Computational Modeling of Human Associative Learning in a Complex Approach-Avoidance Learning Task

Franziska Usée (franziska.usee@uni-marburg.de)
Philipps-Universität Marburg, Department of Psychology
Marburg, 35032, Germany

Sebastian Schmidt (schmidbv@uni-marburg.de)
Philipps-Universität Marburg, Department of Psychology
Marburg, 35032, Germany

Christiane A. Melzig[1] (christiane.melzig@uni-marburg.de)
Philipps-Universität Marburg, Department of Psychology
Marburg, 35032, Germany
Center for Mind, Brain and Behavior (CMBB),
Philipps University Marburg and Justus Liebig University Giessen, Germany

Dirk Ostwald[1] (dirk.ostwald@ovgu.de)
Otto-von-Guericke-Universität Magdeburg, Department of Psychology
Magdeburg, 39106, Germany

[1] These authors contributed equally to this work.

Abstract

Despite its key role in the development, maintenance, and treatment of anxiety disorders, the detailed mechanisms of human avoidance learning remain elusive. Here, we report on a novel approach-avoidance learning paradigm that requires participants to learn associations between complex visual stimuli and combined positive and negative reward states. Using an agent-based behavioral modeling approach, we show that a Rescorla-Wagner learning-based agent with a prior expectation bias parameter best explains the learning behavior of 50 participants, paving the way for a more fine-grained computational understanding of the etiology of anxiety disorders.

Keywords Avoidance learning, clinical psychology, behavioral modeling, Rescorla-Wagner learning, visual foraging

Introduction

From the earliest age, humans rely on the ability to learn associations between differently valenced stimuli and rather complex environmental contexts, allowing for the avoidance of potentially (life-)threatening situations (Krypotos et al., 2015). In the absence of real dangers, however, learned avoidance responses no longer represent adaptive reactions and may in contrast result in substantial psychological impairments. In fact, excessive avoidance represents a core symptom of anxiety disorders, which are among the most common mental disorders worldwide (World Health Organization, 2017, 2004; American Psychiatric Association, 2013). Despite its key role in theories concerning the development, maintenance, and treatment of anxiety disorders, the detailed mechanisms underlying human avoidance learning remain elusive (Seriès, 2020).

Two factors are generally believed to hamper our understanding of avoidance learning to date. First, most studies have relied on the use of overly simplistic and unambiguous stimuli typically associated with only either positive (e.g., monetary gains) or negative rewards (e.g., electrical shocks). However, in real life, maladaptive avoidance behavior becomes most evident in situations of combined positive and negative reward states (e.g., Lissek et al., 2006; Beckers et al., 2013). Second, computational approaches that bear the potential to dissect the fine-grained psychological mechanisms of avoidance learning are only beginning to be employed in the fields of clinical psychology and psychiatry (Mkrtchian et al., 2017; Smith et al., 2021; Sharp & Eldar, 2019; Yamamori et al., 2023).

To contribute to the understanding of avoidance learning, we here report on a novel approach-avoidance learning paradigm that requires participants to learn associations between complex visual stimuli and combined positive and negative reward states while also actively engaging with the experimental environment. We show that a Rescorla-Wagner learning-based agent with a prior expectation bias parameter best explains the learning behavior of 50 participants on this paradigm. For formal and implementational details of this work, please refer to `https://osf.io/bjdse/`.

Behavioral Paradigm and Descriptive Analysis

Behavioral Paradigm On each trial of the experiment, participants performed a foraging task on one of the four visual search fields, each presented 10 times in a randomized order. On each visual search field, participants were requested to uncover coins by clicking on each of the 16 presented blue-colored circles while simultaneously being at risk of receiving electrical shocks. Each visual search field was characterized by a unique spatial pattern of blue-colored circles, representing a certain experimental condition. Specifically, within-subjects factors *profit* (low: two coins, high: six coins) and *punishment* (low: 0.1 probability of electrical shocks, high: 0.8 probability of electrical shocks) were manipulated in a 2 x 2-factorial manner (Figure 1a). Participants were asked to learn the associations between visual search field patterns and experimental conditions. As a behavioral read-out of this learning process, participants indicated the experimental condition associated with the visual search field after each trial (Figure 1b). In line with the reinforcement learning terminology, profits, punishments, and experimental conditions will hereinafter be referred to as *positive rewards*, *negative rewards*, and *reward states*, respectively (Sutton & Barto, 2018).

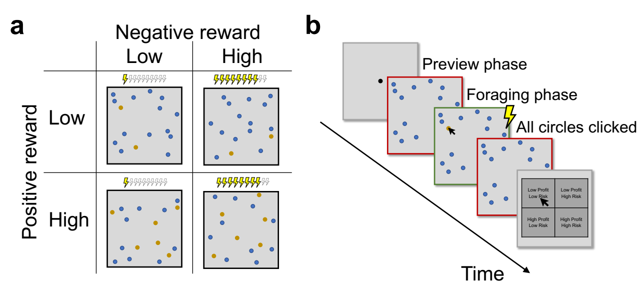

Figure 1: **a** Experimental paradigm. **b** Association of visual search fields with experimental conditions.

Descriptive Analysis Participants achieved an overall average accuracy of 67.95% (± 2.1 *SEM*) for identifying the reward states associated with the visual search fields (Figure 2a). The average accuracy increased over trials from around 38% to around 74%, indicating the successive acquisition of the correct associations (Figure 2b).

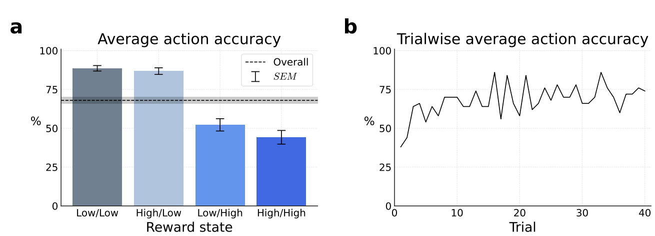

Figure 2: **a** Overall average (dashed line) and average reward state-specific accuracies across n = 50 participants and T = 40 trials. **b** Average trialwise accuracy of n = 50 participants.

Agent-Based Behavioral Modeling

To model the participants' trial-by-trial behavior, we formulated and validated nine agent models that implement different mechanisms for deciding on a given visual search field pattern-reward state association based on the currently available information only and/or successively updated expected reward value estimates. When combined with a behavioral model, the agent models emit the latent reward state indications as actions and their relative model plausibility can be evaluated in light of the participants' experimental data.

Agent Models We designed nine agent models that we denote by A0 to A8 (Table 1). Specifically, agent A0 served as a control agent, implementing a uniform random choice strategy. Its general structure is given by $\mathcal{A} := (D, \phi, p(\delta_t))$, where $(D, \phi, p(\delta_t))$ denotes the agent's decision model, comprising the decision set $D := \{0, 1, 2, 3\}$ of possible reward states, the agent's decision value function ϕ, and the agent's decision distribution $p(\delta_t)$. Agents A1 and A2 implement heuristic choice strategies that only consider the reward information directly observable on a given trial. Their general structure takes the form $\mathcal{A} := ((D, \phi, \delta), r_t)$, where (D, ϕ, δ) refers to the agent's decision model, including the agent's decision function δ, and r_t denotes the trialwise observed positive and negative reward. Agents A3 to A7 implement a Rescorla-Wagner learning-based decision strategy. Their general structure is given by $\mathcal{A} := ((M, \psi), (D, \phi, \delta), (o_t, r_t))$, where (M, ψ) denotes the agent's learning model, (D, ϕ, δ) denotes the agent's decision model, and (o_t, r_t) refer to the trialwise visual search field pattern and reward observations. Importantly, agents A3 to A7 differ in their learning model parameters and initial expected reward values estimates (Table 1). For example, agents A5 and A7 include a prior expectation bias parameter that accounts for the fact that participants were explicitly informed that they can only judge the risk of receiving a negative reward after several trials, potentially biasing their decisions for high negative reward states early in the experiment. Agent A8 implements a hybrid heuristic and Rescorla-Wagner learning-based decision strategy by learning an expected negative reward value estimate and deciding on the positive reward component based on the trialwise observation.

Behavioral Models The agent models were embedded in a statistical inference framework that transforms the agents' decisions into observable actions that correspond to the participants' indicated reward states over trials. For the control agent A0, the behavioral model corresponded to its stochastic decision model, while for agents A1 to A8 the behavioral model took the form of a softmax action distribution with post-decision noise parameter $\tau > 0$.

Parameter Estimation & Model Comparison Model parameters were estimated using a maximum likelihood approach. Specifically, a constrained Nelder-Mead optimization algorithm was applied with the parameter constraints set to $\eta, \eta_p, \eta_n, \tau \in [0.01, 2]$, and $\pi \in [-1, 1]$ (Virtanen et al., 2020).

Table 1: Agent model space

Agent	Learning model parameters	Decision model dependencies
A0	-	-
A1	-	r_t^p
A2	-	r_t^p, r_t^n
A3	Learning rate η	$\hat{\mu}^o_{r^p}$, $\hat{\mu}^o_{r^n}$
A4	Positive reward learning rate η_p, Negative reward learning rate η_n	$\hat{\mu}^o_{r^p}$, $\hat{\mu}^o_{r^n}$
A5	Learning rate η, Prior expectation bias π	$\hat{\mu}^o_{r^p}$, $\hat{\mu}^o_{r^n}$
A6	-	$\hat{\mu}^o_{r^p}$, $\hat{\mu}^o_{r^n}$
A7	Prior expectation bias π	$\hat{\mu}^o_{r^p}$, $\hat{\mu}^o_{r^n}$
A8	Negative reward learning rate η_n	r_t^p, $\hat{\mu}^o_{r^n}$

Notes. r_t^p and r_t^n denote the trialwise observed positive and negative reward, respectively. $\hat{\mu}^o_{r^p}$ and $\hat{\mu}^o_{r^n}$ refer to the expected positive and negative reward value estimates for a specific visual search field pattern observation o. Agent models A6 and A7 implement a trial-adaptive learning rate instead of a free learning rate η.

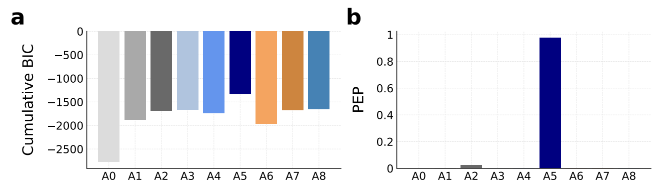

Figure 3: Model evaluation in light of the experimental data. A0 to A8 denote different agent models. Model comparison based on **a** BIC values and **b** PEP. Note that in line with Schwarz (1978), higher BIC values indicate higher model plausibility.

To evaluate the embedded agent models in light of the participants' experimental data, participant-specific Bayesian Information Criterion (BIC) values were computed according to Schwarz (1978) and summed across all participants. In line with the original formulation by Schwarz (1978), higher BIC values indicate higher model plausibility. Moreover, BIC values of all agents and data sets were subjected to a random-effects Bayesian model selection procedure using a Python implementation of the protected exceedance probability (PEP) approach (Stephan et al., 2009; Rigoux et al., 2014). Upon applying the agent-based behavioral modeling approach to the experimental data, we generated artificial data to assess the degree to which the models and parameter values were reliably recoverable. Please refer to `https://osf.io/bjdse/` for detailed information on these model validation results.

Model Comparison Figure 3 shows that, overall, agent A5 could account best for the observed behavioral data.

Conclusion

We show that a Rescorla-Wagner agent can explain human behavior in a complex approach-avoidance learning task better than random or heuristic agents. Given the task particularities, we find that a single fixed learning rate and a prior expectation bias parameter provide the most plausible agent configuration. Our work contributes groundwork for the computational dissection of approach-avoidance learning in complex decision scenarios and emphasizes its context adaptivity.

Acknowledgments

We thank Christoph Benke and Alexander Maximilian Norbert Vielhauer for help with experimental task design and implementation. We thank Clara Kessler for help with participant recruitment and data acquisition.

References

American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders: DSM-5* (5th ed.). American Psychiatric Publishing, Inc. doi: 10.1176/appi.books.9780890425596

Beckers, T., Krypotos, A.-M., Boddez, Y., Effting, M., & Kindt, M. (2013). What's wrong with fear conditioning? *Biological Psychology*, *92*(1), 90–96. doi: 10.1016/j.biopsycho.2011.12.015

Krypotos, A.-M., Effting, M., Kindt, M., & Beckers, T. (2015). Avoidance learning: a review of theoretical models and recent developments. *Frontiers in Behavioral Neuroscience*, *9*, 189. doi: 10.3389/fnbeh.2015.00189

Lissek, S., Pine, D. S., & Grillon, C. (2006). The strong situation: A potential impediment to studying the psychobiology and pharmacology of anxiety disorders. *Biological Psychology*, *72*(3), 265–270. doi: 10.1016/j.biopsycho.2005.11.004

Mkrtchian, A., Aylward, J., Dayan, P., Roiser, J. P., & Robinson, O. J. (2017). Modeling avoidance in mood and anxiety disorders using reinforcement learning. *Biological Psychiatry*, *82*(7), 532–539. doi: 10.1016/j.biopsych.2017.01.017

Rigoux, L., Stephan, K. E., Friston, K. J., & Daunizeau, J. (2014). Bayesian model selection for group studies – revisited. *Neuroimage*, *84*, 971–985.

Schwarz, G. (1978). Estimating the dimension of a model. *The annals of statistics*, 461–464.

Seriès, P. (2020). *Computational psychiatry: A primer*. MIT Press.

Sharp, P. B., & Eldar, E. (2019). Computational models of anxiety: Nascent efforts and future directions. *Current Directions in Psychological Science*, *28*(2), 170-176. doi: 10.1177/0963721418818441

Smith, R., Kirlic, N., Stewart, J. L., Touthang, J., Kuplicki, R., Khalsa, S. S., . . . Aupperle, R. L. (2021). Greater decision uncertainty characterizes a transdiagnostic patient sample during approach-avoidance conflict: a computational modelling approach. *Journal of Psychiatry and Neuroscience*, *46*(1), E74–E87. doi: 10.1503/jpn.200032

Stephan, K. E., Penny, W. D., Daunizeau, J., Moran, R. J., & Friston, K. J. (2009, July). Bayesian model selection for group studies. *NeuroImage*, *46*(4), 1004–1017. doi: 10.1016/j.neuroimage.2009.03.025

Sutton, R. S., & Barto, A. G. (2018). *Reinforcement learning: An introduction*. MIT press.

Virtanen, P., Gommers, R., Oliphant, T. E., Haberland, M., Reddy, T., Cournapeau, D., . . . SciPy 1.0 Contributors (2020). SciPy 1.0: Fundamental Algorithms for Scientific Computing in Python. *Nature Methods*, *17*, 261–272. doi: 10.1038/s41592-019-0686-2

World Health Organization. (2004). *ICD-10: International statistical classification of diseases and related health problems* (10th ed.). World Health Organization.

World Health Organization. (2017). *Depression and other common mental disorders: Global health estimates* (Tech. Rep.). Geneva: World Health Organization.

Yamamori, Y., Robinson, O. J., & Roiser, J. P. (2023, November). Approach-avoidance reinforcement learning as a translational and computational model of anxiety-related avoidance. *eLife*, *12*, RP87720. doi: 10.7554/eLife.87720