A Multi-Level Reinforcement-Learning Model of Wisconsin Card Sorting Test Performance

Alexander Steinke (Steinke.Alexander@MH-Hannover.de)
Hannover Medical School, Carl-Neuberg-Straße 1
30625 Hannover, Germany

Florian Lange (Florian.Lange@KULeuven.be)
KU Leuven, Naamsestraat 69,
3000 Leuven, Belgium

Bruno Kopp (Kopp.Bruno@MH-Hannover.de)
Hannover Medical School, Carl-Neuberg-Straße 1
30625 Hannover, Germany

Abstract:

The Wisconsin Card Sorting Test (WCST) is considered to be gold standard for the clinical assessment of executive functions. However, little is known about cognitive processes corresponding to WCST performance. Recent research suggests that multiple levels of control contribute to WCST performance. In this study, we introduce a reinforcement-learning (RL) model, which incorporates category and response learning. We test this multi-level RL model against single-level models, i.e., a category RL model and the state-of-the-art attentional updating model, by means of relative and absolute model performance. A sample of 375 participants completed a computerized version of the WCST (cWCST). Behavioral outcome measures were traditional perseveration and set-loss errors that we further stratified by response demands. The multi-level RL model outperformed both single-level models, with the state-of-the-art attentional updating model performing worst. Only the multi-level RL model was able to simulate all behavioral phenomena under consideration. In conclusion, results of model comparisons support the hypothesis that control processes at multiple levels contribute to cWCST performance. The multi-level RL model might offer a suitable framework for discerning latent cognitive processes contributing to WCST performance in general.

Keywords: Wisconsin Card Sorting Test; Reinforcement Learning; Model Comparisons

Introduction

The Wisconsin Card Sorting Test (WCST) is considered to be gold standard for the clinical assessment of executive functions. However, little is known about which cognitive processes correspond to WCST performance. In order to identify and isolate the cognitive processes that drive performance on the WCST, behavioral analyses of WCST data need to be complemented with computational modeling analyses. The WCST requires participants to match target cards to one of four key cards by categories that periodically change. Key card choices are followed by positive or negative feedback (see Figure 1). Individual performance relies on the ability to adapt categories by evaluating feedback (e.g., to avoid repeating a category following negative feedback). Recent research (Kopp, Steinke, Bertram, Skripuletz, & Lange, under review) suggests that negative feedback also induces behavioral avoidance of motor responses irrespective of categories. These authors concluded that trial-by-trial learning at multiple levels of control, i.e., at category and response levels, contribute to WCST performance. In a follow-up of this hypothesis, we introduce a multi-level reinforcement-learning (RL) model of WCST performance. We compare it to a single-level RL model that solely operates at the category level. We further compare both RL models with the state-of-the-art attentional updating model (Bishara et al., 2010) as the benchmark for this model comparison.
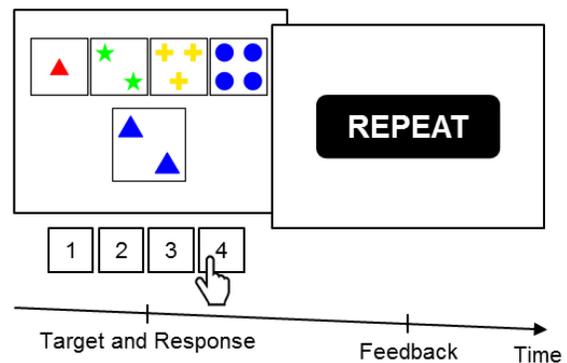

Figure 1: A showcase trial on the WCST. The target card (two blue triangles) can be sorted by the color category (far right key card), the shape category (far left key card), or the number category (inner left key card). In this example, the color category is applied, as response 4 is pressed that spatially maps the far right key card. A positive feedback indicates that the given

response was correct and the color category should be repeated on the next trial.

Methods

Data Collection

Participants A total of *N* = 407 participants (155 male, two preferred not to say; *M* = 23.47 yrs; *SD* = 4.83 yrs) completed a computerized version of the WCST (cWCST). We excluded 32 participants due to invalid test performance, resulting in a final sample of *N* = 375 participants (144 male, one preferred not to say; *M* = 23.17 yrs; *SD* = 4.37 yrs). Test performance was considered invalid when the frequency of a category or an odd choice (i.e., a response that matched no category) fell below or above the overall mean of that category or of odd choices plus/ minus three standard deviations. The study was approved by the local ethics committee of the KU Leuven (G-2016 12 694).

Wisconsin Card Sorting Test The cWCST (see, for example, Steinke, Lange, Seer, & Kopp, 2018) requires participants to match cards according to one of three possible categories. Stimulus cards varied on three dimensions that equaled the three viable categories *U* = {color, form, number}. Participants indicated their choice by pressing one of four keys *V* = {response 1, response 2, response 3, response 4} that were spatially mapped to the position of the key cards *K* = {one red triangle, two green stars, three yellow crosses, and four blue balls}. Responses were followed by a positive or negative feedback cue ("REPEAT" or "SWITCH", respectively). Categories changed in an unpredictable manner after runs of two or more correct category repetitions. Participants were required to complete 42 runs (including 41 category switches). Participants were given a maximum of 250 trials to complete these 42 runs and six practice runs. Prior to the experimental session, participants were explicitly informed about the three possible sorting categories and about the fact that the valid category would change from time to time. For a detailed description of the cWCST, see Steinke et al. (2018).

Traditional set-loss errors (a switch of the applied category after positive feedback) and perseveration errors (a repetition of the applied category after negative feedback) served as behavioral outcome measures. As Kopp et al. (under review) did with a traditional paper-and-pencil version of the WCST (Schretlen, 2010), we stratified these error scores by response demands (i.e. repetition vs. alternation). The resulting four trial types could be grouped based on feedback on *t*-1 trials: Negative feedback *t*-1 trials (cf. perseverative errors) led to reduced error probabilities when response demands shifted from repeat to alternate (in the latter case, a perseveration error was committed by repeating the previously executed response), indicating behavioral avoidance. In contrast, on positive feedback *t*-1 trials (cf. set-loss errors), these authors reported no modulation of error probabilities by response demands.

Computational Modeling

We utilized the RL framework of "Q-learning" to model trial-by-trial WCST performance (Sutton & Barto, 1998). The introduced multi-level RL model is based on the assumption that feedback on the WCST can be attributed to the applied category, but also to the executed response. Following, participants form independent feedback expectations of category (c) and response (r) choices.

Category-Level Reinforcement Learning Category-level RL operates on a 3 (categories) x 1 vector $\mathbf{Q}_c(t)$, which quantifies the expected feedback for each category on trial *t*. For trial-wise updating of $\mathbf{Q}_c(t)$, expected-feedback values decay as:

$$\mathbf{Q}'_c(t) = \gamma_c * \mathbf{Q}_c(t) \tag{1}$$

where γ_c gives the strength of decay. γ_c ranges from 0 to 1, with low values representing higher decay of expected feedback. Next, trial-wise prediction errors $\delta_c(t)$ are computed with regard to the category $u \epsilon U$, which has been applied on trial *t*, as:

$$\delta_c(t) = r(t) - Q'_{c,u}(t) \tag{2}$$

where *r*(t) is 1 for positive and -1 for negative feedback. Expected-feedback values of categories are updated following a delta-learning rule:

$$\mathbf{Q}_c(t + 1) = \mathbf{Q}'_c(t) + \mathbf{Z}_c(t) * \alpha_c * \delta_c(t) \tag{3}$$

where $\mathbf{Z}_c(t)$ is a 3 x 1 dummy vector, which is 1 for the applied category *u* and 0 for all other categories on trial *t*. $\mathbf{Z}_c(t)$ ensured that only the expected feedback value of the applied category is updated in response to the prediction error. We assumed distinct learning rate parameters for positive and negative feedback, α_c^+ and α_c^-, which quantify the degree to which received feedback affects subsequent expected feedback. Learning rates range from 0 to 1.

Response-Level Reinforcement Learning Response-level RL parallels the trial-wise updating on category-level. However, response-level RL operates on a 4 (responses) x 1 vector $\mathbf{Q}_r(t)$, which gives expected-feedback values for the execution of responses on trial *t*. First, the decay of $\mathbf{Q}_r(t)$ is computed as:

$$\mathbf{Q}'_r(t) = \gamma_r * \mathbf{Q}_r(t) \tag{4}$$

where γ_r modulates the strength of decay. Trial-wise prediction errors on response level are computed with regard to the executed response $v \epsilon V$ on trial *t* as:

$$\delta_r(t) = r(t) - Q'_{r,v}(t) \qquad (5)$$

Next, expected-feedback values are updated as:

$$\mathbf{Q}_r(t+1) = \mathbf{Q}'_r(t) + \mathbf{Z}_r(t) * \alpha_r * \delta_r(t) \qquad (6)$$

where $\mathbf{Z}_r(t)$ is a 4 x 1 dummy vector that is 1 for the executed response v and 0 for all other responses on trial t, which, again, ensured that only expected-feedback values of the given response are updated in response to the prediction error. We assumed different learning rate parameters for positive and negative feedback, α_r^+ and α_r^-, respectively.

Level Integration and Choice Probabilities In order to compute choice probabilities for key cards, expected-feedback values on category- and response-level are integrated. The integrated expected-feedback value for key card $k \in K$ on trial t is computed as:

$$Q_{sum,k}(t) = \mathbf{X}_k^T(t)\, \mathbf{Q}_c(t) + \mathbf{Y}_k^T(t)\, \mathbf{Q}_r(t) \qquad (7)$$

with $\mathbf{X}_k(t)$ is a 3 x 1 vector that represents the match between a target card and key card k on trial t with regard to the color, form, and number category. Here, 1 indicates a match and 0 indicates no match. Likewise, $\mathbf{Y}_k(t)$ is a 4 x 1 vector that represents the match between key card k and responses 1 to 4 on trial t. $\mathbf{X}_k^T(t)$ and $\mathbf{Y}_k^T(t)$ denoted the transpose of $\mathbf{X}_k(t)$ and $\mathbf{Y}_k(t)$. We set $\mathbf{X}_k^T(t)\, \mathbf{Q}_c(t)$ in equation 7 to -1, if key card k on trial t matches none of the categories. Finally, the choice probability of key card k on trial t is computed using a "softmax" logistic function on integrated expected-feedback values as:

$$P_k(t) = \frac{e^{\frac{Q_{sum,k}(t)}{\tau}}}{\sum_{j=1}^{4} e^{\frac{Q_{sum,j}(t)}{\tau}}} \qquad (8)$$

with τ is an inverse temperature parameter indicating whether differences in expected rewards are attenuated ($\tau > 1$) or emphasized ($0 < \tau < 1$).

Model Space We considered three computational models of WCST performance. First, we implemented the model of category- and response-level RL as described above. Second, we implemented a category RL model, i.e., trial-by-trial updating of expected-feedback values accorded to equations 1 - 3 and choice probabilities were computed by adapting equation 8 on $\mathbf{Q}_c(t)$. Note that we did not implement a RL model that operates at the response-level only, as it is psychologically implausible. Finally, we implemented the state-of-the art model of attentional updating (Bishara et al., 2010) as the benchmark for model comparison. Note that we used a configuration of the attentional updating model with all four individual parameters set free to vary.

Results

Behavioral Analysis

Results of behavioral analysis are presented in Figure 2 (upper left plot). Observed error probabilities were overall higher after negative feedback than after positive feedback. Error probabilities were reduced when response demands shifted from repeat to alternate. However, this reduction appeared solely after negative feedback, a finding that replicates on the cWCST the WCST-based finding of behavioral avoidance (Kopp et al., under review).

Computational Modeling

We used hierarchical Bayesian analysis for individual parameter estimation by means of RStan (Stan Development Team, 2018). Relative model performance was assessed by 5-fold cross validation according to the procedure outlined by Vehetari, Gelman, and Gabry (2017). Relative model performance was quantified by the difference in expected log pointwise predictive densities (elpd) between the model with the lowest absolute elpd and any other model. Higher absolute elpd indicates better model performance, hence larger negative Δelpd-values indicate worse model performance. We also report standard errors associated with the Δelpd-values. Relative model comparison results are presented in Table 1. The best performing model was based on category- and response-level RL followed by the RL model that operates solely at the category-level. Both models outperformed the state-of-the-art attentional updating model (Bishara et al., 2010), which should be considered as the benchmark for model comparison.

Table 1: Results of 5-fold cross validation.

Model	Level	Pars	Δelpd	SE
Category- and Response-Level RL	2	7	0	0
Category-Level RL	1	4	-85	28
Attentional Updating	1	4	-5499	363

Note. RL = reinforcement learning; Pars = number of free parameters; Δelpd = difference in expected log pointwise predictive density between a model and the best performing model; SE = standard error of Δelpd.

Relative model comparisons are not informative about a model's ability to simulate the behavioral phenomenon of interest (Palminteri, Wyart, & Koechlin, 2017). Therefore, we assessed absolute model performance by simulating individual card choices according to the post hoc absolute fit method (Steingroever, Wetzels, & Wagenmakers, 2014). The

post hoc absolute fit method simulates individual choices on trial *t*, using estimated model parameters as well as observed choices and received feedback on trial 1 to *t*-1. Simulated mean error scores were calculated from 100 iterations. Results are presented in Figure 2. The post-hoc fit method revealed that all considered computational models were able to simulate the finding of higher error probabilities after negative feedback than after positive feedback. However, only the model incorporating category- and response-level RL was able to simulate behavioral avoidance, i.e., the modulation of perseveration error probabilities by response demands.
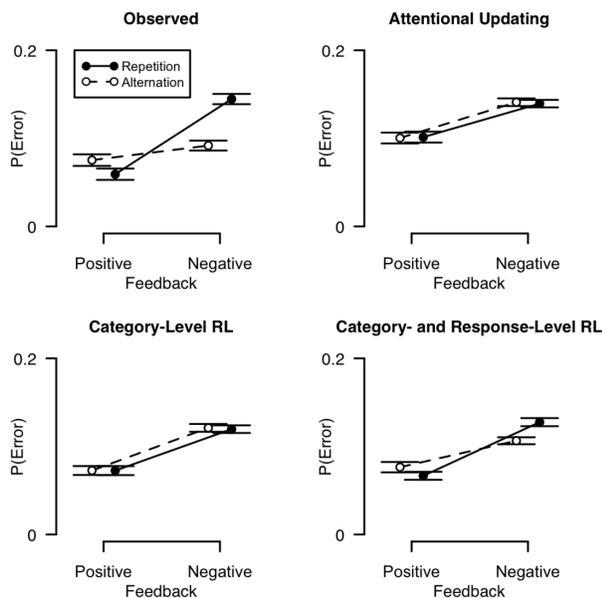

Figure 2: Observed (upper left plot) and simulated error probabilities (all other plots). Group mean error probabilities are presented. Error bars indicate +/- 1 standard error of the mean. Note that errors after positive and negative feedback correspond to set-loss and perseveration errors, respectively. RL = reinforcement learning; Repetition = response repetition demanded; Alternation = response alternation demanded.

Conclusion

Our results suggest that RL provides a generally better framework for understanding WCST performance than does the state-of-the-art attentional updating model (Bishara et al., 2010). Kopp et al. (under review) suggested that WCST performance should be conceptualized at multiple levels of learning and control. The present results of computational modeling support this hypothesis, as a computational model of category- and response-level RL outperformed a pure category-level RL model. However, the multi-level RL

model exceeds the pure category RL model with regard to model complexity (in terms of the number of free parameters, see Table 1). Future research conducted in our lab will address possibilities to reduce model complexity at comparable degrees of model fit. Overall, RL models in general, and the multi-level RL model in particular, seem to offer computational models for understanding latent cognitive processes that contribute to WCST performance.

References

Bishara, A. J., Kruschke, J. K., Stout, J. C., Bechara, A., McCabe, D. P., & Busemeyer, J. R. (2010). Sequential learning models for the Wisconsin Card Sort Task: Assessing processes in substance dependent individuals. *Journal of Mathematical Psychology*, *54*(1), 5–13.

Kopp, B., Steinke, A., Bertram, M., Skripuletz, T., & Lange, F. (under review). Multiple levels of control processes for Wisconsin Card Sorts: An observational study.

Palminteri, S., Wyart, V., & Koechlin, E. (2017). The importance of falsification in computational cognitive modeling. *Trends in Cognitive Sciences*, *21*(6), 425–433.

Schretlen, D. J. (2010). *Modified Wisconsin Card Sorting Test (M-WCST): Professional manual*. Lutz, Florida: Psychological Assessment Resources Inc.

Stan Development Team (2018). *RStan: the R interface to Stan*. R package version 2.18.2. http://mc-stan.org

Steingroever, H., Wetzels, R., & Wagenmakers, E. J. (2014). Absolute performance of reinforcement-learning models for the Iowa Gambling Task. *Decision*, *1*(3), 161–183.

Steinke, A., Lange, F., Seer, C., & Kopp, B. (2018). Toward a computational cognitive neuropsychology of Wisconsin card sorts: A showcase study in Parkinson's disease. *Computational Brain & Behavior*, *1*(2), 137–150.

Sutton, R. S., & Barto, A. G. (1998). *Reinforcement Learning: An Introduction*. Cambridge, MA: MIT Press.

Vehtari, A., Gelman, A., & Gabry, J. (2017). Practical Bayesian model evaluation using leave-one-out cross-validation and WAIC. *Statistics and Computing*, *27*(5), 1413–1432.