Decision-Making with Naturalistic Options

Can Demircan[1] (can.demircan@tuebingen.mpg.de), Leonardo Pettini[2,3], Tankred Saanum[1], Marcel Binz[1], Blazej Baczkowski[2], Christian Doeller[2,3], Mona Garvert[2], & Eric Schulz[1]

[1]Max Planck Institute for Biological Cybernetics, Tübingen, Germany

[2]Max Planck Institute for Cognitive and Brain Sciences, Leipzig, Germany

[3]Max Planck School of Cognition, Leipzig, Germany

Abstract

How do humans generalise to make better decisions? Previous work has investigated this question using reward-guided decision-making tasks with low-dimensional and artificial stimuli. In this article, we extended this work by presenting participants with a naturalistic decision-making task, in which options were images of real-world objects and the underlying reward function was based on one of their latent dimensions. Even though participants received no explicit instruction about object features, they quickly learned to do the task and generalised to unseen objects without problems. To understand how they accomplish this, we tested a range of computational models and found that human behaviour is overall best explained by a linear model but that their strategies change during the experiment. Lastly, we show that combining pixel-based representations extracted from convolutional neural networks with the original latent dimensions further improves our models. Taken together, our study offers new insights into how humans make decisions in more naturalistic settings.

Keywords: naturalistic decision-making, function-learning, heuristics, generalisation

Introduction

Imagine that you have tried and enjoyed a specific brand of chocolate, but the next time you go to the store you cannot find the same brand anymore. Early theories have characterized the process of learning as forming stimulus-reward associations (Rescorla, 1972). However, if you were only forming such associations, you would not be able to find a good alternative in the example above. In reality, on the other hand, generalising what you like about that specific chocolate to pick another option that you would enjoy is trivial.

How do people accomplish this seemingly complex challenge so effortlessly? The study of human generalisation is a central theme in many areas of cognitive science, such as associative learning (Shanks & Darby, 1998), function learning (Schulz et al., 2017), and decision-making (Stojic et al., 2015; Schulz et al., 2020; Saanum et al., 2021). Together, these studies illustrate that humans rely on feature-based representations to make generalisations in situations like the chocolate example above. Previous research suggests that these generalisation capabilities are explained by various computational models, including rule- and similarity-based theories (Shanks & Darby, 1998; Lucas et al., 2015).

While these studies have been important for understanding how humans generalise, they also have important shortcomings. First, the stimuli used in these studies typically have only a few underlying dimensions. This does not reflect the high-dimensional features of the objects we deal with in real life, which makes it unclear whether the proposed theories of learning can scale up to real-world environments. In addition, features of stimuli are often explicitly provided and clearly separable from each other. In real life, this is not the case and an object can be arbitrarily broken down into different sets of features. For example, you can represent a bar of chocolate by how sweet it is, its price, its environmental impact, any combination thereof, or with completely different features.

To address these shortcomings, we conducted a two-alternative forced-choice study, in which stimuli were unique images of real-world objects (Hebart et al., 2019). The underlying reward function was determined by one of the latent features of these objects, which we extracted using a combination of computational modeling and human behavioural data (Hebart et al., 2020). With this design, we address the following questions:

1. Do people make successful generalisations about rewards in high-dimensional feature spaces?

2. Do insights from previously conducted studies with low-dimensional, artificial stimuli transfer to these more naturalistic domains?

3. What kind of representations do people use to solve our task?

We find that people discover the underlying relationship within a few dozen trials. To gain insights into their decision-making processes, we carried out several model-based analyses. We first compared participant behaviour to various models of decision-making, including a linear model, a Gaussian Processes regression model, and two heuristic decision-making strategies. This analysis revealed that while the linear model explains participant behaviour the best overall, humans seem to change their strategies as they progress through the task. We then probed which feature-based representations our participants utilised. We find that while performance is best explained by the originally extracted latent features, it can be further improved by using additional pixel-based representations extracted from a pre-trained convolutional neural network. Taken together, our results provide an initial step towards understanding human decision-making in naturalistic domains.

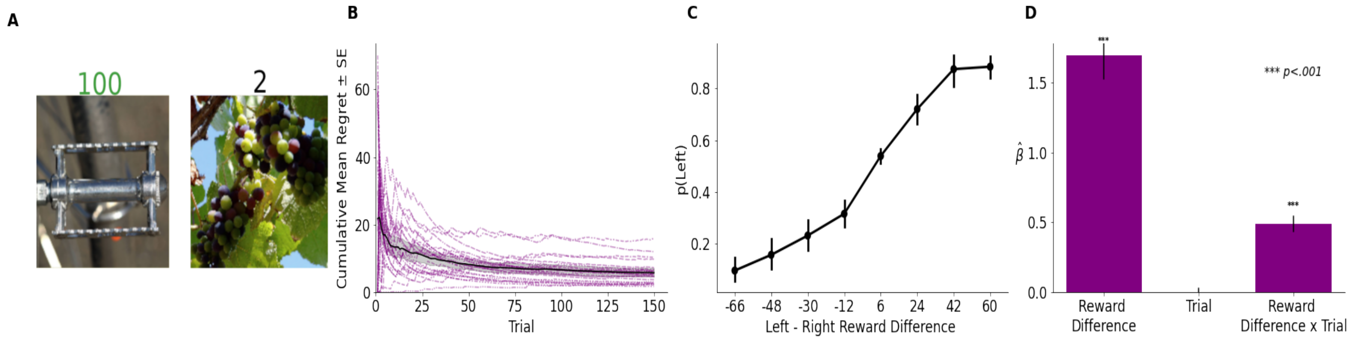

Figure 1: Design & Behavioural Analyses. **A)** An example of a trial outcome. The bike pedal, which is highly rewarding because it is mostly metallic, is chosen over the grapes. **B)** Participants performance over trials. The shaded black line represents the mean and its standard error, and the purple lines show individual participants' learning curves. **C)** Probability of choosing the option on the left for different reward differences between the two options with standard error bars. **D)** Standardised coefficients with standard error bars from the mixed-effects logistic regression model predicting participant choice as a function of trial number and reward differences between the two options.

Methods

Participants

25 healthy participants (7 female, $M_{\text{age}} = 24.55$, $SD_{\text{age}} = 3.86$) were recruited through Prolific. All participants had a Prolific Score of 89 or above. Participants were given 7€ per hour as a base rate, and a bonus of up to 10€ was offered depending on the performance. Participants took 15 minutes to complete the task on average.

Design

Participants were asked to complete a two-alternative forced-choice task with 150 consecutive trials. At the beginning of the experiment, they were instructed that each image was associated with a reward in a non-random way, and they were asked to choose the images that gave the most rewards.

In each trial, participants were first presented with a fixation cross (displayed for 500 ms), followed by a pair of images. They had unlimited time to choose one of the two images by using the left and right arrow buttons on the keyboard. They were then shown the reward associated for both the chosen and the unchosen option, in green and white respectively (displayed for 2000 ms). The order of trials was randomised across participants.

Stimuli & Reward Function

300 images were sampled from the THINGS database, which is a systematically curated image dataset of real-world objects (Hebart et al., 2019). Hebart and colleagues have trained an image embedding model on similarity judgements of humans on the THINGS database in order to extract latent dimensions with continuous loadings that capture participants' mental representations of these objects (Hebart et al., 2020). They extracted 49 latent dimensions, which were validated to be semantically meaningful by further behavioural testing. Latent dimensions include, for example, how metallic, food-related, or colourful an object is (see the original paper for the entire list of latent dimensions).

We normalised the loadings of the first latent dimension, which describes how metallic an object is, as follows to compute our reward function:

$$r_n = \frac{w_n - \min(\mathbf{w})}{\max(\mathbf{w}) - \min(\mathbf{w})} \times 100$$

where r_n is the reward for stimulus n and $\mathbf{w} = [w_1, \ldots, w_N]$ is the vector of first dimension loadings for the sampled stimuli. Crucially, the existence of these latent dimensions was unknown to the participants. We illustrate an example trial from our experiment in Figure 1A.

Behavioural Analyses

We first computed several descriptive statistics to analyse human behaviour on our task. Participants learned to select the options with higher rewards over a few dozen trials. Figure 1B shows that the cumulative mean regret, which is the average of the difference between the best option and the chosen option, decreases over trials. We furthermore plotted participant choices as a function of the reward difference between options in Figure 1C. The resulting function takes a sigmoid shape, indicating participants can use the reward differences between the options to guide their decisions.

To formally test whether participants could learn the task, we used a mixed-effects logistic regression model. We predicted participant choice in each trial as a function of the reward difference between the two options and the trial number. Both predictors were standardized and modelled as fixed and random effects. A greater reward difference between the right and the left options led participants to choose the right option more frequently ($\hat{\beta} = 1.69$, 95% CI $[1.37, 2.01]$, $p < .001$). While the trial number had no significant effect on participant choice ($\hat{\beta} = -0.02$, 95% CI $[-.07, .11]$, $p = .71$), there was an interaction effect between trial number and the reward difference ($\hat{\beta} = .49$, 95% CI $[.39, .59]$, $p < .001$), indicating that participants get better over time at using reward differences between the options when making their decisions (see Figure 1D).

Model-Based Analyses

The previous analyses confirmed that participants can solve our task and that they get better at it over time. In the upcom-

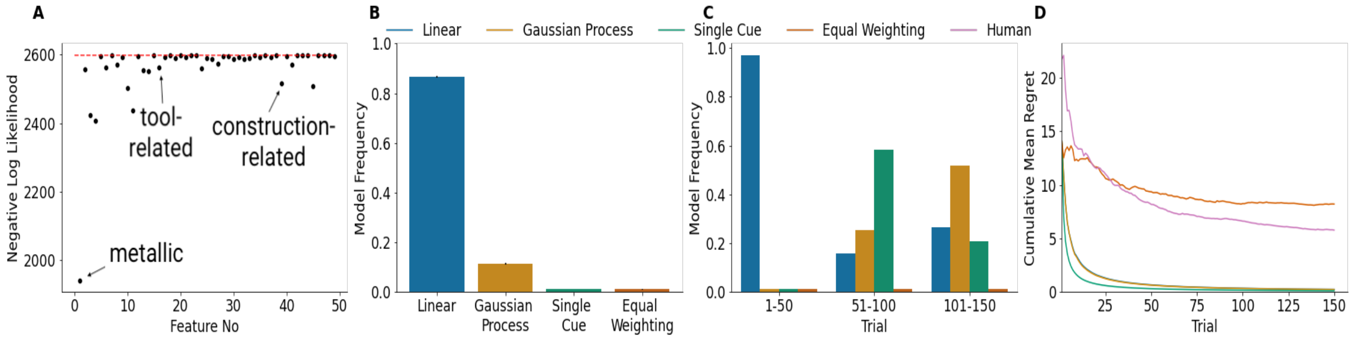

Figure 2: Model-Based Analyses. **A)** Model fits for mixed-effects models predicting participant choice with loadings in different latent dimensions. Red dashed line shows chance level performance. **B)** Model comparison of computational models for all trials. Frequencies plotted with standard error bars. **C)** Model comparison of computational models for the beginning (1-50), the middle (51-100), and the ending (101-150) trials. **D)** Performance of computational models, where models follow a greedy policy.

ing section, we complement these behavioural results with additional model-based analyses in order to gain insights into how they accomplish this.

Which feature predicts behaviour the best?

We started by testing which of the latent dimensions predicted participant choice behaviour the best. To do so, we ran 49 mixed-effects logistic regression models, one for each latent dimension. Each model had scaled differences of a given dimension's loadings between the left and the right options both as fixed and random effects. As can be seen in Figure 2A, the model that used the first latent dimension, which is the one the reward function is based on, was the best performing model (negative log-likelihood = 1940.24). The same figure also shows that there were other latent dimensions, such as tool relatedness and construction relatedness, that predicted participant behaviour above chance level. The above-chance level performance of these models can be explained by the fact that they are correlated with how metallic an object is.

Computational Models

In order to provide a computational account for how participants learned to do our task, we assessed the degree to which their choice behaviour was described by different decision-making models prevalent in the literature. All models presented in this section use the latent dimensions identified by Hebart et al. (2020) as features and are updated after each trial using data from both the chosen and the unchosen option.

The first model under consideration is a *linear model*. This model assumes that rewards are a weighted linear combination of all features:

$$\mathbf{r} = \mathbf{X}\beta + \varepsilon \qquad \varepsilon \sim \mathcal{N}(0, \sigma^2)$$

where the rows of $\mathbf{X}$ are trials and the columns are different features, β are the weights, $\mathbf{r}$ is the reward of optionS, and ε is the noise term. Previously, this class of models has provided good fits in decision-making tasks, both when the reward function was a linear function of single (Niv et al., 2015) and multiple features (Speekenbrink & Shanks, 2010; Stojic et al., 2015). In our case, it was implemented as a Bayesian linear regression model (Bishop, 2006). The prior over the weights was defined as a spherical Gaussian distribution with the scaling parameter λ. The reward prediction for a new stimulus $\mathbf{x}$ was obtained using the mean of the posterior predictive distribution:

$$\hat{r}(\mathbf{x}) = \left(\sigma^{-2}\left(\sigma^{-2}\mathbf{X}^T\mathbf{X} + \lambda\mathbf{I}\right)^{-1}\mathbf{X}^T\mathbf{y}\right)^T \mathbf{x}$$

Gaussian Process (GP) regression models (Schulz et al., 2018) offer a competing explanation for how people could solve our task. In previous work, these models have been successfully used to understand human generalisation across a range of reward-guided decision-making studies (Schulz et al., 2020). A GP defines a multivariate normal probability distribution over functions:

$$f \sim \mathcal{N}(m(\mathbf{x}), k(\mathbf{x}, \mathbf{x}'))$$

where $m(\mathbf{x})$ is the mean function, which we set to 0, and $k(\mathbf{x}, \mathbf{x}')$ is the kernel (also called the covariance function), which defines prior assumptions about how similar two feature vectors $\mathbf{x}$ and $\mathbf{x}'$ are. Here, we employ a Radial Basis Function (RBF) kernel, which represents the similarity between two feature vectors as an exponentially decaying function of their squared Euclidean distance:

$$k(\mathbf{x}, \mathbf{x}') = exp\left(\frac{||\mathbf{x} - \mathbf{x}'||^2}{2\ell^2}\right)$$

where the parameter ℓ controls the rate of decay of similarity. We have picked the RBF kernel as it has previously been shown to explain human behaviour well in decision-making tasks with linear reward functions (Stojić et al., 2020), despite it not capturing the underlying linear task structure. Using GP regression, reward predictions for a new stimulus $\mathbf{x}$ can be made by:

$$\hat{r}(\mathbf{x}) = \mathbf{k}^T \left(\mathbf{K} + \sigma^2\mathbf{I}\right)^{-1}\mathbf{y}$$

where $\mathbf{k}$ is the covariance matrix between the previously observed stimuli and the new stimulus and $\mathbf{K}$ is the covariance matrix between all previously observed stimuli.

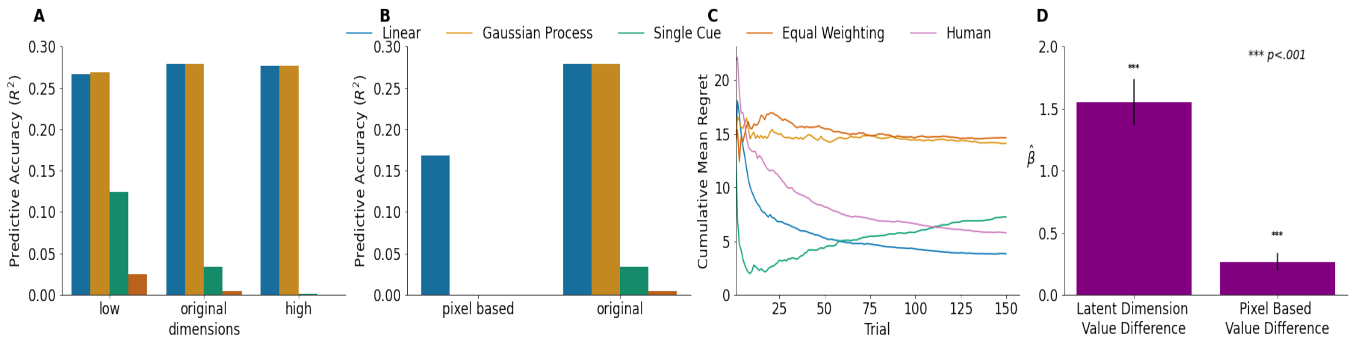

Figure 3: Representational Analyses. **A)** Predictive accuracy of computational models trained with different latent representations. **B)** Predictive accuracy of computational models trained with pixel-based representations and the original latent dimensions. **C)** Performance of computational models trained with pixel-based representations, where models follow a greedy policy. **D)** Standardised coefficients with standard error bars from the mixed-effects logistic regression model predicting participant choice with reward estimates obtained by the linear model trained with pixel-based representations and those obtained by the same model trained on the latent dimensions.

The two previously outlined models take all available features to a varying degree into account when making predictions. It has been argued in previous research that this style of decision-making is too computationally expensive and that people rely on simpler heuristic strategies instead (Gigerenzer & Gaissmaier, 2011). We therefore also considered two common heuristics in our model comparison: a *single cue* model and an *equal weighting* model. The single cue model only uses the single best feature to make decisions. This type of heuristic has been shown to be successful at explaining human behaviour both in real-world and lab settings (Gigerenzer & Goldstein, 1999; Gigerenzer & Gaissmaier, 2011). We assume that the identity of the best feature is unknown, and instead maintain one single cue model for each feature dimension. Each of these models is implemented as a one-dimensional Bayesian linear regression model. To obtain reward estimates, we make predictions based only on the best performing model up until that point, i.e., the one with the highest log-likelihood (Binz et al., 2022). The equal weighting model, on the other hand, does not distinguish between different features, and instead learns a single weight for all of them (Gigerenzer & Gaissmaier, 2011; Dawes & Corrigan, 1974). We implemented this form of heuristic decision-making through a Bayesian linear regression model with a single feature that is obtained by summing up the original features (Binz et al., 2022).

Model Comparison

For our model comparison, we first computed the reward estimates for each computational model as described above. We then ran a separate mixed-effects logistic regression for each model, where we used the difference between the reward estimates of two options as fixed and random regressors to predict participant choice behaviour. We did a leave-one-trial-out cross-validation for each of these models to obtain cross-validated log-likelihoods (Garvert et al., 2021). To compare models, we used these log-likelihoods in a model-frequency analysis (Stephan et al., 2009; Rigoux et al., 2014), which is a Bayesian procedure that estimates the prevalence of a model within the participant population. We report model frequencies (MF), which measure how common a model is in that population, and their exceedance probability (XP), which is the posterior probability that the frequency of a given model is larger than all the other models in that population. We additionally report pseudo-R^2 scores (McFadden, 1974) obtained from the cross-validated log-likelihoods:

$$R^2 = 1 - \frac{\mathcal{L}(M)}{\mathcal{L}(\text{Random})}$$

$\mathcal{L}(M)$ is the log-likelihood of a given model and $\mathcal{L}(\text{Random})$ is the log-likelihood of a random model. While $R^2 = 1$ shows M is infinitely more accurate than chance, $R^2 = 0$ indicates it is a model performing at chance-level.

We found that the linear model is the most frequent model within our population (MF $= .87$, XP $> .99$, $R^2 = .2792$) as illustrated in Figure 2B. While much less frequent, the GP explained participant behaviour to a similar degree (MF $= .11$, $R^2 = .2786$). The single cue and equal weighting models performed poorly in predicting participant behaviour ($R^2 = .03, .005$ respectively).

We also hypothesised that people might rely on different strategies in different stages of the task. For this reason, we ran separate model frequency analyses for different parts of the study (Figure 2C). While the linear model outperforms the other models in the first 50 trials (MF $= .97$, XP$> .99$, $R^2 = .18$), the single cue model was the most frequent model (MF $= .58$, XP $= .97$, $R^2 = .32$) for the second 50 trials. Interestingly, in the last 50 trials, the GP was the most frequent model (MF $= .52$, XP $= .91$, $R^2 = .36$). These results indicate that participants switch strategies as they progress through the task. They start the task with a linear-additive strategy, then switch to make decisions only based on the best cue, and, in the end, use an examplar-based strategy. For a forward simulation of the computational models, see Figure 2D.

Interim Discussion

How do these results relate to the outcomes of previous decision-making studies with low-dimensional, artificial stimuli? An almost universal conclusion from these studies is that people seem to employ linear-additive strategies unless they are explicitly encouraged to use simpler decision-making heuristics instead (Binz et al., 2022). This finding aligns with our main result that people are overall best described by the linear model.

Rieskamp & Otto (2006) furthermore found that participants in their study had initial preferences for linear-additive strategies, but then switched to single cue heuristics during later stages. This hypothesis was confirmed in several other studies (Gluck et al., 2002; Mata et al., 2007). We observed an analogous pattern in our study, with participants having an initial preference for linear-additive strategies, followed by a switch to single cue heuristics.

Finally, Juslin et al. (2003) argued that "people have an inclination to abstract explicit representations whenever possible ..., with exemplar memory acting as a backup in tasks in which explicit representations of cue–criterion relations cannot be abstracted or in which behaviour has become automatic". It would be plausible that behaviour has become automatic by the end of our study, which would, in turn, explain the late emergence of GPs (an exemplar-based model) as the winning hypothesis. Taking everything into consideration, our analysis suggests that key results from decision-making studies with low-dimensional, artificial stimuli also transfer to more naturalistic settings.

Representational Analyses

In the above section, we used the latent dimensions extracted by Hebart et al. (2020) to train our computational models. However, these are not the only representations that can be used to solve our task. Humans may, for example, use more granular or more compressed representations. In addition to testing the dimensionality of the representations, we tested whether our models can be improved by incorporating pixel-based representations extracted from convolutional neural networks.

Learning with Different Latent Dimensions

We first re-trained the model of Hebart et al. (2020) to extract a different number of latent dimensions. The model is trained to predict human responses on an odd-one-out task with three objects. To do so, the model learns weights shared across all the objects. Importantly, the model is penalised for the number of non-zero weights that it learns, and the extent of this penalty is controlled by a hyperparameter in the model's loss function. By changing this hyperparameter, we extracted a low (14) and a high (82) number of latent dimensions from the objects. The test accuracy for the newly trained models on the odd-one-out task was comparable to that of the original model, indicating that the latent dimensions described the objects well. We then replicated the previously discussed model comparison procedure with the newly extracted latent dimensions to test which representations predict human choice behaviour the best.

For the models that were trained with low number of latent dimensions, the GP regression model performed the best (MF $= .58$, XP $= .90$, $R^2 = .27$), followed closely by the linear model ($R^2 = .27$). The single cue and equal weighting models again performed worse comparably ($R^2 = .12$ and $R^2 = .02$ respectively). Out of the models that were trained with the high number of latent dimensions, the linear model performed the best (MF $= .49$, XP $= .50$, $R^2 = .28$), with the GP regression model again performing similarly ($R^2 = .28$). Both the single cue and equal weighting models predicted around chance level ($R^2 < .001$). Overall, the best performing model was the linear model trained with the original latent dimensions, indicating that the original dimensions extracted by Hebart and colleagues captured the representations used by participants to complete the task best (Figure 3A). Another interesting finding here is that the single cue model got better at predicting participant behaviour as the feature space got more compressed, hinting at the possibility that a sufficiently small feature space combined with this model may be able to compete with our currently winning models.

Learning with Pixel-Based Representations

Recently, there has been a surge of interest and success in using end-to-end representations to model human behaviour (see Battleday et al. (2021) for a review). Following this direction, we consider the possibility that participants use such representations to solve our task. Because deep convolutional neural networks have proven to be promising models of the human visual system over the last decade (Yamins & DiCarlo, 2016), we decided to use one to obtain such representations. We passed the images through a pre-trained ResNet18 (He et al., 2016) and extracted the activity pattern of the penultimate layer neurons. We then used the resulting 512 features as inputs to our models.

Models trained with the pixel-based representations can perform the task above chance level as shown in Figure 3C. The linear model predicted human choice behaviour the best compared to the other pixel-based models (MF $= .97$, XP $>$ $.99$, $R^2 = .17$), which is however worse than the linear model trained with the original latent dimensions. All other models trained with pixel-based representations perform around chance level (Figure 3B).

Even though the pixel-based computational models do not outperform the models trained with the original latent dimensions, it is possible that the reward estimates obtained by training on the pixel-based representations capture some features of the objects that are used by the participants but that are not captured by the original latent dimensions. To test this hypothesis, we used a mixed-effects logistic regression model, where we used reward estimates of the linear model trained with the original latent dimensions and reward estimates of the linear model trained with pixel-based representations as predictors. The estimates obtained from both the original dimensions ($\hat{\beta} = 1.55$, 95% CI [1.37, 1.73], $p < .001$) and the pixel-based representations ($\hat{\beta} = .27$, 95% CI [.21, .33], $p < .001$) were significant predictors (Figure 3D). This model performed significantly better in predicting human choice behaviour compared to the mixed-effects model that only used the linear model's estimates coming from the original latent dimensions ($\chi(3) = 17.9$, $p < .001$, $R^2 = .28$). These results provide support for our hypothesis that humans may use representations that are not captured by the symbolic feature space of latent dimensions but can be extracted using end-to-end methods.

General Discussion

How do people learn to make good decisions in settings with high-dimensional options? We have studied this question in a two-alternative forced-choice task with naturalistic stimuli. Importantly, participants in our study were not explicitly guided about the existence of the high-dimensional features of objects but nevertheless got better at the task over time simply by reward guidance. This marks a novel contribution to the existing literature on reward-guided decision-making as previous work has largely focused on designs where options were not composed of high-dimensional features and where features were explicitly signalled to the participants. The fact that none of the stimuli appear more than once also shows that participants did not simply learn stimulus-reward associations but that they learned features about the stimuli, allowing them to generalise effectively.

We furthermore tested various models used in the decision-making literature to provide a computational explanation of human generalisation in high-dimensional feature spaces. When trained on the latent dimensions extracted from the stimuli, a linear model provided the best fit to the human data. Interestingly, however, comparing models at different points in the experiment revealed that participants changed their strategy as they progressed through the task. Participants initially learned a linear function until they were certain about which aspect was relevant for obtaining rewards and switched to a strategy that only cares about the single reward-relevant dimension once they were certain. In the later stages of the experiment, they switched to a GP regression model, suggesting that their behaviour has become automated.

While our models could predict participant behaviour well, the features they received do not necessarily reflect the representations used by the participant. To investigate the possibility that humans use different representations while doing our task, we first compared models trained on the original latent dimensions with more compressed and more granular latent dimensions. We showed that the original latent dimensions that capture the structure of the task provided the best fit to the human choice data. It is worth pointing out that using more compressed dimensions led the single cue model to perform better compared to the other conditions. This is interesting because the individual cues this model uses to make decisions do not directly correspond to the reward function of our task. It is possible that humans represent the objects with even fewer dimensions than we have tested here and that a single cue model trained with more compressed representations can explain human choice behaviour even better.

While the cognitive sciences have mostly used symbolic representations when modeling human behaviour, more recent work has shown that using distributed representations obtained by deep neural networks can provide a better account of human behaviour in similarity judgement (Peterson et al., 2018) and categorisation tasks (Battleday et al., 2020). To test whether such representations can be useful in explaining human behaviour in our reward-guided decision-making task, we trained our models on pixel-based representations obtained from a pre-trained convolutional neural network. The linear model again provided the best fit for the human data with this form of representation. More interestingly reward estimates obtained from this linear model were better at predicting participants' choice when combined with the reward estimates coming from the linear model trained on the original latent dimensions, compared to using the reward estimates coming from the latter alone. In the future, trying different neural networks architectures and fine-tuning them for the task at hand can extend our work and provide better predictions of human behaviour.

Lastly, we have assumed that participants receive reward information of both the chosen and unchosen options in our task. We have made this choice to simplify our experimental design. However, this assumption also removes the need for exploratory choices. Future work could lift this restriction and reveal reward information for only the chosen option. In turn, this will allow us to adapt our paradigm to study the relationship between exploration and generalisation (Wu et al., 2018) in a more naturalistic setting.

In summary, our work provides three important insights into human decision-making. First, we established that people could learn to pick more rewarding options by generalising their knowledge about object features in a naturalistic setting. Second, we showed that humans employ different strategies at different stages and that their behaviour can be explained by similar models utilized in decision-making tasks with low-dimensional, artificial stimuli. Lastly, distributed representations obtained from neural networks capture aspects of how humans represented the objects in our task, that were not captured by the original latent dimensions. These results offer some of the first insights into how humans make decisions with naturalistic options and pave the way for further research in this domain.

References

Battleday, R. M., Peterson, J. C., & Griffiths, T. L. (2020, October). Capturing human categorization of natural images by combining deep networks and cognitive models. *Nature Communications*, *11*(1). Retrieved 2022-01-26, from https://www.nature.com/articles/s41467-020-18946-z

Battleday, R. M., Peterson, J. C., & Griffiths, T. L. (2021). From convolutional neural networks to models of higher-level cognition (and back again). *Annals of the New York Academy of Sciences*, *1505*(1), 55–78. Retrieved 2022-01-17, from https://onlinelibrary.wiley.com/doi/abs/10.1111/nyas.14593 (_eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1111/nyas.14593)

Binz, M., Gershman, S. J., Schulz, E., & Endres, D. (2022, January). Heuristics from bounded meta-learned inference. *Psychological Review*.

Bishop, C. M. (2006). *Pattern recognition and machine learning (information science and statistics)*. Berlin, Heidelberg: Springer-Verlag.

Dawes, R. M., & Corrigan, B. (1974). Linear models in decision making. *Psychological Bulletin*, *81*(2), 95–106. (Place: US Publisher: American Psychological Association)

Garvert, M. M., Saanum, T., Schulz, E., Schuck, N. W., & Doeller, C. F. (2021, October). *Hippocampal spatio-temporal cognitive maps adaptively guide reward generalization* (Tech. Rep.). Retrieved 2022-01-21, from https://www.biorxiv.org/content/10.1101/2021.10.22.465012v1 (Company: Cold Spring Harbor Laboratory Distributor: Cold Spring Harbor Laboratory Label: Cold Spring Harbor Laboratory Section: New Results Type: article)

Gigerenzer, G., & Gaissmaier, W. (2011). Heuristic Decision Making. *Annual Review of Psychology*, *62*(1), 451–482. Retrieved 2022-01-17, from https://doi.org/10.1146/annurev-psych-120709-145346 (_eprint: https://doi.org/10.1146/annurev-psych-120709-145346)

Gigerenzer, G., & Goldstein, D. G. (1999). Betting on one good reason: The take the best heuristic. In *Simple heuristics that make us smart* (pp. 75–95). New York, NY, US: Oxford University Press.

Gluck, M. A., Shohamy, D., & Myers, C. (2002). How do people solve the "weather prediction" task?: Individual variability in strategies for probabilistic category learning. *Learning & Memory*, *9*(6), 408–418.

He, K., Zhang, X., Ren, S., & Sun, J. (2016, June). Deep Residual Learning for Image Recognition. In *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)* (pp. 770–778). (ISSN: 1063-6919)

Hebart, M. N., Dickter, A. H., Kidder, A., Kwok, W. Y., Corriveau, A., Wicklin, C. V., & Baker, C. I. (2019, October). THINGS: A database of 1,854 object concepts and more than 26,000 naturalistic object images. *PLOS ONE*, *14*(10), e0223792. Retrieved 2022-01-11, from https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0223792 (Publisher: Public Library of Science)

Hebart, M. N., Zheng, C. Y., Pereira, F., & Baker, C. I. (2020, November). Revealing the multidimensional mental representations of natural objects underlying human similarity judgements. *Nature Human Behaviour*, *4*(11), 1173–1185. Retrieved 2022-01-11, from https://www.nature.com/articles/s41562-020-00951-3

Juslin, P., Olsson, H., & Olsson, A.-C. (2003). Exemplar effects in categorization and multiple-cue judgment. *Journal of Experimental Psychology: General*, *132*(1), 133.

Lucas, C. G., Griffiths, T. L., Williams, J. J., & Kalish, M. L. (2015). A rational model of function learning. *Psychonomic bulletin & review*, *22*(5), 1193–1215.

Mata, R., Schooler, L. J., & Rieskamp, J. (2007). The aging decision maker: cognitive aging and the adaptive selection of decision strategies. *Psychology and aging*, *22*(4), 796.

McFadden, D. (1974). Conditional logit analysis of qualitative choice behavior. *Frontiers in econometrics*.

Niv, Y., Daniel, R., Geana, A., Gershman, S. J., Leong, Y. C., Radulescu, A., & Wilson, R. C. (2015, May). Reinforcement Learning in Multidimensional Environments Relies on Attention Mechanisms. *The Journal of Neuroscience*, *35*(21), 8145–8157. Retrieved 2022-01-11, from https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4444538/

Peterson, J. C., Abbott, J. T., & Griffiths, T. L. (2018, November). Evaluating (and Improving) the Correspondence Between Deep Neural Networks and Human Representations. *Cognitive Science*, *42*(8), 2648–2669.

Rescorla, R. (1972). A theory of Pavlovian conditioning : Variations in the effectiveness of reinforcement and nonreinforcement..

Rieskamp, J., & Otto, P. E. (2006). Ssl: a theory of how people learn to select strategies. *Journal of Experimental Psychology: General*, *135*(2), 207.

Rigoux, L., Stephan, K. E., Friston, K. J., & Daunizeau, J. (2014, January). Bayesian model selection for group studies — Revisited. *NeuroImage*, *84*, 971–985. Retrieved 2022-01-24, from https://www.sciencedirect.com/science/article/pii/S1053811913009300

Saanum, T., Schulz, E., & Speekenbrink, M. (2021). Compositional generalization in multi-armed bandits.

Schulz, E., Franklin, N. T., & Gershman, S. J. (2020). Finding structure in multi-armed bandits. *Cognitive Psychology*, *119*. (Place: Netherlands Publisher: Elsevier Science)

Schulz, E., Speekenbrink, M., & Krause, A. (2018, August). A tutorial on Gaussian process regression: Modelling, exploring, and exploiting functions. *Journal of Mathematical Psychology*, *85*, 1–16. Retrieved 2022-01-15, from https://www.sciencedirect.com/science/article/pii/S0022249617302158

Schulz, E., Tenenbaum, J. B., Duvenaud, D., Speekenbrink, M., & Gershman, S. J. (2017). Compositional inductive biases in function learning. *Cognitive psychology*, *99*, 44–79.

Shanks, D. R., & Darby, R. J. (1998). Feature-and rule-based generalization in human associative learning. *Journal of Experimental Psychology: Animal Behavior Processes*, *24*(4), 405.

Speekenbrink, M., & Shanks, D. R. (2010, May). Learning in a changing environment. *Journal of Experimental Psychology. General*, *139*(2), 266–298.

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.

Stojic, H., Analytis, P., & Speekenbrink, M. (2015, July). Human behavior in contextual multi-armed bandit problems..

Stojić, H., Schulz, E., P. Analytis, P., & Speekenbrink, M. (2020, October). It's new, but is it good? How generalization and uncertainty guide the exploration of novel options. *Journal of Experimental Psychology: General*, *149*(10), 1878–1907. Retrieved 2022-01-11, from http://doi.apa.org/getdoi.cfm?doi=10.1037/xge0000749

Wu, C. M., Schulz, E., Speekenbrink, M., Nelson, J. D., & Meder, B. (2018, December). Generalization guides human exploration in vast decision spaces. *Nature Human Behaviour*, *2*(12), 915–924. Retrieved from https://doi.org/10.1038/s41562-018-0467-4

Yamins, D. L. K., & DiCarlo, J. J. (2016, March). Using goal-driven deep learning models to understand sensory cortex. *Nature Neuroscience*, *19*(3), 356–365. Retrieved 2022-01-17, from https://www.nature.com/articles/nn.4244