Figures and figure supplements

Catecholaminergic challenge uncovers distinct Pavlovian and instrumental mechanisms of motivated (in)action

Jennifer C Swart et al
**Figure 1.** Distinct mechanisms by which motivational valence may bias behavioural activation. (A) Pavlovian response bias: appetitive cues (green edge) elicit generalized behavioural activation (‘Go’), whereas aversive cues (red edge) elicit behavioural inhibition (‘NoGo’). This Pavlovian response bias is introduced in model M3a as the parameter $\pi$ (c.f. Figure 3). (B) Instrumental learning bias: rewarding outcomes (upper panel) facilitate learning of action (‘Go’, thick arrow) relative to inaction (‘NoGo’, thin arrow). Thus, learning effects at the individual trials $t$ will result in a cumulative selective increase of the rewarded action on later trials $t_n$. Punishment outcomes (lower panel) hamper the unlearning of inaction (‘NoGo’, dashed arrow) relative to action (‘Go’, solid arrow), resulting in sustained inaction. Neutral outcomes are equally well associated with actions and inactions, and are not illustrated here. The instrumental learning bias is introduced as the parameter $\kappa$ in model M3b (c.f. Figure 3). We assess whether these two mechanisms (i) act in parallel, and (ii) are modulated by the catecholamine system. To test the

---

**Swart et al. eLife 2017;6:e22169. DOI: 10.7554/eLife.22169**

2 of 20
latter, we administered methylphenidate (MPH), which prolongs the effects of catecholamine release via blockade of the catecholamine receptors. We first assess whether MPH affects the strength of the Pavlovian response bias, introduced as the parameter $\pi_{\text{MPH}}$ in model M5a, and instrumental learning bias, implemented as the parameter $\kappa_{\text{MPH-selective}}$ in model M5b (c.f. Figure 5). (C) We hypothesise that prolonged effects of dopamine release following reward outcomes will reduce (temporal) specificity, leading to spread of credit. Credit is assigned to other recent actions (thin arrow), in addition to the performed (and rewarded) Go response (thick arrow), resulting in additional learning of the alternative (not-performed) Go response. This MPH-induced diffuse learning bias is implemented by the parameter $\kappa_{\text{MPH-diffuse}}$ in model M5c (c.f. Figure 5). DOI: 10.7554/eLife.22169.003
Figure 2. Motivational Go/NoGo learning task and performance. (A) On each trial, a Win or Avoid cue appears on screen. Subjects can respond during cue presentation. Response-dependent feedback follows. (B) In total eight cues are presented for which the correct response needs to be learned. (C) Each cue has only one correct response (Go-left, Go-right, or NoGo), which subjects can learn from the feedback. (D) Feedback is probabilistic. Correct responses are followed by reward (Win cues) or a neutral outcome (Avoid cues) in 80% of the time and by a neutral outcome (Win cues) or punishment (Avoid cues) otherwise. For incorrect responses, these probabilities are reversed. (E) Trial-by-trial proportion of Go responses (±SEM) for Go cues (solid lines) and NoGo cues (dashed lines), collapsed over Placebo and MPH. Left: All cue types. From the first trial onwards, subjects made more Go responses to Win vs. Avoid cues (i.e. green lines are above red lines), reflecting the motivational bias. Additionally, subjects clearly learn whether to make a Go response or not (proportion of Go responses increases for Go cues and decreases for NoGo cues). Right: Go cues only. For the Go cues, a Go response could be either correct or incorrect. The motivational bias is present in both correct and incorrect Go responses, but incorrect Go responses are unlearnt. Note that the total p(Go) in this plot sums up to the solid lines in the left plot. (F) Mean (±SED) proportion Go responses. Proportion Go responses is higher for Go vs. NoGo cues, indicative of task learning. Additionally, subjects made more correct and incorrect Go responses to Win vs. Avoid cues. Source data of task performance are available in Figure 2—source data 1. Figure 2 continued on next page
Figure 2 continued

DOI: 10.7554/eLife.22169.004

The following source data is available for figure 2:

**Source data 1.** Source data for task performance under MPH and placebo.

DOI: 10.7554/eLife.22169.005
Figure 2—figure supplement 1. Individual traces (black lines) and group average (coloured lines) of correct and incorrect Go responses using a sliding average of 5 trials. Traces are averaged within cue types and over sessions. Individual traces are semi-transparent, so that darker areas reflect more overlaying subjects. Across trials, subjects increased correct Go responses (top) and decreased incorrect Go responses (bottom). Subjects performed at ceiling level more rapidly for the Go-to-Win cues (top-left) than Go-to-Avoid cues (top-right).

DOI: 10.7554/eLife.22169.006
Figure 3. Model evidence and parameter inference of base models. (A) Model evidence, relative to simplest model M1, clearly favours M4. The simplest model M1 contains a feedback sensitivity ($\rho$) and learning rate ($\epsilon$) parameter. Stepwise addition of the go bias ($b$), Pavlovian bias ($\pi$; Figure 1A), and instrumental learning bias ($\kappa$; Figure 1B) parameter improves model fit, quantified by WAIC (estimated log model evidence). Lower (i.e. more negative) WAIC indicates better model fit. (B) Temporal dynamics of the correlation between the motivational bias parameters (M4) and the predicted motivational bias, i.e. probability to make a Go response to Win relative to Avoid cues. The impact of the Pavlovian bias ($\pi$) on choice decreases over time (although, importantly, the parameter itself remains constant). This is because the instrumental values of the actions are learnt and thus will increasingly diverge. As a result, $\pi$ is less and less ‘able’ to tip the balance in favour of the responses in direction of the motivational bias (i.e. it can no longer overcome the difference in instrumental action values). In contrast, the impact of $\kappa$ on choice increases over time, reflecting the cumulative impact of biased learning (also Figure 3—figure supplement 2). (C) Posterior densities of the winning base model M4. Appendix 5 shows posterior densities for all models. (D) One-step-ahead predictions and posterior predictive model simulations of winning base model M4 (coloured lines), to assess whether the winning model captures the behavioural data (grey lines). Both absolute model fit methods use the fitted parameters to compute the choice probabilities according to the model. The one-step-ahead predictions compute probabilities based on the history of each subject’s actual choices and outcomes, whereas the simulation method generates new choices and outcomes based on the response probabilities (see Materials and methods for details). Both methods capture the key features of the data, i.e. responses are learnt (more ‘Go’ responding for ‘Go’ cues relative to ‘NoGo’...
Figure 3 continued

‘NoGo’ cues) and a motivational bias (more Go responding for Win relative to Avoid cues). We note that the model somewhat underestimates the initial Pavlovian bias (i.e. difference in Go responding between Win and Avoid trials is, particularly trial 1–2), while it overestimates the Pavlovian bias on later trials. This is likely the result from the fact that while the modelled Pavlovian bias parameter ($\alpha$) is constant over time, the impact of the Pavlovian stimulus values weakens over time, as the subjects’ confidence in the instrumental action values increases. Interestingly, notwithstanding the constancy of the Pavlovian bias parameter, we do capture some of these dynamics as Figure 3B shows that the impact of the Pavlovian bias on choice decreases over time. Source data of M4 simulated task performance are available in Figure 3—source data 1.

DOI: 10.7554/eLife.22169.007

The following source data is available for figure 3:

Source data 1. Source data for model M4 simulated task performance.
DOI: 10.7554/eLife.22169.008
**Figure 3—figure supplement 1.** Subject traces of model M4 (green/red) overlaid on observed behavior (black). M4 one-step-ahead predictions capture the individual variability in task performance.

DOI: 10.7554/eLife.22169.009
A. Strong vs. weak Pavlovian bias

Strong (π > 66%)  
Weak (π < 33%)

B. Strong vs. weak instrumental learning bias

Strong (κ > 66%)  
Weak (κ < 33%)

Figure 3—figure supplement 2. Illustration of the behavioural effects associated with the Pavlovian bias and instrumental learning bias parameters. Model M4 one-step-ahead predictions (coloured) overlaid on real data (grey) for the subjects with the upper versus lower tertile of parameter estimates. (A) Effects of Pavlovian bias (π). A strong Pavlovian bias (top 33% of π estimates) predicts higher Go responding for the Win than Avoid cues from the first trial onward, vice versa for a weak Pavlovian bias (33% lowest π estimates). (B) Effects of instrumental learning bias (κ). A strong instrumental learning bias (33% highest κ estimates) predicts steeper Go-to-Win learning and shallower Go-to-Avoid learning, vice versa for a weak instrumental learning bias (33% lowest κ estimates). See also Figure 3B for the temporal dynamics of the parameter-behaviour correlations.

DOI: 10.7554/eLife.22169.010
Figure 3—figure supplement 3. M4 subject-level parameters in model space (i.e. untransformed). The diagonal panels contain the posterior densities for the subject-level parameter means. The off-diagonal panels show the correlation over subjects in mean parameter estimates. Importantly, the two key parameters, Pavlovian bias (\(p\)) and instrumental learning bias (\(k\)) are not correlated to any of the other parameters. We do note that the feedback sensitivity parameter (\(r\)) is anti-correlated with the learning rate (\(e\)), such that the impact of high feedback sensitivity estimates is restricted by low learning rates. This correlation is not problematic, because independent estimation of learning rate and feedback sensitivity is no direct interest to the questions we ask.

DOI: 10.7554/eLife.22169.011
Figure 4. MPH-induced changes in motivational bias (i.e. proportion of Go responses to Win relative to Avoid cues). (A) Mean (±SED) proportion Go responses under MPH relative to Placebo. MPH did not significantly alter the motivational bias across the group (p=0.22; ns indicates p>0.05). (B) MPH increased the motivational bias in high span subjects, yet decreased it in low span subjects (R = 0.21; p=0.016). (C) MPH altered the motivational bias particularly for incorrect Go proportional to working memory span (incorrect Go: p<0.001; correct Go: p=0.152).

DOI: 10.7554/eLife.22169.013
**Figure 4—figure supplement 1.** Simple effects of MPH-induced changes in motivational bias. (A) The span-dependent motivational bias emerged under MPH (right; $p=0.032$), and was not significant under placebo (left; $p=0.34$). (B) MPH did not significantly alter the motivational bias proportional to working memory span for correct Go responses (correct Go: $p=0.15$).

DOI: 10.7554/eLife.22169.014
Figure 5. Model evidence and parameter inference of extended MPH models. (A) Model evidence (WAIC) relative to winning base model M4. We tested whether MPH alters the strength of the Pavlovian response bias ($\pi_{MPH}$; M5a), the instrumental learning bias ($k_{MPH-selective}$; M5b), or has a diffuse effect on the learning bias ($k_{MPH-diffuse}$; M5c; Figure 1C). Model selection favoured the composite model M6, including the $\pi_{MPH}$ and $k_{MPH-diffuse}$ parameters. (B) Posterior densities of the top-level parameters of M6. (C) Subject-level estimates of MPH-induced Pavlovian bias parameter (upper) and the MPH-induced diffuse learning bias parameter (lower; logistic scale) correlated significantly with Listening Span. (D) One-step-ahead model predictions and posterior predictive model simulations of M6 using subject-level parameter estimates. The model predictions and simulations echo the observed data, i.e. that the motivational bias correlates positively with working memory span (Figure 4B), confirming the winning model M6 captures the MPH-induced increase in Go responses to Win vs. Avoid cues.

DOI: 10.7554/eLife.22169.015
Figure 5—figure supplement 1. Illustration of the behavioural effects of MPH related to the Pavlovian bias and diffuse learning bias parameters. Model M6 one-step-ahead predictions (coloured) overlaid on real data (grey) for the subjects with the 33% strongest vs. weakest parameter estimates. The coloured bars at the bottom indicate the trial-by-trial correlation across all subjects, of the parameter estimate with the effect of MPH on Go responding per cue. The R value indicates the average correlation. (A) The effect of MPH on Pavlovian bias ($p_{\text{MPH}}$). Strong $p_{\text{MPH}}$ estimates predict that MPH increases the motivational bias (increased Go to Win cues and decreased Go to Avoid cues), and vice versa for weak $p_{\text{MPH}}$ estimates. The influence of $p_{\text{MPH}}$ is present from the first trial onward and decreases over time as indicated by the correlation coefficients. (B) Effect of MPH on diffuse learning bias ($\kappa_{\text{MPH-diffuse}}$). Strong $\kappa_{\text{MPH-diffuse}}$ estimates predict that MPH increases the motivational bias for Win cues specifically, whereas this effect is diminished for subjects with relatively weak $\kappa_{\text{MPH-diffuse}}$ estimates. The effect of $\kappa_{\text{MPH-diffuse}}$ is experience-dependent and evolves over time. These one-step-ahead predictions illustrate how each parameter results in an increased motivational bias under MPH, but with unique temporal dynamics, even though the parameter themselves are constant.

DOI: 10.7554/eLife.22169.016
Figure 5—figure supplement 2. M6 subject-level parameters in model space (i.e. untransformed). The diagonal panels contain the posterior densities for the subject-level parameter means. The off-diagonal panels contain the parameter correlations over subjects. Importantly, the parameters estimating the effects of MPH on Pavlovian bias ($\pi$) and diffuse learning bias ($\kappa$) are not correlated with any of the other parameters.
Appendix 3—figure 1. Logistic mixed model estimates of the probability of Go responses to verify that exclusion of a subset of subjects (7) did not affect our inference. Left: N = 99; Right: N = 106. Fixed effect estimates and 95% confidence interval (CI) are plotted on probability scale. Effects are sorted by lower bound of the CI. The results including all 106 subjects replicate the findings when discarding the subset of subjects (four subjects dropped out after the first test day, two subjects dissolved the capsules before swallowing, one subject did not sufficiently engage in the task).

DOI: 10.7554/eLife.22169.021
Appendix 5—figure 1. Model selection and parameter inference for base models and extended MPH models including all subjects (N = 99). (A–B) Model selection favours M4 of the base models and M6b of the extended MPH models as reported in the main article. Note that M6b in this figure corresponds to M6 in the main manuscript. (C–D) Posterior densities of top-level parameters of the winning base and MPH model, in model space (i.e. transformed). Only κ is presented untransformed (i.e. in sample space), as it is added to τ0 prior to transformation. (E) As reported in the main article, πMPH and κMPH-diffuse of M6b positively correlate with Listening Span. (F) In the composite model M6a, κMPH-selective Correlates negatively with Listening Span (N = 94; Rho = −0.22, p=0.036), which further supports that this parameter cannot capture the positive relation between listening span and the effect of MPH on motivational bias. Note that we report the correlation here for the 94 subjects for whom the parameters were reliably estimated, i.e. model convergence was reached.

DOI: 10.7554/eLife.22169.022
Appendix S—figure 2. Average one-step-ahead predictions for the base models M2-4 overlaid on the observations in grey. The one-step-ahead predictions indicate the action probabilities as predicted by the model, using each subject’s actual choices and outcomes.

DOI: 10.7554/eLife.22169.026
Appendix 5—figure 3. Average one-step-ahead predictions for the extended MPH models M5-6 overlaid on the observations in grey. The one-step-ahead predictions generate the action probability of each choice, based on the history of the subject’s actual choices and outcomes preceding the choice. The predictions are separately plotted for MPH (top) and placebo (bottom). We observed no main effect of MPH on the motivational bias (i.e. more Go to Win cues relative to Avoid cues). Accordingly, all models make highly similar predictions under MPH and placebo across the group.

DOI: 10.7554/eLife.22169.027