

# Time-dependent competition between goal-directed and habitual response preparation

Robert M. Hardwick<sup>1,2,4\*</sup>, Alexander D. Forrence<sup>1</sup>, John W. Krakauer<sup>1,3</sup> and Adrian M. Haith<sup>1</sup>

**Habits are commonly conceptualized as learned associations whereby a stimulus triggers an associated response<sup>1-3</sup>. We propose that habits may be better understood as a process whereby a stimulus triggers only the preparation of a response, without necessarily triggering its initiation. Critically, this would allow a habit to exist without ever being overtly expressed, if the prepared habitual response is replaced by a goal-directed alternative before it can be initiated. Consistent with this hypothesis, we show that limiting the time available for response preparation<sup>4,5</sup> can unmask latent habits. Participants practiced a visuomotor association for 4 days, after which the association was remapped. Participants easily learned the new association but habitually expressed the original association when forced to respond rapidly (~300–600 ms). More extensive practice reduced the latency at which habitual responses were prepared, in turn increasing the likelihood of their being expressed. The time-course of habit expression was captured by a computational model in which habitual responses are automatically prepared at short latency but subsequently replaced by goal-directed responses. Our results illustrate robust habit formation in humans and show that practice affects habitual behaviour in two distinct ways: by promoting habit formation and by modulating the likelihood of habit expression.**

We are all familiar with the experience of committing habitual slips of action. For instance, if you are driving a right-hand-drive car but are used to a left-hand-drive car, you might habitually reach out for the gear stick with the wrong hand. Such slips of action are consistent with the theory, supported by behavioural, computational and neurobiological evidence, that behaviour is governed by two distinct systems<sup>1-3,6</sup>. The goal-directed system selects actions based on a prospective evaluation of which options will best achieve task goals, given knowledge of the current task and environment. By contrast, the habitual system selects actions purely based on what has been successful in the past. In principle, goal-directed selection should yield the best possible outcome, but can be computationally intensive and slow. Habitual action selection bypasses these time-consuming computations and is therefore simpler and faster than goal-directed selection. However, the habitual system is inflexible; it will persistently select the same actions even when they are no longer appropriate. These two systems are posited to operate in parallel and compete for control of action selection, but the exact nature of this competition remains unclear.

Studies in animals have demonstrated that extensive repetition leads to a transition from goal-directed to habitual behaviour<sup>2,7</sup> and have implicated different regions of the cortex and striatum in expressing goal-directed versus habitual behaviour<sup>8</sup>. However,

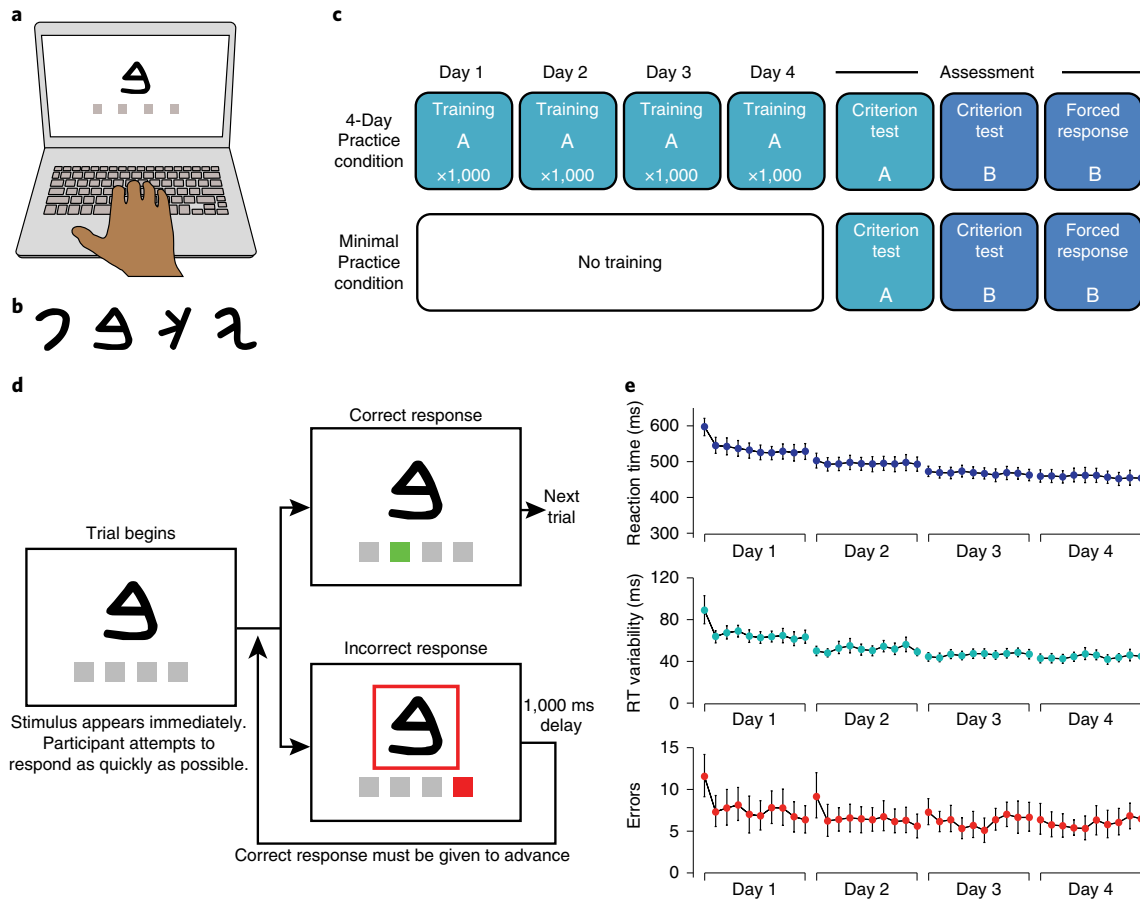
while everyday experience indicates that humans form habits in a similar manner, it has proved surprisingly difficult to induce habits in humans experimentally<sup>9</sup>. A common approach has been to train people to press specific buttons in response to visual stimuli (an arbitrary visuomotor association) then, after a period of practice, test whether this behaviour has become habitual by either switching the required responses for particular stimuli<sup>10</sup> or asking participants to withhold responses for certain stimuli<sup>9</sup>. Under both approaches, however, participants exhibit little evidence of becoming habitual, easily accommodating changed task requirements even after extensive practice of the original task<sup>9,10</sup>.

We propose that a major reason why habit formation has been difficult to demonstrate in humans is that, although people do form habits, these habits are masked by goal-directed processes. Specifically, we suggest that observation of a stimulus may trigger an associated response to be prepared by the motor system. Critically, however, the prepared response may not be immediately initiated, allowing it to be replaced by a more appropriate, goal-directed alternative before it can be expressed. This view is supported by recent work establishing that movement preparation occurs independently from, and often substantially earlier than, movement initiation<sup>4</sup>. Critically, we would expect that the habitual system can select potential responses rapidly, while the goal-directed system requires longer processing times<sup>11</sup>. We therefore reasoned that limiting the time participants had to prepare responses would prevent preparation of goal-directed responses and unmask otherwise latent habits.

Experiment 1 examined whether practicing a visuomotor association led to the association becoming habitual (Fig. 1a,b). We contrasted behaviour in two conditions: a 4-Day Practice condition and a Minimal Practice condition, both of which participants ( $n=22$ ) completed in a counterbalanced order (Fig. 1c). In the 4-Day Practice condition, participants first trained on a previously unseen stimulus–response mapping, completing 4,000 reaction-time-based trials (100 trials  $\times$  10 blocks  $\times$  4 days) in which they responded as quickly as possible to visual stimuli presented in rapid succession (Fig. 1d). Performance at this task improved with practice (Fig. 1e), with significant reductions in reaction times (average of first versus last day, paired samples  $t$ -test, difference 81 ms,  $t_{21}=11.96$ ,  $P<0.001$ ,  $d_z$  (Cohen's  $d$  for repeated measures) = 2.55, 95% confidence interval (CI) 66–95), reaction-time variability (paired samples  $t$ -test on reaction-time median absolute deviation, difference 23 ms,  $t_{21}=9.38$ ,  $P<0.001$ ,  $d_z=2.00$ , 95% CI 18–28) and errors (paired samples  $t$ -test, difference 1.7 errors,  $t_{21}=2.18$ ,  $P=0.041$ ,  $d_z=0.47$ , 95% CI 0.1–3.4).

We assessed whether practice led the association to become habitual by transposing the required responses for two of the four stimuli (Fig. 2a,b). Participants first learned this revised mapping in

<sup>1</sup>Department of Neurology, Johns Hopkins University, Baltimore, MD, USA. <sup>2</sup>Movement Control & Neuroplasticity Research Group, KU Leuven, Leuven, Belgium. <sup>3</sup>Department of Neuroscience, Johns Hopkins University, Baltimore, MD, USA. <sup>4</sup>Present address: Institute of Neuroscience, UC Louvain, Brussels, Belgium. \*e-mail: [robert.hardwick@uclouvain.be](mailto:robert.hardwick@uclouvain.be)



**Fig. 1 | Task and training schedule for Experiment 1.** **a**, Experimental setup. Participants ( $n = 22$ ) responded to the appearance of a stimulus by pressing a corresponding keyboard button. **b**, Example stimuli (letters of the Phoenician alphabet). **c**, Experiment 1 overview. Participants completed two conditions in a counterbalanced order. In the 4-Day Practice condition, participants completed 4,000 reaction-time-based training trials on original mapping A over 4 days. On day 5 (assessment) they were tested on their knowledge of mapping A in a criterion test block, then learned revised mapping B until they achieved a steady accuracy criterion (see Fig. 2d). They then completed forced-response trials (Fig. 2e) under this new mapping. In the Minimal Practice condition, participants only performed the assessment session (and therefore practiced the original mapping A only until they achieved a steady accuracy criterion). **d**, Trial structure of the reaction-time-based training task. Participants attempted to complete blocks of 100 trials as quickly as possible, incurring a 1 s time penalty for incorrect responses. **e**, Data from the reaction-time-based training in the 4-Day Practice condition. Participants' median reaction times, reaction-time (RT) variability (median absolute deviation) and error rates improved with training. Error bars represent bootstrapped 95% CI.

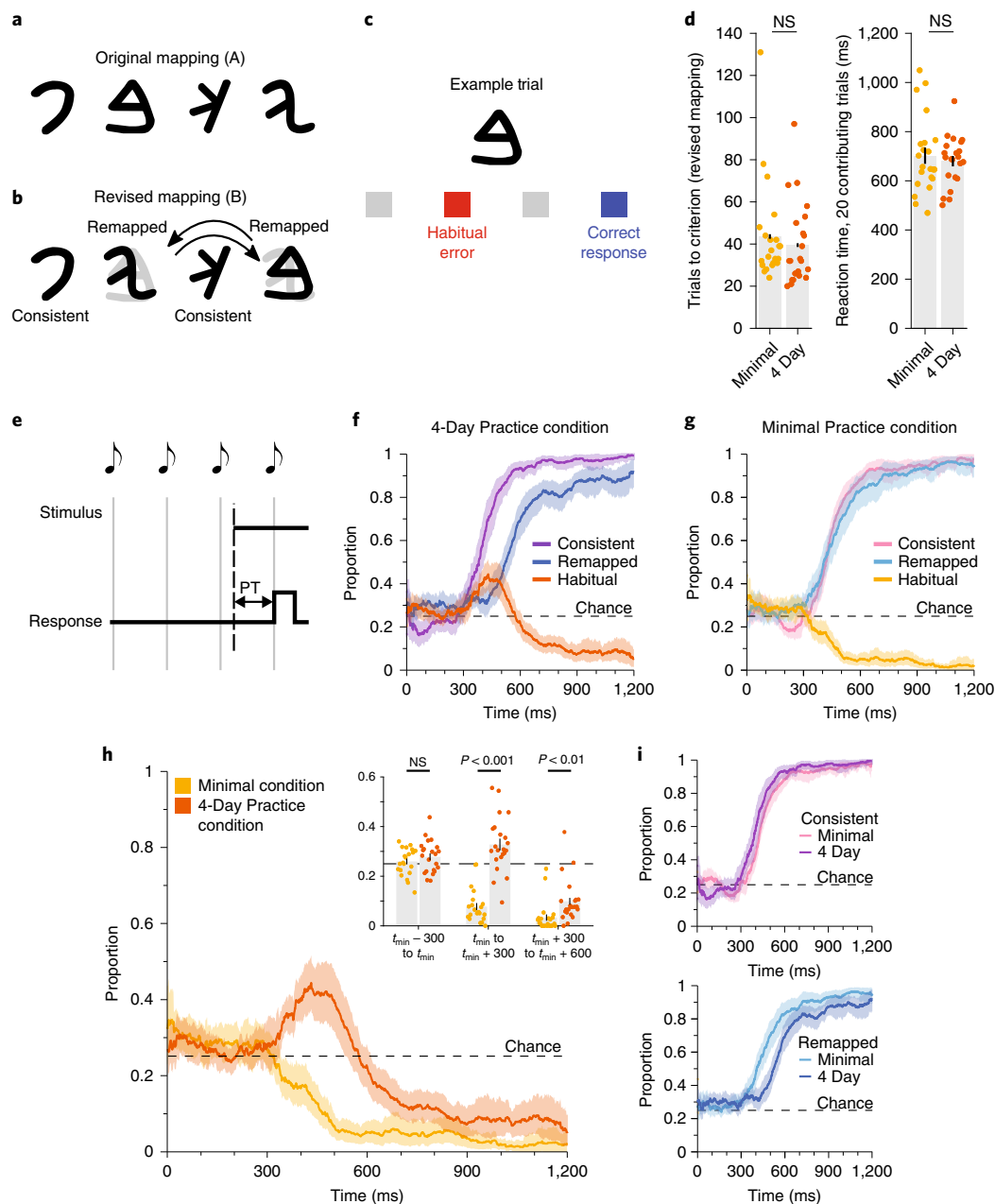
a criterion test block; they were instructed that there were no time constraints in this block and that they should focus on learning a new stimulus–response mapping. They trained on the revised mapping until satisfying an accuracy criterion of five consecutive correct responses to each of the four stimuli (a total of 20 responses). We did not expect participants to have any difficulty learning the revised mapping, even if the original mapping had become habitual, since their goal-directed system should be able to override any habit. Indeed, we found that participants in the 4-Day Practice condition required an average of  $44 \pm 5$  (mean  $\pm$  s.e.m.) trials to meet the accuracy criterion, which was not significantly different from the number of trials required to learn the revised mapping at the start of the Minimal Practice condition ( $40 \pm 4$  trials) (Fig. 2d; paired samples *t*-test, difference = 4 trials,  $t_{21} = 0.64$ ,  $P = 0.53$ ,  $d_z = 0.14$ , 95% CI  $-9.3$ – $17.4$ ). This suggested that 4 days of practice had not led the association to become habitual.

We then examined whether limiting response preparation time would unmask habitual preparation of the initially practiced response. We did so using a ‘forced-response’ paradigm<sup>4,5,12</sup> (see Methods). Participants heard a sequence of four equally spaced tones and were instructed to respond synchronously with the fourth

tone. We varied the onset of the stimulus in relation to the fourth tone, effectively enabling us to control the time the participants had to prepare their action across the range 0–1,200 ms (Fig. 2e). This allowed us to probe the responses that were prepared at different times relative to the presentation of the stimulus.

To visualize the time-course of response preparation, we assessed the probability of expressing different responses within a 100-ms sliding window (Fig. 2f). Focusing first on responses to stimuli that did not change between the mappings (consistently mapped stimuli; purple curve in Fig. 2f), the probability of generating the correct response was near chance (0.25) for preparation times less than  $\sim 300$  ms. This indicates that participants had insufficient time to process the stimulus and therefore had to guess which response was required. The probability of generating the correct response then rose gradually as preparation time increased, appearing to reach asymptote between 700 and 900 ms.

Habitual response preparation was revealed by examining the time-course of responses to remapped stimuli. In the 4-Day Practice condition, the probability of generating the originally learned response (Fig. 2f, orange curve) began at chance for low ( $< 300$  ms) preparation times but then transiently increased above chance at



**Fig. 2 | Switch manipulation, retraining and forced-response task results.** **a, b**, Following minimal practice or 4 days of practice on original stimulus–response mapping A (example mapping shown) (**a**), two stimulus–response associations were switched to create a revised mapping (**b**). **c**, This revised mapping allowed the identification of habitual behaviour—when participants acted according to the original mapping (habitual error) rather than the revised mapping (correct response). **d**, Participants trained on the revised mapping without reaction-time constraints until they reached a criterion of five consecutive correct responses to each of the four stimuli (i.e., 20 correct trials split across the four stimuli). Participants required ~40 trials to learn this revised mapping, regardless of the volume of training they had completed on the original mapping (left). We found no reaction-time differences in the 20 trials that contributed to this assessment between conditions (right). **e**, Participants then completed a forced-response task. Participants were instructed to respond synchronously with the final tone in a sequence of four equally spaced tones. Stimulus onset was varied relative to this time to impose differing preparation times (PT), distributed randomly and uniformly from 0 to 1,200 ms before the fourth tone. **f**, Distribution of responses as a function of allowed preparation time for the 4-Day Practice group. Each line presents the proportion of different types of responses within a 100 ms sliding window. Purple, correct responses, consistently mapped stimuli; blue, correct responses, remapped stimuli; orange, habitual errors, remapped stimuli. The brief increase in the probability of habitual errors between 300 and 600 ms indicates transient habitual preparation of the originally practiced mapping. **g**, Analogous results for the Minimal Practice condition, indicating no evidence of habitual action preparation. **h**, Direct comparison of the time-varying probability of expressing the habitual response across the groups. Inset compares the proportion of habitual responses binned over 300 ms intervals relative to the minimum time at which participants could respond to stimuli ( $t_{\min}$ , see text for details). **i**, Direct comparisons of speed–accuracy trade-offs across conditions for consistent (top) and remapped (bottom) stimuli. Shaded error regions (**f–i**) represent bootstrapped 95% CIs. Bar chart error bars represent  $\pm 1$  s.e.m. NS, not significant. All data  $n=22$ , counterbalanced crossover design.

preparation times of 300–600 ms, before declining towards zero as the proportion of correct responses (Fig. 2f, blue curve) began to increase. Therefore, although participants had no difficulty

in acquiring the revised mapping after 4 days of practice, forcing them to respond at low preparation times unmasked latent habitual responses. By contrast, in the Minimal Practice condition (Fig. 2g),

the proportion of habitual errors (Fig. 2g, yellow curve) began at chance, then declined as preparation time increased.

We summarized and confirmed these observations by analysing the proportion of habitual responses in a 300-ms interval aligned to the minimum possible time at which participants could generate an accurate response ( $t_{\min}$ , identified as the time at which a cumulative Gaussian fit to the speed–accuracy trade-off for consistently mapped stimuli first reached 5% of its height). This showed a significant interaction between condition and allowed preparation time (repeated measures ANOVA,  $F_{1,21}=39.1$ ,  $P<0.001$ , partial eta squared  $\eta_p^2=0.65$ ). Subsequent simple main effects analysis (paired samples  $t$ -tests) identified no statistically significant difference between conditions in the 300 ms before  $t_{\min}$  (difference = 0.02,  $t_{21}=1.31$ ,  $P=0.21$ ,  $d_z=0.28$ , 95% CI –0.01–0.05) but that the proportion of errors was significantly higher in the 4-Day Practice condition from  $t_{\min}$  to  $t_{\min}+300$  ms (difference = 0.25,  $t_{21}=11.16$ ,  $P<0.001$ ,  $d_z=2.38$ , 95% CI 0.20–0.30) and from  $t_{\min}+300$  to  $t_{\min}+600$  ms (difference = 0.07,  $t_{21}=3.14$ ,  $P=0.005$ ,  $d_z=0.67$ , 95% CI 0.02–0.11).

This pronounced tendency to express the previously practiced association at low preparation time suggested that practice brought about a qualitative change in participants' behaviour. This assertion was further supported by two additional observations. First, participants' ability to prepare appropriate responses for revised stimuli depended on whether the original association had been practiced (repeated measures ANOVA, significant condition by response type interaction,  $F_{1,21}=41.87$ ,  $P<0.001$ ,  $\eta_p^2=0.67$ ). In the 4-Day Practice condition, the speed–accuracy trade-off for remapped stimuli was shifted later relative to that for consistently mapped stimuli (Fig. 2f; purple versus blue curves, significant difference between the centre of the speed–accuracy trade-offs, difference 155 ms,  $t_{21}=10.68$ ,  $P<0.001$ ,  $d_z=2.28$ , 95% CI 125–186). In the Minimal Practice condition, there was no statistically significant difference in the speed of preparation between remapped and consistently mapped stimuli (difference = 33 ms,  $t_{21}=1.53$ ,  $P=0.14$ ,  $d_z=0.33$ , 95% CI –12–78). Second, even when long preparation times (900–1,200 ms) were allowed, participants were still liable to generate habitual responses after 4 days of practice but not after minimal practice; the probability of committing a habitual error for responses in this time range was clearly different for the 4-Day Practice and Minimal Practice conditions (Wilcoxon signed-rank test,  $Z=2.58$ ,  $P=0.01$ ,  $r=0.55$ ).

We also note that the observed effect could not have been attributable to participants simply becoming more familiar with the general task structure over 4 days of practice, rather than because they practiced a specific mapping. Due to the counterbalanced design, half of the participants in the Minimal Practice condition had previously completed the 4-Day Practice condition with a different set of symbols. However, these participants did not exhibit any evidence of habitual behaviour following remapping.

We developed a computational model of response preparation that accurately described the time-varying expression of habitual or goal-directed behaviour. This model assumes that different responses became prepared at different latencies after presentation of the stimulus (Fig. 3). In the case of a single learned association<sup>4</sup>, before the association was remapped, we assume that the response became prepared at some time,  $T_A$ , after presentation of the stimulus, with this time varying from trial to trial according to a Gaussian distribution (Fig. 3a). Critically, we assumed that this response is not necessarily generated immediately after being prepared but instead is held in a prepared state until the time of response initiation, at which time the response will be generated. So long as a response is initiated after  $T_A$ , the appropriate response will be generated. If a response is initiated before  $T_A$ , however, then the participant will not have had enough time to process the stimulus and will respond at a chance level of accuracy. These assumptions predict a speed–accuracy trade-off that qualitatively matches that observed for the

consistently mapped stimuli (Fig. 3a, lower panel). Improvements in the speed–accuracy trade-off through practice can be accounted for by improvements in the mean and variance of the distribution of  $T_A$  (Fig. 3b).

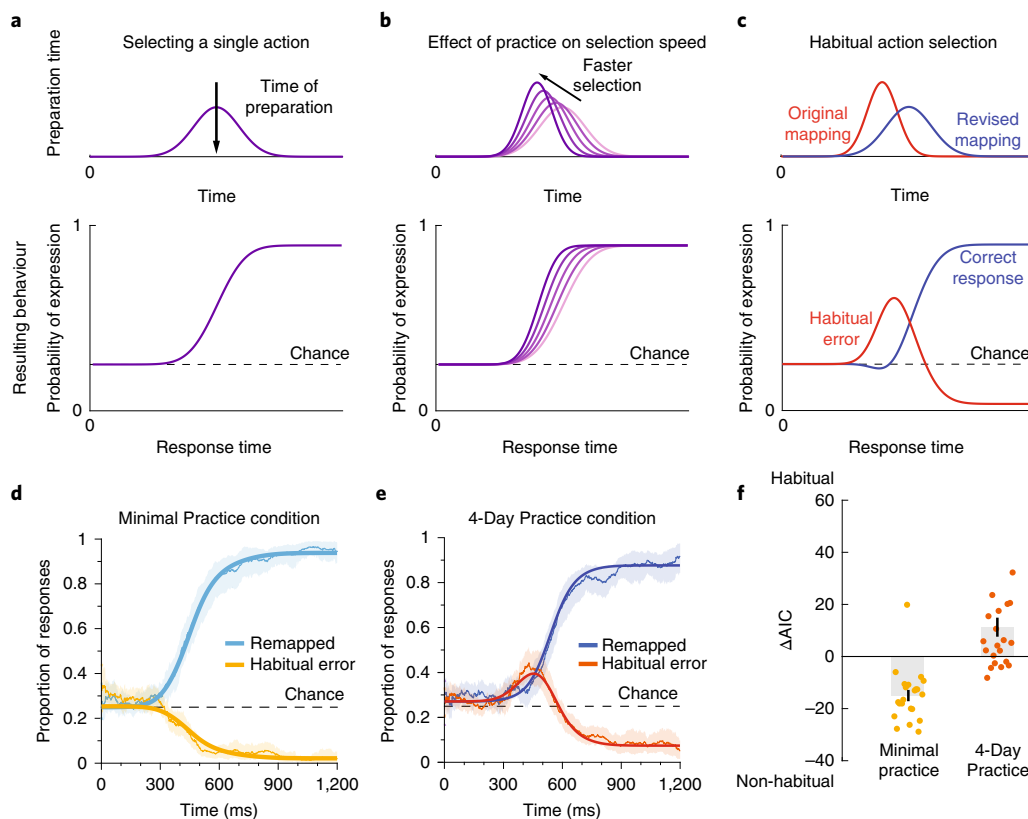
We extended this model to account for the possibility of multiple, competing response-selection processes (Fig. 3c). In this model, habitual and goal-directed processes select potential responses in parallel but compete for preparation of a single action. We assume that the habitual response becomes available at time  $T_A$ , while the correct, remapped response becomes available at a later time  $T_B$ . We assume that  $T_A$  and  $T_B$  are both Gaussian distributed and independent but with  $T_B$  having a greater mean and variance than  $T_A$ . Critically, participants will prepare the habitual response as soon as it is available (at time  $T_A$ ) but this will be replaced by the goal-directed response once that becomes available (at time  $T_B$ ). Participants will express whichever response is prepared at the time of movement initiation. Therefore a habitual response will be generated if the response is initiated after  $T_A$  but before  $T_B$ . This model replicated the stereotypical time-course we observed experimentally (Fig. 3c, lower panel). We also extended the model to include baseline error rates for goal-directed and habitual response selection, and to allow for the possibility of an occasional lapse by the goal-directed system (see Methods for details).

We contrasted this habitual preparation model ('habit' model) with an alternative model in which there was no habitual response preparation ('no-habit' model), equivalent to the single-process model described above (Fig. 3a). We first fit each model to data pooled across participants and compared the Akaike information criterion (AIC; a measure of the quality of a model's fit, penalized according to the number of free parameters it contains) for each fit, to determine which model best described the data. The habit model accounted for the data in the 4-Day Practice condition significantly better than the no-habit model (difference in AIC in favour of the habit model = 293.1). By contrast, data from the Minimal Practice condition was best explained by the no-habit model (difference in AIC in favour of the no-habit model = 217.0).

Our models not only accounted for the data aggregated across participants but also accounted well for the behaviour of individual participants (see individual fits in Supplementary Figs. 1 and 2). Conducting the same model comparison on individual-participant data for the 4-Day Practice condition, the habit model explained the data better for 17/22 participants (mean difference in AIC = 11.5, Fig. 3f) suggesting that most participants had become habitual following practice. In the Minimal Practice condition, the no-habit model better explained behaviour in 21/22 participants (mean difference AIC = 14.9, Fig. 3f). Thus, the results of Experiment 1, along with the computational model, established that 4 days of practice led to a qualitative change in how participants performed the task, with the originally practiced mapping being habitually prepared following practice. This habit was not apparent when participants completed the remapped task under self-paced conditions (in the criterion test blocks), but was unmasked by forcing them to respond rapidly.

In a second experiment, we investigated the effects of more extensive practice on habitual behaviour. In Experiment 2, a new group of participants ( $n=14$ ) trained on an original mapping over a period of 4 weeks, completing 20 days of practice in total (20-Day Practice condition; 1,000 trials per day, 20,000 total trials). As in Experiment 1, participants trained in reaction-time-based trials. As expected, participants' reaction times improved significantly with training (paired samples  $t$ -test, first versus last day of training, difference = 139 ms,  $t_{13}=12.61$ ,  $P<0.001$ ,  $d_z=3.37$ , 95% CI 115–162). The prolonged practice enabled participants to reduce their reaction times by 75 ms relative to participants at the end of the 4-Day Practice condition (mixed model ANOVA, group-by-day interaction for first/last day comparison,





**Fig. 3 | Computational model of response preparation.** **a**, We assume that, in each trial, a response is prepared at a random time (following a Gaussian distribution) after stimulus onset (top), giving rise to the observed speed–accuracy trade-off across trials (bottom). **b**, As the response becomes able to be selected more rapidly (top) the speed–accuracy trade-off improves (bottom). **c**, After the mapping is revised, the response appropriate for the original mapping (red) is still habitually prepared. The appropriate, revised response (blue) becomes available later and replaces the habitually prepared response. This sequence of events leads to a time-varying probability of each response being expressed (bottom). **d**, Averaged behaviour (thin lines,  $n=22$ ) and average model fit (bold lines) for the Minimal Practice condition. Fit shown here for the model with no habitual preparation (no-habit model; as in **a**), which had a lower AIC for this condition. Error bars show  $\pm$ s.e.m. **e**, Behaviour and model fit in the 4-Day Practice condition ( $n=22$ ). Fit shown is for the model with habitual preparation (habit model; as in **c**), which had a lower AIC for this condition. **f**, Difference in AIC between the habit and no-habit models for individual-participant fits to the data ( $n=22$ , counterbalanced crossover design). Results favoured the habitual preparation model in most participants in the 4-Day Practice condition but in only one participant in the Minimal Practice condition.

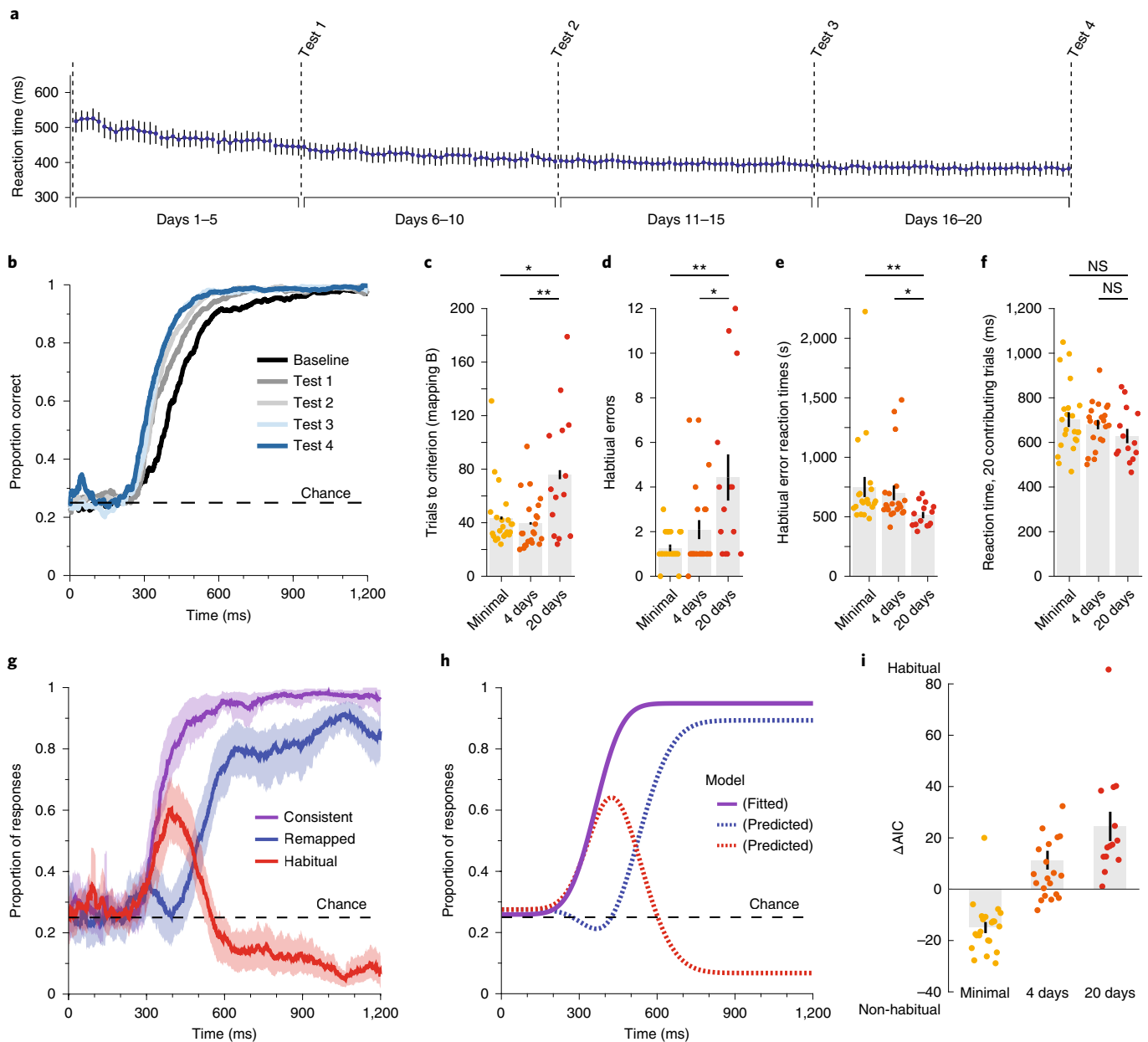
$F_{1,34} = 22.53$ ,  $P < 0.001$ ,  $\eta_p^2 = 0.40$ ;  $t$ -test showed no statistically significant difference between groups at baseline on day 1 of training, difference 17 ms,  $t_{34} = 0.89$ ,  $P = 0.38$ ,  $d = 0.30$ , 95% CI  $-22$ – $55$  and a statistically significant difference after training,  $75 \pm 16$  ms,  $t = 4.73$ ,  $P < 0.001$ ,  $d = 1.62$ , 95% CI  $43$ – $107$ ). The speed of participants' response preparation was also periodically assessed during learning through forced-response trials, yielding speed–accuracy trade-offs for the practiced mapping (Fig. 4). This data revealed a clear improvement in the latency at which participants could generate an accurate response (repeated measures ANOVA on the centre of the speed–accuracy trade-offs,  $F_{4,52} = 40.1$ ,  $P < 0.001$ ,  $\eta_p^2 = 0.76$ ; Fig. 4b).

Following practice, we tested whether response preparation had become habitual by remapping two of the four stimuli, as in Experiment 1. Participants trained on this new mapping until they made five consecutive correct responses to each of the four stimuli. Participants needed an average of  $76 \pm 13$  trials to achieve this criterion; significantly more than the  $44 \pm 5$  trials needed in the Minimal Practice condition ( $t$ -test, difference = 32 trials,  $t_{34} = 2.74$ ,  $P = 0.01$ ,  $d = 0.94$ , 95% CI  $8$ – $56$ ) or the  $40 \pm 4$  trials needed in the 4-Day Practice condition ( $t$ -test, difference = 36 trials,  $t_{34} = 3.24$ ,  $P = 0.003$ ,  $d = 1.11$ , 95% CI  $14$ – $59$ ). This increase was largely attributable to participants who had trained for 20 days committing more habitual errors (Fig. 4d, Mann–Whitney  $U$ -test versus Minimal Practice

condition,  $Z = 3.16$ ,  $P = 0.002$ ,  $n^2 = 0.29$  and versus 4-Day Practice condition,  $Z = 2.29$ ,  $P = 0.02$ ,  $n^2 = 0.15$ ).

Notably, when participants learned the revised mapping after 20 days of practice, the habitual errors that they made during the learning phase also had shorter reaction times during the criterion test block than those who practiced for only 4 days (Fig. 4e, linear mixed model, effect of condition,  $\chi^2 = 13.5$ ,  $P = 0.001$ , Mann–Whitney  $U$ -test versus 4-Day Practice condition,  $Z = 2.14$ ,  $P = 0.03$ ,  $n^2 = 0.13$ ), or those that had minimal practice (versus Minimal Practice condition,  $Z = 2.27$ ,  $P = 0.02$ ,  $n^2 = 0.15$ ). Average reaction times for those 20 trials that contributed to the accuracy criterion were not statistically significantly different across conditions (linear mixed model, no significant effect of condition,  $\chi^2 = 2.8$ ,  $P = 0.23$ ). Thus, in the 20-Day Practice condition, habitual responses were strongly associated with short response times, even though participants were allowed to act under self-paced conditions.

Participants were then tested under forced-response conditions (Fig. 4g), allowing more detailed examination of the effect of additional practice on their response preparation. As expected from the practice-based improvements in reaction time (Fig. 4a) and the speed–accuracy trade-off (Fig. 4b), the centre of the speed–accuracy trade-off for consistently mapped stimuli (Fig. 4g, purple line) was significantly earlier than that of participants in



**Fig. 4 | Results for Experiment 2 (20 days/20,000 trials of practice).** **a**, Participants' ( $n=14$ ) median reaction times over training. Each circle represents data for one block of 100 trials. Separations in the lines between points indicate between-day breaks. Error bars represent bootstrapped 95% CI. **b**, The speed-accuracy trade-off for the original stimulus-response association (mapping A), assessed using blocks of forced-response trials periodically during training: at baseline (just after achieving criterion) and at the end of each week of practice ( $n=14$ ). The leftward shift indicates that participants learned to execute the association more rapidly with practice. **c**, The number of trials required to achieve an accuracy criterion of five consecutive correct trials for each of the four stimuli in the revised mapping. The number of trials needed was greater in the 20-Day Practice condition ( $n=14$ ) than in the Minimal Practice or 4-Day Practice conditions ( $n=22$ ) in experiment 1. **d, e**, Habitual errors committed during the criterion test block (**d**). Participants in the 20-Day practice condition made significantly more habitual errors compared to the other conditions and did so with significantly faster reaction times (**e**) compared to conditions with less practice (note some participants made zero habitual errors in this condition (**d**), therefore in **e**  $n$  values from left to right are 20, 21 and 14). **f**, Reaction times for the 20 trials contributing to criterion success were comparable across conditions (Minimal/4-Day Practice  $n=22$ , 20-Day Practice  $n=14$ ). **g**, Forced-response data for the 20-Day Practice condition ( $n=14$ ). Speed-accuracy trade-off for consistently mapped stimuli (purple) and remapped stimuli (blue). The peak probability of expressing the originally practiced response (red) was even greater than in the 4-Day Practice condition (see text for statistical comparisons). **h**, Predicted behaviour in the 20-Day Practice condition (dotted lines), based on group-level model fits to behaviour in the 4-Day Practice condition ( $n=22$ ) and fits to the speed of preparation of responses to consistently mapped stimuli (solid line,  $n=14$ ). **i**, Difference in AIC between the habit and no-habit models for individual-participant fits to the data in all three conditions (Minimal and 4-Day Practice,  $n=22$  per condition; 20-Day Practice,  $n=14$ ). Positive values favour the habit model.

the 4-Day Practice condition in Experiment 1 ( $t$ -test on centre of the speed-accuracy trade-offs, 48 ms,  $t_{34}=2.36$ ,  $P=0.02$ ,  $d=0.81$ , 95% CI 6–90).

The speed at which responses to remapped stimuli could be prepared (Fig. 4g, blue line) was slower than for consistently mapped stimuli (paired samples  $t$ -test on centre of the speed-accuracy

trade-offs, difference = 156 ms,  $t_{13} = 10.9$ ,  $P < 0.001$ ,  $d_z = 2.92$ , 95% CI 125–187), though not significantly different from that observed in the 4-Day Practice condition ( $t$ -test on revised response speed–accuracy trade-offs, 48 ms,  $t_{34} = 1.80$ ,  $P = 0.08$ ,  $d = 0.62$ , 95% CI –6–102). Thus, given sufficient time, participants were capable of producing accurate, goal-directed responses.

As expected, forced-response trials revealed habitual response preparation were also noted in the 20-Day Practice condition (Fig. 4g, red line). The overall time-course of response preparation was similar to that for the 4-Day Practice condition, except that participants who had practiced for 20 days were more likely to produce habitual responses when forced to respond at low latencies ( $t$ -test on 20-Day Practice versus 4-Day Practice groups for 0–300 ms following  $t_{\min}$ , difference = 0.10,  $t_{34} = 2.47$ ,  $P = 0.019$ ,  $d = 0.84$ , 95% CI 0.02–0.18). The probability of committing a habitual error at longer preparation times did not differ significantly from that in the 4-Day Practice condition ( $t$ -test on 300–600 ms following  $t_{\min}$ , difference = 0.01,  $t_{34} = 0.30$ ,  $P = 0.77$ ,  $d = 0.10$ , 95% CI –0.07–0.05).

As with the 4-Day Practice condition, data from the 20-Day Practice group favoured the habitual preparation model over the no-habitual-preparation model, even more strongly than for Experiment 1 ( $\Delta$ AIC in favour of habitual preparation model = 329.7). At the individual-participant level, the data favoured the habit model in all 14 participants who completed the experiment (mean  $\Delta$ AIC = 24.5; Fig. 4i and Supplementary Fig. 3).

Our data showed that 20 days of practice led to more pronounced habitual behaviour compared to 4 days of practice. This was apparent in two distinct ways. First, participants made more habitual errors in the criterion test block after 20 days of practice. While this could suggest that the habit had become more difficult to override with goal-directed responses, data from the timed-response condition did not support this view; the speed–accuracy trade-off for generating an accurate, goal-directed response (blue lines in Fig. 2f versus Fig. 4g) was not statistically significantly different from that in the 4-Day Practice condition. Instead, the increase in the number of habitual errors occurring during this block was probably attributable to the fact that participants who trained for 20 days had responded with lower reaction times during the criterion test block. Participants may have persisted in responding with short reaction times that had been successful during their practice<sup>13</sup>, which in turn increased the likelihood that they would express the habitually prepared action.

Second, the peak probability of expressing a habitual error under forced-response conditions was greater in the 20-Day Practice condition than the 4-Day Practice condition. Although this might be interpreted as evidence that the habit became stronger, the model suggested an alternative interpretation: the probability of expressing a habitual error is strongly influenced by the speed at which responses can be prepared; more rapid preparation of a habitual response broadens the window of opportunity for it to be expressed before it is replaced by the appropriate, goal-directed response (Supplementary Fig. 4a). After 20 days of practice, participants had significantly reduced the latency at which they could select and prepare the habitual response (Fig. 4b), which may have driven the increase in peak probability of a habitual error. Adjusting the fitted model for the 4-Day Practice condition to account for this increased speed (that is, adjusting  $\mu_A$  and  $\sigma_A$  in the model) predicted an increase in the peak probability of a habitual response that closely matched experimental observations (see Fig. 4h; see Supplementary Information for further details). This suggests that the improved speed of response was probably the primary factor driving the increase in the peak probability of a habitual error.

In its present form, our response-selection model considers habits as an all-or-nothing phenomenon; either the previously practiced response will always be habitually prepared or will never be

habitually prepared. We also entertained the possibility of a continuum between these extremes in which the habitual response might be prepared in only a random subset of trials. This extension would correspond to a notion of variable habit ‘strength’, with ‘stronger’ habitual responses being more likely to be prepared in a given trial. In practice, however, a model-recovery analysis revealed that it is difficult to distinguish between such a continuum model and an all-or-nothing model. While it remains possible that habits could have an intermediate strength rather than being all-or-nothing, it is not necessary to invoke any such notion to account for our data.

In summary, Experiment 2 illustrates that extensive practice led to two distinct effects on behaviour. First, it enabled participants to prepare their habitual response more rapidly, that is their skill at the task improved. Second, participants tended to initiate their responses more rapidly when acting under self-paced conditions (in the criterion assessment block).

Our results and model clarify the nature of competition between habitual and goal-directed response selection. Although both systems can select specific actions in parallel with one another, they compete for which of these potential responses becomes prepared and ready to execute. This is consistent with the emerging view that, although multiple movement goals might be considered in parallel, only a single action is ever prepared<sup>14,15</sup>. Our results are also consistent with the view that preparation and initiation of actions occur separately<sup>4,16</sup>. Although the habitual response was always prepared, this did not necessarily trigger initiation. A brief delay between preparation and initiation instead allowed the habitual response to be replaced by a goal-directed one.

Habitual behaviour is expressed when a habitual response is prepared and initiated before it can be replaced by a goal-directed one. We have previously shown that the timing of movement initiation is subject to use-dependent biases<sup>13</sup>. Our data are consistent with this: in Experiment 2, during the criterion test block, participants appeared to persist with the low reaction times that had been successful during practice blocks and were consequently far more likely to express their habitually prepared responses. This is in contrast to Experiment 1, in which participants maintained longer reaction times during the criterion test block, flexibly delaying initiation to avoid committing habitual errors.

Whereas our paradigm directly probed preparation of a particular response by switching required responses, other approaches have sought to assess habitual behaviour based on whether participants could withhold a previously practiced response<sup>9,17</sup>. This latter approach failed to identify habits in humans in a study by de Wit et al.<sup>9</sup>. It is possible that forcing participants to respond more rapidly might similarly unmask habits, as in our experiments. However, being required to withhold any response at all is qualitatively different from being required to select a different response; it depends on flexibility of initiation, rather than preparation. Participants in the de Wit study could still have habitually prepared the practiced responses but were nevertheless able to avoid initiating any response. Such flexible control over response initiation is consistent with our results in Experiment 1 showing that, even though participants exhibited habitual response preparation after 4 days of practice, they were still able to delay response initiation in the criterion training block to avoid habitual errors.

The relative expression of different components of learning has previously been shown to be influenced by restricting cognitive resources, such as by inducing stress<sup>18,19</sup>, or by imposing a dual task<sup>10</sup>, as well as by limiting the time available to prepare an action<sup>20–24</sup>. However, previous research has manipulated these factors using relatively simple ‘high-or-low’ approaches. The forced-response paradigm used here allowed for a more fine-grained manipulation. Examining responses across a continuum of preparation times allowed us to precisely track the evolving competition between habitual and goal-directed response selection.

The fine-grained control over preparation time also allowed us to establish that practice enabled participants to generate accurate responses at lower latencies, as reflected in an improved speed–accuracy trade-off (they became more skilled at the task<sup>25–27</sup>). Both skilled and habitual behaviour are hallmarks of automatic performance<sup>28</sup>. Although definitions of automaticity vary, it is typically considered to involve improvements in skill, the ability to perform actions with little or no conscious attention and with a loss of flexibility<sup>29–31</sup>. Notably, our data indicate that skill improved continuously with practice, while the presence of habitual action preparation could be explained as an all-or-nothing phenomenon, suggesting a potential dissociation between skill acquisition and habit formation in the course of establishing automatic behaviour.

Previous research has failed to achieve any clear consensus on the neural basis of automaticity, proposing that it arises either through increases in network efficiency<sup>29,32</sup> or through shifts in the brain regions that control behaviour, either within the basal ganglia<sup>33</sup>, within cortex<sup>34</sup>, from the cortex to the cerebellum<sup>35</sup>, or from the basal ganglia to the cortex<sup>36</sup> (note that similar regions are also frequently implicated in both skill acquisition and habit formation<sup>6,37,38</sup>). We propose that these differing conclusions may have arisen because the tasks used to examine automaticity probably engaged multiple distinct learning processes associated with skill learning and habit formation, potentially to differing degrees. Using separate measures of skill acquisition and habit formation could therefore help to clarify the underlying neural basis of automaticity.

In our paradigm, participants were required to generate a response to a presented stimulus. As such, it is similar to most other paradigms that have been used to assess habit formation in humans<sup>9,10</sup>. It is different in many respects, however, from instrumental learning tasks typically used to examine habit formation in animals. In those paradigms, an animal is free to pursue any behaviour but, after learning that pressing a lever will earn a food reward, they will tend to perform that action more and frequently and persist in doing so, even after the reward is devalued. Such paradigms thus assess the rate of spontaneous engagement in a particular behaviour, rather than how a particular action is selected to react to a cuing stimulus. The extent to which such spontaneous engagement habits relate to the more reactive action-selection habits addressed in our paradigm remains unclear. However, both types of habit have been conceptually characterized in terms of stimulus–response associations—with some static feature of the environment (for example, the room or the presence of a lever) typically designated as the ‘stimulus’ in the case of instrumental learning<sup>1,39</sup>. Interestingly, in those cases, it has even been proposed that the ‘stimulus’ does not directly trigger the habitual response but instead leads to an ‘imperative to act’<sup>39</sup>—echoing our conclusion that habits occur at the level of movement preparation.

More generally, latent habits can emerge in any scenario in which goal-directed action is compromised, for instance when cognitive resources are occupied by another task or under stress<sup>19</sup>. However, we suggest that the strategy of forcing participants to initiate a response prematurely by limiting preparation time may provide a practical and effective means of revealing latent habits in any task in which participants must respond to an imperative stimulus. It is also plausible that preparation time might be limited by natural circumstances, such as avoiding a hazard while driving. Latent habits might therefore still have considerable behavioural relevance—particularly if they are most liable to be expressed when the stakes are high. Our ability to avoid committing habitual errors by delaying or withholding initiation is an example of ‘freedom from immediacy’<sup>40</sup> and is, we suggest, the key explanation for why it has proved so challenging to establish habit formation in humans compared to other animals.

In summary, our present results establish the existence of habitual response preparation and demonstrate how fine-grained behavioural assessment can unmask habits that might not be apparent

under conventional, self-paced approaches. Practice led to the formation of a habit in 4 days. However, further practice also brought about additional changes that made an existing habit more likely to be expressed. Thus, our results highlight an important distinction between formation and expression of habits. We suggest that dissociating these aspects of habitual responding is critical to achieve a complete understanding of habitual behaviour.

## Methods

**Participants.** A total of 39 participants took part in the study. Experiment 1 included 24 individuals. Two participants withdrew from Experiment 1 having completed only one of the two required experimental conditions, leaving 22 full datasets for the experiment (17 right-handed, 13 female, mean age 21 years). Experiment 2 included 15 participants. One participant withdrew due to computer hardware failure, leaving a total of 14 participants (12 right-handed, 4 female, mean age 26 years) to complete the experiment. All participants gave written informed consent and all procedures were approved by the Johns Hopkins School of Medicine Institutional Review Board. Participants received financial compensation (US\$15 per h) for their participation.

**General procedures.** Participants sat with the fingers of their dominant hand resting on a computer keyboard (Fig. 1a), with each of the four fingers positioned on a different key. Participants were required to respond to the appearance of one of four stimuli (letters of the Phoenician alphabet) by depressing a specific finger. The stimulus corresponding to each response was counterbalanced across participants, controlling for potential effects whereby participants would find some stimuli easier to recognize and learn to respond to than others. As Experiment 1 comprised two conditions and used a within-subjects design, we used two sets of distinct stimuli (see Supplementary Fig. 5) and counterbalanced the condition to which they corresponded across participants. Participants in Experiment 1 also completed the two conditions (Minimal Practice and 4-Day Practice) in a counterbalanced order. In all conditions, stimuli were presented pseudorandomly within 20 trial sub-blocks, with each stimulus appearing five times within a sub-block and the same stimulus appearing in two consecutive trials at the most. Participants attempted to respond to stimuli in training, criterion test or forced-response trial blocks.

**Training blocks.** During training, participants completed a gamified task in which they attempted to complete blocks of 100 reaction-time-based trials as quickly as possible (Fig. 1d). In each trial, a stimulus appeared in the centre of the screen and a tone played to signal the participant that a trial had started. On correct responses a pleasant tone sounded and after a 300 ms delay the task advanced to next trial. Errors were punished with a buzzer sound and a compulsory delay of 1,000 ms, after which the participant could once again respond to the same stimulus; this process repeated until participants produced the correct response, at which point the task progressed to the next trial. At the end of each block, participants received feedback on the time taken to complete each block and how this compared to their ‘personal best’ block completion time. Participants were encouraged to improve their performance by aiming to beat their personal best time each time they completed the task.

**Criterion test of mapping knowledge.** We assessed the ability of participants to learn new or revised stimulus–response associations using criterion test blocks. Participants were instructed that reaction-time constraints were removed, that their goal was to learn the correct set of stimulus–response associations, and that the block would end once they had made enough correct responses in a row. These blocks ended once participants had made five consecutive correct responses to each of the four stimuli (minimum of 20 trials) and the number of trials required to reach this steady, high-accuracy criterion was recorded.

**Forced-response blocks.** We used blocks of forced-response trials to probe the speed of response preparation and to assess whether participants habitually prepared their responses. Each block comprised 100 trials. In each trial, the participant heard a series of four tones, spaced 400 ms apart, and was instructed to synchronize his/her response with the onset of the fourth and final tone. The stimulus appeared at a random (uniformly distributed) time during the series of tones, effectively controlling the amount of time in which the participant could prepare his/her response. As such, in cases in which participants did not have sufficient time to process the stimulus (when it appeared less than ~300 ms before the deadline of the fourth tone), they were essentially forced to guess the correct response (and thus had a one in four chance of responding correctly). If participants selected the correct response, then the on-screen box corresponding to the button they pressed turned green; if their response was incorrect, the box turned red. On-screen feedback informed participants if their responses were ‘too early’ or ‘too late’, that is, if they responded more than 100 ms before or after the fourth tone. By contrast to the reaction-time condition, no time penalties were enforced for providing incorrect responses, as this would encourage participants to



ignore timing demands and instead focus on taking longer to provide an accurate response that avoided a penalty.

**Protocol. Experiment 1.** In Experiment 1, participants completed a counterbalanced, crossover design comprising two conditions. Both conditions began with a warm-up/familiarization task. Participants completed two blocks (200 trials total) of reaction-time-based trials in response to non-arbitrary stimuli (pictures of the hand with one finger coloured black to indicate the desired response—see Supplementary Fig. 6). This was followed by two blocks (200 trials total) of forced-response trials to the same non-arbitrary stimuli. This familiarization period allowed the experimenter to explain the practice and forced-response paradigms to the participant and to ensure that the participant could comply with the demands of each task.

Following this familiarization procedure, participants in the Minimal Practice condition then learned an original stimulus–response mapping (mapping A) in a block of criterion test trials, after which a second block of criterion test trials was used to introduce and assess the ability to learn a revised mapping (mapping B). We then probed for habitual response preparation using forced-response trials. The 4-Day Practice condition used the same assessment but this was completed after four consecutive days of practice (10 × 100-trial reaction-time training blocks each day) on mapping A.

**Experiment 2.** The second experiment comprised a single condition. All participants first completed the same familiarization procedure as in Experiment 1. Participants then completed a criterion test block in which they learned a set of stimulus–response associations through trial and error (mapping A). Once they had achieved this criterion, they completed 500 forced-response trials on this original mapping (to allow assessment of baseline performance), followed by 500 reaction-time-based training trials. Each day thereafter, participants completed training sessions in which they completed 10 × 100-trial blocks of reaction-time-based training trials. On the final (fifth) day of training for each week of practice, participants completed a training-and-probe session, in which they completed 500 (5 × 100-trial blocks) reaction-time-based training trials, followed by 500 (5 × 100-trial blocks) of forced-response trials. Participants completed 20 sessions in this manner (aiming to complete five sessions of practice in each 7-day week), allowing us to measure changes in performance at baseline and after each of the 4 weeks of practice.

On a separate day after all training sessions were complete, participants were exposed to the same assessment as in Experiment 1; they learned a revised set of stimulus–response associations in a criterion test block and their performance on this new mapping was then probed in 5 × 100-trial blocks of forced-response trials.

**Data analysis. Reaction-time trials.** Performance for each block was measured by taking the median reaction time (measured from stimulus onset to the time at which the key was pressed) for correct trials, the median absolute deviation of the reaction time, and by calculating the error rate for each block (that is, the number of erroneous responses in each block; note that it was possible for participants to make multiple errors in the same trial, as the trial did not advance until the participant provided the correct answer).

**Criterion test trials.** Criterion test trials were primarily analysed by counting the number of trials required for a participant to make five consecutive correct responses to each of the four stimuli. The reaction time and accuracy was recorded for each response (although participants were made aware that there were no reaction-time requirements for these trials).

**Forced-response trials.** Preparation times were calculated as the time between stimulus presentation and the participant’s first response. We examined the probability of three types of response: correct responses to consistently mapped stimuli (stimuli for which the same key press was required throughout the experiment); correct responses to the remapped associations; and responses consistent with the original mapping. A sliding window was used to visualize the time-varying probability for each of these response types; responses were binned over 100-ms windows, and the proportion of correct versus total responses was calculated and recorded for the centre of each window.

**Statistical analyses.** Data were analysed using parametric ANOVA or two-tailed *t*-tests as indicated in the text. In cases where the assumption of normality was violated, we instead used Mann–Whitney *U*-tests to compare independent samples and the Wilcoxon signed rank test to compare related samples.

**Response preparation models.** First, we consider the simple case of generating responses in the context of a single mapping A between stimuli and actions (see ref. 4). We assume that preparation of the response (becoming ready to generate the response) occurred as a discrete event at a random time  $T_A \approx N(\mu_A, \sigma_A^2)$ . Responses generated before  $T_A$  are assumed to be uniform across the four possible keys, reflecting the fact that participants were guessing. Responses generated later than  $T_A$  are assumed to be correct with probability  $q_A$  (with other responses distributed

uniformly across the other three keys). The probability of observing a correct response, given that the response was generated at time  $t$  is therefore given by

$$P(\text{correct}|t) = P(\text{correct}|t < t_A)P(t < t_A) + P(\text{correct}|t \geq t_A)P(t \geq t_A) = \frac{1}{4}(1 - \Phi_A(t)) + q_A \Phi_A(t)$$

where  $\Phi_A(t) = \Phi(\mu_A, \sigma_A^2)$  is the cumulative distribution of  $T_A$ , and  $\Phi$  represents the cumulative normal distribution function. The probability of an error is likewise given by

$$P(\text{error}|t) = \frac{3}{4}(1 - \Phi_A(t)) + (1 - q_A)\Phi_A(t)$$

We extended this model to include the possibility of two distinct mappings A (the original, habitual mapping) and B (the goal-directed mapping). We assumed that the associated responses become available at independent times  $T_A \approx N(\mu_A, \sigma_A^2)$  and  $T_B \approx N(\mu_B, \sigma_B^2)$  with  $T_B > T_A$ , reflecting the fact that we expect response B to be available, on average, later than response A. In this instance, when the presented symbol is one that has been remapped, there are three possible response types: (1) a correct response (according to the ‘true’ mapping, B), (2) a habitual error (correct according to the original mapping, A) and (3) an ‘other error’ (not correct for either A or B). The probability of each response type depends on which of the events, A or B, has occurred by the time of responding. To simplify notation, we use  $\bar{A}$  to denote the event  $t < T_A$  and  $A$  to denote the event  $t \geq T_A$ , with likewise definitions for  $\bar{B}$  and B. The probability of a given response being generated at time  $t$  is then given by:

$$P(r | t) = \begin{matrix} P(r | \bar{A}, \bar{B})P(\bar{A})P(\bar{B}) & P(r | \bar{A}, \bar{B})(1 - \Phi_A(t))(1 - \Phi_B(t)) \\ + P(r | A, \bar{B})P(A)P(\bar{B}) & + P(r | A, \bar{B})\Phi_A(t)(1 - \Phi_B(t)) \\ + P(r | \bar{A}, B)P(\bar{A})P(B) & + P(r | \bar{A}, B)(1 - \Phi_A(t))\Phi_B(t) \\ + P(r | A, B)P(A)P(B) & + P(r | A, B)\Phi_A(t)\Phi_B(t) \end{matrix}$$

The model is then finally specified by the distribution of different response types given each of these event types (that is,  $P(r | \bar{A}, \bar{B})$  and so on). We assume that the habitual and goal-directed processes select the response associated with their respective mappings with probability  $q_A$  and  $q_B$ , respectively. Critically, we assumed that when the goal-directed response B becomes available it will immediately be prepared, displacing any habitually prepared response A. So the probability of a correct response given that B has occurred will be  $q_B$  and we assume an equal probability of selecting any of the other three responses. Additionally, the likelihood of different responses might have been influenced by a possible bias in guessing either remapped or non-remapped keys, which we represented through the parameter  $q_i$ , reflecting the baseline probability of pressing one of the two remapped keys. The resulting distribution of responses can be expressed in matrix form:

$$\begin{pmatrix} P(\text{correct}|t) \\ P(\text{habitual error}|t) \\ P(\text{other error}|t) \end{pmatrix} = \begin{pmatrix} P(\text{correct}|\bar{A}, \bar{B}) & P(\text{correct}|A, \bar{B}) & P(\text{correct}|\bar{A}, B) & P(\text{correct}|A, B) \\ P(\text{habit}|\bar{A}, \bar{B}) & P(\text{habit}|A, \bar{B}) & P(\text{habit}|\bar{A}, B) & P(\text{habit}|A, B) \\ P(\text{other}|\bar{A}, \bar{B}) & P(\text{other}|A, \bar{B}) & P(\text{other}|\bar{A}, B) & P(\text{other}|A, B) \end{pmatrix} \times \begin{pmatrix} (1 - \Phi_A(t))(1 - \Phi_B(t)) \\ \Phi_A(t)(1 - \Phi_B(t)) \\ (1 - \Phi_A(t))\Phi_B(t) \\ \Phi_A(t)\Phi_B(t) \end{pmatrix} = \begin{pmatrix} q_1 & \frac{1 - q_A}{3} & q_B & q_B \\ q_1 & q_A & \frac{1 - q_B}{3} & \frac{1 - q_B}{3} \\ 1 - 2q_1 & \frac{2(1 - q_A)}{3} & \frac{2(1 - q_B)}{3} & \frac{2(1 - q_B)}{3} \end{pmatrix} \begin{pmatrix} (1 - \Phi_A(t))(1 - \Phi_B(t)) \\ \Phi_A(t)(1 - \Phi_B(t)) \\ (1 - \Phi_A(t))\Phi_B(t) \\ \Phi_A(t)\Phi_B(t) \end{pmatrix}$$

This model can be more compactly expressed by using  $P$ ,  $Q$  and  $\Psi$  to denote the matrices in the above equation:

$$P(t) = Q\Psi(t)$$

Finally, we extended this model to allow for the possibility that the goal-directed response might fail to override the habitual one. We assumed that the goal-directed response successfully replaced the habitual response with probability  $\rho$ . Note this parameter is distinct from  $q_A$  and  $q_B$ , which represent the probability of updating

the prepared response to the appropriate key, given that an update occurred. The parameter  $\rho$  allows the model to capture the possibility of committing habitual errors even at long reaction times, due to a lapse in replacing the habitual response with the goal-directed response. The possibility of lapses can be incorporated into the model through an additional matrix,  $R$ , which represents the probability of the preparatory state given different events (A, B) having occurred. The probability of each response type can then be expressed as:

$$P(t) = Q \begin{pmatrix} 1 & 0 & (1-\rho) & 0 \\ 0 & 1 & 0 & (1-\rho) \\ 0 & 0 & \rho & 0 \\ 0 & 0 & 0 & \rho \end{pmatrix} \Psi(t) \equiv QR\Psi(t)$$

The habit model therefore included eight parameters in total. However, since we found  $q_A$  to be difficult to reliably estimate in practice, we set it to  $q_A = 0.95$ , a nominal value reflecting typical success rates for already well-known mappings, leaving seven free parameters. The no-habit model, by contrast, has four free parameters:  $\mu_B, \sigma_B, q_A, q_B$ .

We estimated the parameters of these models based on maximum likelihood, using the MATLAB function `fmincon`. To achieve more robust fits to data that avoided the possibility of unrealistically steep speed-accuracy trade-offs ( $\sigma \approx 0$ ), we regularized the fits by adding an additional term to the overall log-likelihood that penalized values of  $\sigma_A$  and  $\sigma_B$  that deviated from some typical value  $\sigma_0$ :

$$L^* = L - \lambda(\sigma_A - \sigma_0)^2 - \lambda(\sigma_B - \sigma_0)^2$$

where  $\lambda$  reflects the strength of the penalty for deviating from  $\sigma_0$ . We set  $\sigma_0 = 0.1$  s and  $\lambda = 500$ , which avoided overfitting the slope of the speed-accuracy trade-off without biasing parameter recovery too much. Our results were qualitatively unaffected by the exact values of  $\sigma_0$  and  $\lambda$  used. A parameter recovery and model-recovery analysis indicated that we could reliably recover the correct model and its parameters from simulated datasets (see Supplementary Information).

**Reporting Summary.** Further information on research design is available in the Nature Research Reporting Summary linked to this article.

## Data availability

Data are available at <https://osf.io/3fjez>.

## Code availability

Analysis code are available at <https://osf.io/3fjez>.

Received: 8 October 2018; Accepted: 9 August 2019;

Published online: 30 September 2019

## References

- Robbins, T. W. & Costa, R. M. Habits. *Curr. Biol.* **27**, R1200–R1206 (2017).
- Dolan, R. J. & Dayan, P. Goals and habits in the brain. *Neuron* **80**, 312–325 (2013).
- Wood, W. & Rünger, D. Psychology of habit. *Annu. Rev. Psychol.* **67**, 289–314 (2016).
- Haith, A. M., Pakpoor, J. & Krakauer, J. W. Independence of movement preparation and movement initiation. *J. Neurosci.* **36**, 3007–3015 (2016).
- Ghez, C. et al. Discrete and continuous planning of hand movements and isometric force trajectories. *Exp. Brain Res.* **115**, 217–233 (1997).
- Morris, L. S. et al. Fronto-striatal organization: defining functional and microstructural substrates of behavioural flexibility. *Cortex* **74**, 118–133 (2016).
- Dickinson, A. Actions and habits: the development of behavioural autonomy. *Phil. Trans. R. Soc. Lond. B* **308**, 67–78 (1985).
- Killcross, S. & Coutureau, E. Coordination of actions and habits in the medial prefrontal cortex of rats. *Cereb. Cortex* **13**, 400–408 (2003).
- de Wit, S. et al. Shifting the balance between goals and habits: five failures in experimental habit induction. *J. Exp. Psychol. Gen.* **147**, 1043–1065 (2018).
- Hélie, S., Waldschmidt, J. G. & Ashby, F. G. Automaticity in rule-based and information-integration categorization. *Atten. Percept. Psychophys.* **72**, 1013–1031 (2010).
- Keramati, M., Dezfouli, A. & Piray, P. Speed/accuracy trade-off between the habitual and the goal-directed processes. *PLoS Comput. Biol.* **7**, e1002055 (2011).
- Schouten, J. F. & Bekker, J. A. M. Reaction time and accuracy. *Acta Psychol.* **27**, 143–153 (1967).
- Wong, A. L., Goldsmith, J., Forrence, A. D., Haith, A. M. & Krakauer, J. W. Reaction times can reflect habits rather than computations. *eLife* **6**, e28075 (2017).
- Wong, A. L. & Haith, A. M. Motor planning flexibly optimizes performance under uncertainty about task goals. *Nat. Commun.* **8**, 14624 (2017).
- Dekleva, B. M., Kording, K. P. & Miller, L. E. Single reach plans in dorsal premotor cortex during a two-target task. *Nat. Commun.* **9**, 3556 (2018).
- Kaufman, M. T. et al. The largest response component in the motor cortex reflects movement timing but not movement type. *eNeuro* **3**, ENEURO.0085-16.2016 (2016).
- Wit, S., de, Corlett, P. R., Aitken, M. R., Dickinson, A. & Fletcher, P. C. Differential engagement of the ventromedial prefrontal cortex by goal-directed and habitual behavior toward food pictures in humans. *J. Neurosci.* **29**, 11330–11338 (2009).
- Otto, A. R., Gershman, S. J., Markman, A. B. & Daw, N. D. The curse of planning: dissecting multiple reinforcement-learning systems by taxing the central executive. *Psychol. Sci.* **24**, 751–761 (2013).
- Schwabe, L. & Wolf, O. T. Stress prompts habit behavior in humans. *J. Neurosci.* **29**, 7191–7198 (2009).
- Otto, A. R. & Daw, N. The opportunity cost of time modulates cognitive effort. Preprint at *bioRxiv* <https://doi.org/10.1101/201863> (2017).
- Keramati, M., Smittenaar, P., Dolan, R. J. & Dayan, P. Adaptive integration of habits into depth-limited planning defines a habitual-goal-directed spectrum. *Proc. Natl Acad. Sci. USA* **113**, 12868–12873 (2016).
- Katnani, H. A. & Gandhi, N. J. Time course of motor preparation during visual search with flexible stimulus-response association. *J. Neurosci.* **33**, 10057–10065 (2013).
- Fernandez-Ruiz, J., Wong, W., Armstrong, I. T. & Flanagan, J. R. Relation between reaction time and reach errors during visuomotor adaptation. *Behav. Brain Res.* **219**, 8–14 (2011).
- Haith, A. M., Huberdeau, D. M. & Krakauer, J. W. The influence of movement preparation time on the expression of visuomotor learning and savings. *J. Neurosci.* **35**, 5109–5117 (2015).
- Reis, J. et al. Noninvasive cortical stimulation enhances motor skill acquisition over multiple days through an effect on consolidation. *Proc. Natl Acad. Sci. USA* **106**, 1590–1595 (2009).
- Hardwick, R. M., Rajan, V. A., Bastian, A. J., Krakauer, J. W. & Celnik, P. A. Motor learning in stroke: trained patients are not equal to untrained patients with less impairment. *Neurorehabil. Neural Repair* **31**, 178–189 (2017).
- Shmuelof, L., Krakauer, J. W. & Mazzoni, P. How is a motor skill learned? Change and invariance at the levels of task success and trajectory control. *J. Neurophysiol.* **108**, 578–594 (2012).
- Haith, A. M. & Krakauer, J. W. The multiple effects of practice: skill, habit and reduced cognitive load. *Curr. Opin. Behav. Sci.* **20**, 196–201 (2018).
- Hélie, S. & Cousineau, D. The cognitive neuroscience of automaticity: behavioral and brain signatures. *Cogn. Sci. (Hauppauge)* **6**, 35–53 (2011).
- Shiffrin, R. M. & Dumais, S. T. in *Cognitive Skills and Their Acquisition* (ed. Anderson, J. R.) 111–140 (Lawrence Erlbaum Associates, 1981).
- Moors, A. & De Hower, J. Automaticity: a theoretical and conceptual analysis. *Psychol. Bull.* **132**, 297–326 (2006).
- Wu, T., Kansaku, K. & Hallett, M. How self-initiated memorized movements become automatic: a functional MRI study. *J. Neurophysiol.* **91**, 1690–1698 (2004).
- Ashby, F. G., Turner, B. O. & Horvitz, J. C. Cortical and basal ganglia contributions to habit learning and automaticity. *Trends Cogn. Sci.* **14**, 208–215 (2010).
- Grol, M. J., Lange, F. P., de, Verstraten, F. A. J., Passingham, R. E. & Toni, I. Cerebral changes during performance of overlearned arbitrary visuomotor associations. *J. Neurosci.* **26**, 117–125 (2006).
- Balsters, J. H. & Ramnani, N. Cerebellar plasticity and the automation of first-order rules. *J. Neurosci.* **31**, 2305–2312 (2011).
- Helie, S., Roeder, J. L. & Ashby, F. G. Evidence for cortical automaticity in rule-based categorization. *J. Neurosci.* **30**, 14225–14234 (2010).
- Hardwick, R. M., Rottschy, C., Miall, R. C. & Eickhoff, S. B. A quantitative meta-analysis and review of motor learning in the human brain. *NeuroImage* **67**, 283–297 (2013).
- Graybiel, A. M. & Grafton, S. T. The striatum: where skills and habits meet. *Cold Spring Harb. Perspect. Biol.* **7**, a021691 (2015).
- Gardner, B. A review and analysis of the use of ‘habit’ in understanding, predicting and influencing health-related behaviour. *Health Psychol. Rev.* **9**, 277–295 (2015).
- Shadlen, M. N. & Kiani, R. Decision making as a window on cognition. *Neuron* **80**, 791–806 (2013).

## Acknowledgements

We thank E. Lesage and Y. Du for helpful comments on the manuscript and M. Adputra for producing the stimuli. This project was supported by an NSF grant (no. 1358756). This project has received funding from the European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant agreement No 702784

(R.M.H.). The funders had no role in the conceptualization, design, data collection, analysis, decision to publish or preparation of the manuscript.

### Author contributions

R.M.H. and A.M.H. conceived and designed the experiments. R.M.H. collected the data. R.M.H. and A.M.H. analysed the data. R.M.H. and A.M.H. wrote the manuscript. R.M.H., A.D.F., J.W.K. and A.M.H. reviewed and edited the manuscript.

### Competing interests

The authors declare no competing interests.

### Additional information

**Supplementary information** is available for this paper at <https://doi.org/10.1038/s41562-019-0725-0>.

**Correspondence and requests for materials** should be addressed to R.M.H.

**Peer review information** Primary Handling Editor: Marike Schiffer.

**Reprints and permissions information** is available at [www.nature.com/reprints](http://www.nature.com/reprints).

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

© The Author(s), under exclusive licence to Springer Nature Limited 2019