, we created a simple obstacle avoidance aiming task. We

638 found participants exhibited hysteresis effects when the obstacle moved systematically in one

639 direction between trials. In blocks where the obstacle moved randomly, there was no global

640 hysteresis effect. The random blocks did, however, show trial-by-trial biases - with action selection

641 being influenced by the previous movement. We were also interested in exploring the impact on

642 hysteresis of changing the temporal constraints of the task. The rationale for manipulating the

643 temporal constraints was based on the growing evidence that human decision-making involves

644 evidence accumulation processes. The existence of such processes suggests that humans may

645 choose to act before a full evaluation of the costs associated with the available action options has

646 been completed (i.e. as soon as an action reaches a threshold it is selected). It follows that providing

647 a longer time period for decision-making might cause a different action to be selected (as the

648 available time can be used to more fully evaluate the costs, reducing the reliance of the current

649 decision on previous successes and failures). We used two manipulations to alter the temporal

650 constraints of the task. In Experiment 2, we directly manipulated the temporal constraints by

651 preventing action until a 1500ms time window had elapsed. In Experiment 3, we indirectly altered

652 the constraints by decreasing the cognitive demands of the task (reasoning that less time spent

653 identifying the task goal would provide more time for evaluating the respective costs of the

654 available actions). The results showed that the hysteresis effect was practically eliminated in

655 Experiment 2 and attenuated in Experiment 3.

656 Our investigation of hysteresis was motivated by our hypothesis that hysteresis is the

657 naturally emergent property of a dynamical learning system that is operating in an uncertain world.

658 In order to test our hypothesis we created a simple POMDP type computational model that

659   incorporated dynamic probabilistic estimate updating. We used this model to simulate behavioural

660   responses for the experimental tasks and found that it was able to capture the empirical data. This

661   finding suggests that there is no need to invoke the existence of a bespoke 'hysteresis function'

662   within the sensorimotor system, and provides support for our hypothesis that hysteresis is an

663   emergent property of a dynamical learning system.

664   The results reported within this manuscript emphasise the dynamical nature of human

665   sensorimotor decision making. The challenge for the human nervous system is to maintain optimal

666   action selection in a noisy and uncertain world. The only way that the nervous system can maintain

667   its efficiency is through an ongoing evaluation of the accuracy of its internal representation of the

668   external world, and frequent updating of its probability estimates. The current findings suggest that

669   this updating occurs on a trial-by-trial basis (though the biases we observed also accumulated

670   across multiple trials). This paints a picture of a system that is continually adapting, and ensuring

671   that its actions are precisely tailored to the external environment. This observation calls into

672   question the classical distinctions between sensorimotor control and sensorimotor learning. It

673   appears that human control systems appear stable because they have been refined over long

674   periods of time through interactions with a world that obeys consistent rules described by

675   Newtonian mechanics– but controllers are nevertheless updated continually on the basis of

676   feedback from every interaction. Our findings also highlight the dynamics of decision-making in

677   terms of the system needing to make choices under time constraints (where there are strong

678   evolutionary pressures favouring species who react swiftly). In line with the existence of evidence

679   accumulation processes, we found that the hysteresis effect was attenuated when the temporal

680   constraints of the task were eased. This emphasises the tendency within the system to select the

681   first action to cross a pre-specified threshold, favouring recently successful actions, rather than wait

682   until the full costs of all options have been exhaustively evaluated. The success of such a strategy is

683   witnessed by the fact that Homo sapiens remained standing after the evolutionary arms race of the

684   past gigayear.

685   Our experiments have focussed on the sensorimotor system but the general phenomenon of

686   hysteresis can be observed in other aspects of human behaviour. For example, hysteresis effects

687   have been observed in perceptual decision-making tasks where choices are biased by previous

688   decisions (Abrahamyan et al., 2016; Akaishi et al., 2014; Urai et al., 2019). The magnitude of the

689   perceptual bias tends to depend on whether the previous decision was rewarded or not

690   (Abrahamyan et al., 2016; Hermoso-Mendizabal et al., 2018). These hysteresis effects (typically

691   described as choice-history biases) have been successfully implemented in evidence accumulation

692   models of decision-making. We argue that the existence of hysteresis within the perceptual system

693   can be explained through the same mechanisms that we used to account for hysteresis in

694   sensorimotor decision-making (i.e. the presence of Bayesian type processes operating within the

695   brain, where priors are continually updated with new sensory information to create posterior

696   probability estimates). It is possible that similar mechanisms can account for reports of hysteresis

697   in higher order cognition (often described as 'perseveration'). There is a growing consensus that

698   the sensorimotor system provides the phylogenetic and ontogenetic foundations of higher order

699   cognition (Raw et al., 2019; Wilson, 2002). The postulated links between the sensorimotor and

700   cognitive system might suggest a close relationship between sensorimotor hysteresis and

701   perseveration type behaviours. This may prove a fruitful line of investigation for future studies.

702   Previous accounts of hysteresis have assumed that the sensorimotor system has a 'hysteresis'

703   function whose purpose is to create an advantage when planning a new movement. It is argued that

704   modifying a previously used plan would be more cognitively efficient than planning a new one from

705   scratch, so hysteresis exists to improve planning efficiency (Meulenbroek et al., 1993; Rosenbaum

706   et al., 2007; Schütz & Schack, 2019; Weiss & Wark, 2009), as indexed by reduced RTs when

707   performing the same action as previously (Valyear et al., 2018). On the basis of previous reports we

708   fully expected to find reduced RTs when the hysteresis effect was present. Moreover, we

709   anticipated the presence of reduced RTs on the theoretical basis that evidence accumulation

710   processes will cause actions to be selected more rapidly when there is a bias towards one action

711   versus another. In line with these expectations, we observed a decrease in the average RT on the

712   sequential trials in Experiment 1 relative to the random trials. In Experiment 2, participants were

713   given a substantial time to select the goal directed action and the RTs were similar to the sequential

714   trials in Experiment 1 regardless of trial type. In Experiment 3, the task demands were decreased

715   and there was a commensurate reduction in RT (consistent with a large body of literature showing

716   that RT is a function of task complexity). Notably, the sequential trials within Experiment 3 showed

717   hysteresis (relative to the random trials within the experiment) and were associated with faster

718   RTs than the random trials (producing the fastest RTs across all three experiments as predicted by

719   the presence of hysteresis and the reduced task complexity).

720   The work presented within this manuscript addresses issues from the field of 'Human-Like-

721   Computing' where researchers attempt to bridge the gap between models of human decision-

722   making and the models used in artificial intelligence and robot motion control. Stochastic models of

723   actions, observations, costs and rewards are the main tools used in modelling and planning robot

724   motion, including tasks that involve reaching behind obstacles (Dogar & Srinivasa, 2012). An

725   improved understanding of human decision-making can inform the development of such robot

726   motion models. The identification of the hysteresis bias allows roboticists and computer scientists

727   to decide whether their agents are operating within environments that are sufficiently constrained

728   so that control schemes can seek to ameliorate hysteresis. Alternatively, hysteresis may suggest

729   mechanisms through which a robotic agent can show human-like flexibility and adaptability in

730   complex dynamic environments. Moreover, the identification of hysteresis as an emergent property

731   can help improve the legibility and predictability available within human-robot interactions. It

732   follows that investigations into human biases (such as hysteresis), and their formal description

733   through mathematical models can be useful in robot motion and control. Thus, the approach

734   adopted within this manuscript provides an interesting avenue for future investigations by

735   roboticists, psychologists and computer scientists.

References

736

737 Abrahamyan, A., Silva, L. L., Dakin, S. C., Carandini, M., & Gardner, J. L. (2016). Adaptable history biases

738   in human perceptual decisions. *Proceedings of the National Academy of Sciences*, *113*(25), E3548–

739   E3557. https://doi.org/10.1073/pnas.1518786113

740 Akaishi, R., Umeda, K., Nagase, A., & Sakai, K. (2014). Autonomous Mechanism of Internal Choice

741   Estimate   Underlies   Decision   Inertia.   *Neuron*,   *81*(1),   195–206.

742   https://doi.org/10.1016/j.neuron.2013.10.018

743 Barr, D. J., Levy, R., Scheepers, C., & Tily, H. J. (2013). Random effects structure for confirmatory

744   hypothesis testing: Keep it maximal. *Journal of Memory and Language*, *68*(3), 255–278.

745   https://doi.org/10.1016/j.jml.2012.11.001

746 Barton, K. (2020). *MuMIn: Multi-model inference* (Version 1.43.6) [Computer software]. https://CRAN.R-

747   project.org/package=MuMIn

748 Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4.

749   *Journal of Statistical Software*, *67*(1), 1–48. https://doi.org/10.18637/jss.v067.i01

750 Bogacz, R., Brown, E., Moehlis, J., Holmes, P., & Cohen, J. D. (2006). The physics of optimal decision

751   making: A formal analysis of models of performance in two-alternative forced-choice tasks.

752   *Psychological Review*, *113*(4), 700–765. https://doi.org/10.1037/0033-295X.113.4.700

753 Brody, C. D., & Hanks, T. D. (2016). Neural underpinnings of the evidence accumulator. *Current Opinion*

754   *in Neurobiology*, *37*, 149–157. https://doi.org/10.1016/j.conb.2016.01.003

755 Brookes, J., Warburton, M., Alghadier, M., Mon-Williams, M., & Mushtaq, F. (2019). Studying human

756   behavior with virtual reality: The Unity Experiment Framework. *Behavior Research Methods*, *52*,

757   455–463. https://doi.org/10.3758/s13428-019-01242-0

758 Carpenter, B., Gelman, A., Hoffman, M. D., Lee, D., Goodrich, B., Betancourt, M., Brubaker, M., Guo, J.,

759   Li, P., & Riddell, A. (2017). *Stan*: A Probabilistic Programming Language. *Journal of Statistical*

760   *Software*, *76*(1), 1–32. https://doi.org/10.18637/jss.v076.i01

761   Cohen, R. G., & Rosenbaum, D. A. (2004). Where grasps are made reveals how grasps are planned:

762   Generation and recall of motor plans. *Experimental Brain Research*, *157*(4), 486–495.

763   https://doi.org/10.1007/s00221-004-1862-9

764   Cohen, R. G., & Rosenbaum, D. A. (2011). Prospective and retrospective effects in human motor control:

765   Planning grasps for object rotation and translation. *Psychological Research*, *75*(4), 341–349.

766   https://doi.org/10.1007/s00426-010-0311-6

767   Dixon, P., McAnsh, S., & Read, L. (2012). Repetition effects in grasping. *Canadian Journal of*

768   *Experimental Psychology/Revue Canadienne de Psychologie Expérimentale*, *66*(1), 1–17.

769   https://doi.org/10.1037/a0026192

770   Dogar, M. R., & Srinivasa, S. S. (2012). A Planning Framework for Non-Prehensile Manipulation under

771   Clutter and Uncertainty. *Autonomous Robots*, *33*(3), 217–236. https://doi.org/10.1007/s10514-012-

772   9306-z

773   Dreher, J.-C., & Tremblay, L. (Eds.). (2009). *Handbook of reward and decision making*. Academic Press.

774   Friston, K. (2010). The free-energy principle: A unified brain theory? *Nature Reviews Neuroscience*, *11*(2),

775   127–138. https://doi.org/10.1038/nrn2787

776   Gelman, A. (2004). *Bayesian data analysis* (Second edition.). Chapman & Hall/CRC.

777   Gelman, A., & Rubin, D. B. (1992). Inference from Iterative Simulation Using Multiple Sequences.

778   *Statistical Science*, *7*(4), 457–472. https://doi.org/10.1214/ss/1177011136

779   Gold, J. I., & Shadlen, M. N. (2007). The Neural Basis of Decision Making. *Annual Review of*

780   *Neuroscience*, *30*(1), 535–574. https://doi.org/10.1146/annurev.neuro.29.051605.113038

781   Hermoso-Mendizabal, A., Hyafil, A., Rueda-Orozco, P. E., Jaramillo, S., Robbe, D., & de la Rocha, J.

782   (2018). *Response outcomes gate the impact of expectations on perceptual decisions* [Preprint].

783   https://doi.org/10.1101/433409

784   Hothorn, T., Bretz, F., & Westfall, P. (2008). Simultaneous inference in general parametric models.

785   *Biometrical Journal*, *50*(3), 346–363. https://doi.org/10.1002/bimj.200810425

786 Jax, S. A., & Rosenbaum, D. A. (2007). Hand path priming in manual obstacle avoidance: Evidence that

787   the dorsal stream does not only control visually guided actions in real time. *Journal of Experimental*

788   *Psychology: Human Perception and Performance*, *33*(2), 425–441. https://doi.org/10.1037/0096-

789   1523.33.2.425

790 Jax, S. A., & Rosenbaum, D. A. (2009). Hand path priming in manual obstacle avoidance: Rapid decay of

791   dorsal stream information. *Neuropsychologia*, *47*(6), 1573–1577.

792   https://doi.org/10.1016/j.neuropsychologia.2008.05.019

793 Kaelbling, L. P., Littman, M. L., & Cassandra, A. R. (1998). Planning and acting in partially observable

794   stochastic domains. *Artificial Intelligence*, *101*(1–2), 99–134. https://doi.org/10.1016/S0004-

795   3702(98)00023-X

796 Kelso, J. A. S., Buchanan, J. J., & Murata, T. (1994). Multifunctionality and switching in the coordination

797   dynamics of reaching and grasping. *Human Movement Science*, *13*(1), 63–94.

798   https://doi.org/10.1016/0167-9457(94)90029-9

799 Kent, S. W., Wilson, A. D., Plumb, M. S., Williams, J. H. G., & Mon-Williams, M. (2009). Immediate

800   Movement History Influences Reach-to-Grasp Action Selection in Children and Adults. *Journal of*

801   *Motor Behavior*, *41*(1), 10–15. https://doi.org/10.1080/00222895.2009.10125921

802 Meulenbroek, R. G. J., Rosenbaum, D. A., Thomassen, A. J. W. M., & Schomaker, L. R. B. (1993). Limb-

803   Segment Selection in Drawing Behaviour. *The Quarterly Journal of Experimental Psychology*

804   *Section A*, *46*(2), 273–299. https://doi.org/10.1080/14640749308401047

805 Mushtaq, F., Wilkie, R. M., Mon-Williams, M. A., & Schaefer, A. (2016). Randomised prior feedback

806   modulates neural signals of outcome monitoring. *NeuroImage*, *125*, 868–879.

807   https://doi.org/10.1016/j.neuroimage.2015.10.046

808 Nakagawa, S., Johnson, P. C. D., & Schielzeth, H. (2017). The coefficient of determination R2 and intra-

809   class correlation coefficient from generalized linear mixed-effects models revisited and expanded.

810   *Journal of the Royal Society Interface*, *14*(134), 1–11. https://doi.org/10.1098/rsif.2017.0213

811   R Core Team. (2018). *R: A Language and Environment for Statistical Computing* (Version 3.5.2)

812   [Computer software]. https://www.R-project.org/

813   Raw, R. K., Wilkie, R. M., Allen, R. J., Warburton, M., Leonetti, M., Williams, J. H. G., & Mon-Williams,

814   M. (2019). Skill acquisition as a function of age, hand and task difficulty: Interactions between

815   cognition and action. *PLoS ONE*, *14*(2). https://doi.org/10.1371/journal.pone.0211706

816   Rosenbaum, D. A., Cohen, R. G., Jax, S. A., Weiss, D. J., & van der Wel, R. (2007). The problem of serial

817   order in behavior: Lashley's legacy. *Human Movement Science*, *26*(4), 525–554.

818   https://doi.org/10.1016/j.humov.2007.04.001

819   Rosenbaum, D. A., & Jorgensen, M. J. (1992). Planning macroscopic aspects of manual control. *Human*

820   *Movement Science*, *11*(1–2), 61–69. https://doi.org/10.1016/0167-9457(92)90050-L

821   Rostoft, M. S., Sigmundsson, H., Whiting, H. T. A., & Ingvaldsen, R. P. (2002). Dynamics of hand

822   preference in 4 year-old children. *Behavioural Brain Research*, *132*(1), 59–68.

823   https://doi.org/10.1016/S0166-4328(01)00415-6

824   Schütz, C., & Schack, T. (2019). A Simple Model to Estimate the Percentage of Motor Plan Reuse From

825   Hysteresis Effect Size. *Frontiers in Psychology*, *10*, 1–9. https://doi.org/10.3389/fpsyg.2019.00561

826   Schütz, C., Weigelt, M., Odekerken, D., Klein-Soetebier, T., & Schack, T. (2011). Motor Control Strategies

827   in a Continuous Task Space. *Motor Control*, *15*(3), 321–341. https://doi.org/10.1123/mcj.15.3.321

828   Schweighofer, N., Xiao, Y., Kim, S., Yoshioka, T., Gordon, J., & Osu, R. (2015). Effort, success, and

829   nonuse determine arm choice. *Journal of Neurophysiology*, *114*(1), 551–559.

830   https://doi.org/10.1152/jn.00593.2014

831   Short, M. W., & Cauraugh, J. H. (1997). Planning macroscopic aspects of manual control: End-state

832   comfort and point-of-change effects. *Acta Psychologica*, *96*(1–2), 133–147.

833   https://doi.org/10.1016/S0001-6918(97)00006-1

834   Singmann, H., Bolker, B., Westfall, J., & Aust, F. (2019). *afex: Analysis of factorial experiments* (Version

835   0.23.0) [Computer software]. https://CRAN.R-project.org/package=afex

836   Todorov, E. (2004). Optimality principles in sensorimotor control. *Nature Neuroscience*, *7*(9), 907–915.

837   https://doi.org/10.1038/nn1309

838   Tresilian, J. (2012). *Sensorimotor Control & Learning: An Introduction to the Behavioural Neuroscience*

839   *of Action.* Palgrave Macmillan.

840   Trommershäuser, J., Maloney, L. T., & Landy, M. S. (2008). Decision making, movement planning and

841   statistical   decision   theory.   *Trends   in   Cognitive   Sciences*,   *12*(8),   291–297.

842   https://doi.org/10.1016/j.tics.2008.04.010

843   Unity Technologies. (2018). *Unity* (Version 2018.1) [Computer software]. https://unity3d.com

844   Urai, A. E., de Gee, J. W., Tsetsos, K., & Donner, T. H. (2019). Choice history biases subsequent evidence

845   accumulation. *ELife*, *8*, 1–34. https://doi.org/10.7554/eLife.46331

846   Valyear, K. F., Fitzpatrick, A. M., & Dundon, N. M. (2018). Now and then: Hand choice is influenced by

847   recent   action   history.   *Psychonomic   Bulletin   &   Review*,   *26*,   305–314.

848   https://doi.org/10.3758/s13423-018-1510-1

849   van der Wel, R. P. R. D., Fleckenstein, R. M., Jax, S. A., & Rosenbaum, D. A. (2007). Hand path priming

850   in manual obstacle avoidance: Evidence for abstract spatiotemporal forms in human motor control.

851   *Journal of Experimental Psychology: Human Perception and Performance*, *33*(5), 1117–1126.

852   https://doi.org/10.1037/0096-1523.33.5.1117

853   Weigelt, M., Rosenbaum, D. A., Huelshorst, S., & Schack, T. (2009). Moving and memorizing: Motor

854   planning modulates the recency effect in serial and free recall. *Acta Psychologica*, *132*(1), 68–79.

855   https://doi.org/10.1016/j.actpsy.2009.06.005

856   Weiss, D. J., & Wark, J. (2009). Hysteresis effects in a motor task with cotton-top tamarins (Sanguinus

857   oedipus). *Journal of Experimental Psychology: Animal Behavior Processes*, *35*(3), 427–433.

858   https://doi.org/10.1037/a0013964

859   Wilson, M. (2002). Six views of embodied cognition. *Psychonomic Bulletin & Review*, *9*(4), 625–636.

860   https://doi.org/10.3758/BF03196322

861