Cross-Context Value Dynamics: The Impact of Contextually Irrelevant Values on Choice Behavior

Neele K. Elbersgerd (neele.elbersgerd@uni-hamburg.de)
Universität Hamburg, Von-Melle-Park 5, 20146 Hamburg, Germany

Nir Moneta (moneta@mpib-berlin.mpg.de)
Max Planck Institute for Human Development
Lentzeallee 94, 14195 Berlin, Germany

Nicolas W. Schuck (nicolas.schuck@uni-hamburg.de)
Universität Hamburg, Von-Melle-Park 5, 20146 Hamburg, Germany

Abstract

In everyday decisions, the same choice option can yield different rewards depending on the context. Two coats will, for instance, have different utility depending on whether the goal is to be shielded from rain or to dress for a dinner. We reanalyzed behavioral data from a previous study, where participants switched between contexts with different object features predicting rewards: color in one context and motion in another (Moneta et al., 2023). Participants were explicitly cued in each trial which features to focus on, and outcome depended only on the cued context. Our analysis focused on how competing contexts influence choice and learning trial by trial. We identified two potential learning signals from irrelevant features, carrying information about the value expectation for (1) the chosen object's irrelevant feature (Object Prediction Error) and for (2) the best feature in the irrelevant context (Context Prediction Error). Combining reinforcement learning and general linear models, we found evidence for both learning signals influencing participants' behavior, i.e. that value updates were also guided by the difference between context-irrelevant values and outcomes. The strongest support was found for Object Prediction Error updates. Through examining how outcome-irrelevant information influences subjective value assignment, we deepen our understanding of how competing goals are processed and shape future choice behavior.

Keywords: Prediction Errors; Reinforcement Learning; Latent Value Updating; Context-Dependent Decision-Making

Introduction

Determining feature relevance in a dynamic, multidimensional environment is an essential challenge for goal-directed behavior (Niv et al., 2015). Choice options often have multiple features, with relevance varying according to the specific context at hand (Frömer, Dean Wolf, & Shenhav, 2019; Martino & Cortese, 2023). However, features initially deemed irrelevant may become relevant in the future, necessitating flexible retrieval and updating of context-dependent value information. Moneta, Garvert, Heekeren, and Schuck (2023) showed that when choosing between options, an alternative context and its values interfered with the relevant one, competing to guide behavior. Here, we reanalyze the behavioral data of their main task and ask whether such cross-context interference extends beyond choice deliberation to the processing of the outcome via contextually irrelevant value-updating signals, i.e prediction errors.

Results

Thirty-seven participants switched between contexts where reward was predicted by either stimulus color or motion direction. Participants were first trained to associate four color and four motion features with discrete rewards (balanced across participants). In each trial of the main task, they were explicitly cued which context to focus on and were instructed to choose the higher-valued of two simultaneously presented random dot kinematograms (Fig.1**a**). Outcome was dependent only on the features of the cued context

Moneta et al. (2023) found that while accuracy was overall high ($\mu = 0.9, \sigma = 0.05$, one-sample t-test against chance: $t_{(36)} = 50.27,\ p < .001$), participants were slower and less accurate when the irrelevant context implied a different action (as the example in Fig.1**a**). This effect was found to be further modulated by the best feature of the irrelevant context (EV_i, Fig.1**a**), showing that during choice-deliberation, participants fully processed the counterfactual choice according to the currently irrelevant context.

Prediction Errors of the Irrelevant Context. We hypothesized that irrelevant features also influence subjective value ($\widehat{sv}$) updating, which becomes evident in trial-by-trial fluctuations in choice behavior, despite deterministic outcomes on which participants were trained extensively. To test this, we propose two learning signals that might arise from irrelevant outcome expectations: First, an Object Prediction Error (OPE) captures the difference between the obtained reward and the value expectation of the chosen object's irrelevant feature, i.e. the irrelevant feature of the chosen cloud (irrelevant target: IT; Fig.1**a**, teal). Second, a Context Prediction Error (CPE) reflects the difference between the obtained reward and the reward participants would expect to receive were they in the other context (best irrelevant feature or expected value of the irrelevant context: EV_i; Fig.1**a**, orange).

Each PE can update values via two pathways: First, they

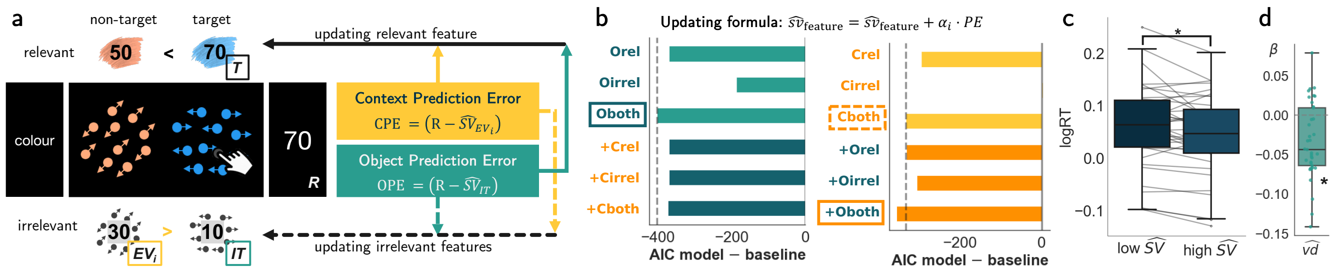

Figure 1: **a.** Schematic of a color trial (left): the cued color features (top) indicate to choose right, while the outcome-irrelevant motion features (bottom) imply otherwise. We hypothesized that after receiving reward (R), two contextually irrelevant prediction errors (right; CPE, orange; OPE, teal) can update latent subjective value estimations ($\widehat{sv}$). Each PE could send an update signal either to the relevant feature (solid arrows, T) or to the irrelevant features (dashed arrows, EV_i highest-valued feature of irrelevant context, or IT irrelevant feature of chosen option). **b.** The summed AIC showed that using RL-estimated $\widehat{sv}$ improved the fit using either PE updates in almost any pathway according to the schematic updating formula over a baseline GLM using objective values. Both OPE (left, light teal) and CPE (right, light orange) contribute most when updating both pathways simultaneously (*Oboth*, *Cboth*), more than when updating only the relevant (*Orel*, *Crel*, solid arrows in **a**) or only the irrelevant pathway (*Oirrel*, *Cirrel*, dashed arrows in **a**). Darker bars show model results where the relevant, irrelevant or both pathways of the other group were added to the best fitting model (in trials where IT and EV_i are different, see main text). **c.** Participants reacted faster in single-feature trials (only relevant features present) when their $\widehat{sv}_T$ were in the upper tertial versus in the lower (within-subject tertial split of trials ($p = .007$). **d.** Fitted regressor betas for the subjective value-difference ($\widehat{vd}$) in the GLM of *Oboth* ($p < .001$), which is introduced only through the OPE updates. Points are participants.

can update the choice's relevant feature (Fig.1**a**, solid arrows), implying that participants update the value of the chosen feature (relevant target feature: T) based on what they would have expected to get in the irrelevant –but competing– context (for the same cloud: OPE, for the best feature: CPE). For instance, for the same chosen cloud in Fig.1**a**, participants might hold an expectation tied to its horizontal motion ($\widehat{sv}_{IT}$, here 10 p). Thus, the $\widehat{sv}$ of blue may be updated also based on the expectation of 10 p (solid OPE arrow). The next time participants see blue, its $\widehat{sv}$ might be slightly reduced due to the OPE update before. Second, the PEs could update the $\widehat{sv}$ of the irrelevant features themselves, even though outcome depended solely on the relevant features (OPE: updates the same choice's irrelevant feature IT, CPE: updates the best feature of the irrelevant context EV_i; Fig1**a**, dashed arrows). We employed reinforcement learning (RL) models (Sutton & Barto, 2018) to estimate $\widehat{sv}$, which were then used as predictors for a GLM explaining reaction time. We compared our models to a baseline GLM which assumes stable value estimates, i.e. true values with no cross-trials influence from the irrelevant context (best model of Moneta et al., 2023).

Strongest support for Object Prediction Error. Model comparison support that RT is best explained by accounts that incorporate value updating of the relevant and irrelevant features through OPE & CPE signals (Fig.1**b**, light colors, AIC differences to baseline from winning models $\Delta\text{AIC}_{Oboth} = -403.2$, $\Delta\text{AIC}_{Cboth} = -331.0$) Note that in some trials the features EV_i and IT are the same (if the blue cloud in Fig.1**a** had diagonal motion), which means that OPE and CPE are the same. To disentangle their contributions we refitted the winning model and added the updating pathways of the respective other PE *only* in trials where OPE and CPE are different. We found that only adding OPE-updates to the winning CPE model improved the fit, but not when adding CPE-updates to the winning OPE model (Fig.1**b**, dark colors). Hence, the dominating features behind irrelevant value updates are likely those associated with the same choice (reflected in the OPE). Future work is planned to further investigate the unique contributions of each PE.

Effects of subjective values on RT. Next, we further assessed the contribution of subjective value estimates derived from the winning OPE model. We tested this on single-feature trials, where only the relevant features were presented (these were interleaved with the dual-feature trials of Fig.1), using a within-subject tertial split on the estimated $\widehat{sv}_T$. Indeed, average RTs in single-trials were slower for choices when the current $\widehat{sv}_T$ was lower compared to when it was higher ($t_{(36)} = 2.85$, $p = .007$, Fig.1**c**). Since there is no information from an alternative context in these trials, this effect provides evidence for irrelevant influences from previous trials. Lastly, by design the value difference of the relevant features was fixed at 20 p (and outcomes were deterministic). However, the $\widehat{sv}$ of features fluctuated over time. This introduced a *subjective* value difference between relevant features ($\widehat{vd}$) which entered the GLM as a new regressor. We found that this contributed significantly to explaining RT, such that participants reacted faster for higher $\widehat{vd}$ (two-sided 1-sample $t_{(36)} = -4.24$, $p < .001$ Fig.1**d**), consistent with prior research reporting faster decisions with larger value differences between options (Hunt et al., 2012).

Conclusion

Our results provide evidence that learning signals arising from contextually irrelevant features lead to latent subjective value fluctuations, and consequently influence future behavior. We found the strongest evidence for the Object Prediction Error interfering with the value-updating process post-outcome. This indicates a within-object spillover, making participants more prone to assigning credit to the irrelevant feature of the cloud they deliberately clicked on. Expanding our understanding of how irrelevant information interferes with goal-directed value computations hints towards how context separation in complex environments is performed in humans and might inform further research on its implementation in the brain.

Acknowledgments

NWS is funded by a Starting Grant from the European Union (ERC-2019-17 StG REPLAY-852669) the Federal Ministry of Education and Research (BMBF) and the Free and Hanseatic City of Hamburg under the Excellence Strategy of the Federal Government and the Länder. NM was partially funded by the Einstein Center for Neuroscience (ECN) Berlin.

References

Frömer, R., Dean Wolf, C. K., & Shenhav, A. (2019). Goal congruency dominates reward value in accounting for behavioral and neural correlates of value-based decision-making. *Nature Communications*, *10*(1), 4926. doi: 10.1038/s41467-019-12931-x

Hunt, L. T., Kolling, N., Soltani, A., Woolrich, M. W., Rushworth, M. F. S., & Behrens, T. E. J. (2012). Mechanisms underlying cortical activity during value-guided choice. *Nature Neuroscience*, *15*(3), 470–476. doi: 10.1038/nn.3017

Martino, B. D., & Cortese, A. (2023). Goals, usefulness and abstraction in value-based choice. *Trends in Cognitive Sciences*, *27*(1), 65–80. doi: 10.1016/j.tics.2022.11.001

Moneta, N., Garvert, M. M., Heekeren, H. R., & Schuck, N. W. (2023). Task state representations in vmPFC mediate relevant and irrelevant value signals and their behavioral influence. *Nature Communications*, *14*(1), 1–21. doi: 10.1038/s41467-023-38709-w

Niv, Y., Daniel, R., Geana, A., Gershman, S. J., Leong, Y. C., Radulescu, A., & Wilson, R. C. (2015). Reinforcement Learning in Multidimensional Environments Relies on Attention Mechanisms. *Journal of Neuroscience*, *35*(21), 8145–8157. doi: 10.1523/JNEUROSCI.2978-14.2015

Sutton, R. S., & Barto, A. G. (2018). *Reinforcement learning: An introduction* (Second edition ed.). The MIT Press.