Social Value Learning Shifts Conceptual Representations of Faces

Ariana M. Familiar (afam@sas.upenn.edu)
Sharon L. Thompson-Schill (sschill@psych.upenn.edu)
Department of Psychology
University of Pennsylvania, Philadelphia, PA 19104 USA

Abstract

Values drive our behavioral choices. Ample research has explored the cognitive and neural underpinnings of value-based computations related to decision-making. However, behaviorally relevant values that we associate with real-world objects are often not monetary. For instance, social values associated with specific people are crucial for social behaviors and interactions. Moreover, understanding and attributing social values allows for proper evaluations of potential interactions with others, and can lead to more beneficial social behaviors and relationships. Learning social values has been shown to recruit the same systems as reward values, however how they become associated with specific people remains to be established. The present study examined social value learning of other people using naturalistic face images. We found that before learning, distances between the faces in conceptual similarity spaces were organized corresponding to their perceptual similarity. However, after learning, faces were shifted in a manner that reflected similarity of their associated social values (generosity). Furthermore, distances were positively correlated with a post-learning index of preference to interact with a person in a future cooperative game. In other words, learned social values of the faces seemed to influence their representations in conceptual space, and such representational changes were related to propensities in future behavior.

Keywords: value learning, face perception, social decision-making

Introduction

Humans possess the remarkable ability to learn about hundreds of other people and objects. One important piece of information about encountered items is their value, which comes to be associated through positive and negative experiences. Remembering and comparing values of alternate choices is crucial to making decisions. While values of items can be monetary, they are often more abstract in nature. Arguably one of the most biologically relevant values are social values that are associated with other people, such as personality traits that inform social behavior. Computing and encoding social values from interactions with others is essential for social behavior across different environmental situations.

Understanding social information is an important skill, as social interactions are prevalent in behavior. Learning social values has implications for both the individual as well as the wider community. For instance, understanding the personality traits of other people that are relevant to social behavior allow for knowledge of social norms and realistic evaluations of the consequences of future interactions with other people. This in turn helps to guide choices in social decisions, which can lead to more beneficial social behaviors and relationships. Thus, understanding how social values are learned and associated with other people is important for promoting social well being at an individual level, as well as creating environments conducive to social learning.

Former studies on social value learning have largely focused on the computations involved in value learning and value-related decision-making. This work has found evidence that suggests social value information is learned via similar mechanisms as reward-based learning (e.g., Behrens, Hunt, Woolrich, & Rushworth, 2008). Moreover, a recent study found that when making decisions in a social game, social value (generosity) information was weighed more heavily than reward value (point) information, even though these sources of information were orthogonal to one another and it would have been more beneficial to focus on reward values (Hackel, Doll, & Amodio, 2015). Additionally, social value learning has been shown to involve the same neural systems as reward value learning. For instance, areas of the prefrontal cortex and ventral striatum are recruited during social valuation and social exchange (Behrens et al., 2008; Hackel, Doll & Amodio, 2015; Izuma, Saito, & Sadato, 2008; for review, see Lee, 2008). Together this work has shed light on the behavioral and neural implications of social value processing.

Although such research has elucidated the underlying computations and neural systems involved in social value learning, it has yet to be studied how values become associated with the specific people to which they belong. Importantly, interacting with another person involves both recognizing the person based on their physical appearance, namely their face, as well as remembering information associated with them from former social interactions. From an encoding perspective, a person-level representation that incorporates both perceptual (facial identity) and social information would facilitate efficient recognition and decision performance.

Some studies have shown that learning information associated with faces modulates perceptual responses to those faces. Electrophysiological studies have shown how neural responses to faces are modulated by learned information. For instance, electroencephalography (EEG) studies have established a specific event-related potential (ERP) response in occipito-temporal areas to face stimuli that likely corresponds to identity processing, the N170 repetition effect, however this effect is only found for unfamiliar, and not familiar, faces. Two studies have found evidence that this familiarity difference is due to person-specific information associated with the familiar faces, suggesting that perceptual processing of faces is changed with associations of semantic (e.g. biographical information) or in-depth social information (Heisz & Shedden, 2009; Herzmann & Sommer, 2010). While such studies have shown perceptual processing of facial identity to be modulated by learned information about a person, it remains to be established how underlying perceptual representations of faces are modulated by such learned information. In other words, behavioral and neuroimaging studies have indicated differences in face processing depending on acquired knowledge about a person, however a crucial gap remains in our understanding of how different information about a person becomes associated.

While face processing and social value learning have been widely studied in isolation, the mechanisms allowing for association of social value information with faces remain unclear. In the present study, we examine whether learned social values (generosity) influence conceptual representations of facial identity. We found that distances between faces in a conceptual similarity space were related to social values after learning, and such representational changes were related to a measure of preference for future interactions. The results of this study suggest that conceptual representations of facial identities are integrated with learned social value information, which becomes associated after value learning, and this conceptual re-organization is related to future expectations of social behavior.

Materials & Methods

Participants

Twenty participants (five males, ages 18-36) were recruited from the University of Pennsylvania and received course credit or financial compensation for their time.

Stimuli

Nine face images were chosen from the Psychological Image Collection at Stirling database (pics.stir.ac.uk), such that they depicted people who were of the same race and gender and of similar ages. These images were cropped with an oval mask and matched in mean luminance using the SHINE Toolbox (Willenbockel et al., 2010).

Value Learning Task

A modified version of the task used in Hackel, Doll, & Amodio (2015) was conducted online using PsyToolkit (Stoet, 2017). Participants were recruited under the impression that they would be randomly assigned to either a social choice or a social learning role, however all were assigned to a social learning role. After this assignment, participants were told that during training, they would learn about the actions of other 'players' assigned to the social choice role, who allocated a pool of points between themselves and a future player (the participant). On a given trial, participants chose to 'play' with one of two players, presented side by side on a computer screen. To make a choice, participants pressed one of two keys to indicate the player on the left (F-key) or the player on the right (J-key); they had 2 seconds to respond before the next trial commenced (inter-trial-interval of 3 seconds). Upon choosing, they were presented with feedback for 3 seconds about how many points that player gave them and the point pool the player was allocated on that trial (Figure 1). If no choice was made, no feedback was presented.

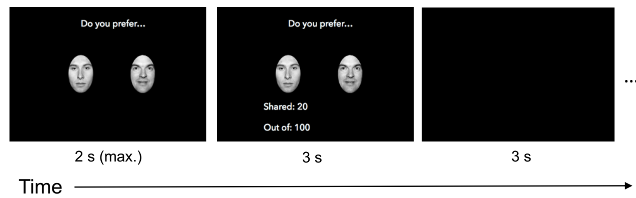


Figure 1. Example of one trial in the learning task. In this case, the participant would have chosen the left player.

Participants were instructed to maximize their accrued points, as the total number of points they earn amounted to a bonus. Moreover, they were told that some players were given more points on average to allocate, and some gave more points on average, so they would have to learn about both sources of information over the course of training in order to maximize their total points.

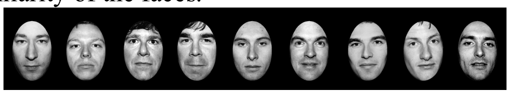
On average, players shared 20, 50, or 80% of their point pool and were assigned 15, 45, or 75 points. On a given trial, noise was added to these values by randomly selecting a value from a normal distribution centered on zero (standard deviation of 5; constraint of minimum 1 for generosity noise and 2 for reward noise) and adding it to the average value for that face. Point pools for that trial were calculated by dividing the rounded point value by the generosity value for that trial. Participants completed an hour of training each day (288 trials per session), for four days, and were shown their accrued number of points at the end of each session.

Social (generosity) and reward (point) values were assigned to the faces as follows. Pairwise differences in perceptual similarity were calculated (described below), and combinations of social and reward values were assigned to the faces such that they were orthogonal to perceptual similarity, and orthogonal to one another (Fig. 2).

Behavioral Similarity Measures

Conceptual Similarity A free sorting task was used to quantify conceptual similarity spaces. Participants were shown the nine face images on a white background and were instructed to organize the images in a spatial manner that reflected their similarity. The closer together the images were in space, the more similar the people depicted were, and the farther apart, the more dissimilar. There were no time constraints on the completion of the task. Participants performed the task once before and once after completing the value learning task.

Perceptual Similarity A separate group of 20 participants performed the same free sorting task for course credit, however they were specifically instructed to organize the images in a manner that reflected the perceptual, or physical, similarity of the faces.



Generosity	20	20	20	50	50	50	80	80	80
Points	15	45	75	15	45	75	15	45	75
Point Pool	75	225	375	30	90	150	19	56	94

Figure 2: Experimental parameters for learning task. Average generosity (percentage of point pool shared), point values, and point pools assigned to each of the nine face images. Generosity values, point values, and perceptual similarity were orthogonal to one another.

Post-Learning Ratings

After the last learning session, participants completed the following ratings, conducted using Qualtrics.

Social Preference Participants were told they may be invited back for an additional study involving a cooperative non-economic task, and were asked to indicate their preference to be paired with the other players on that task. For each player, participants rated on a scale of 1-7 how much they preferred to be paired with that player (1 = not at all; 7 = definitely yes).

Social Value Ranking Participants were instructed to rank the players in order of their overall generosity in the social choice role.

Point Value Ranking Participants were instructed to rank the players in order of their overall points in the social choice role (i.e. how many points the players were allocated on average).

Results

Accuracy

A correct response on a given trial was choosing the player with the higher average point value, and if they had the same average point value then choosing the face with the higher average generosity. Accuracy (percent correct) was computed across trials within a session. If no response was made during a trial, that trial was not included in the analysis (2% excluded trials, across participants and sessions).

Overall, accuracy improved over the four days of learning, as indicated by an increase in the average accuracy across days (Day 1: $M = 61\%$, $SEM = 2\%$; Day 4: $M = 74\%$, $SEM = 1\%$; Fig. 3). A one-way repeated-measures analysis of variance (ANOVA) revealed a significant effect of session ($F(1.9, 36.1) = 23.84$, $p < 0.001$; Greenhouse-Geisser corrected). A paired two-tailed t-test between accuracies on first and last days of learning confirmed a significant increase in accuracy ($t(19) = 5.56$; $p < 0.001$), and an additional t-test showed accuracies on the last day to be significantly greater than chance-performance (50%; $t(19) = 7.46$ $p < 0.001$). This established that our value learning manipulation was successful.

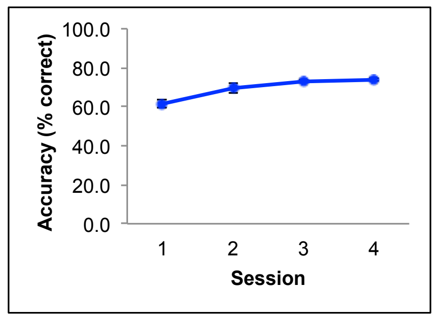


Figure 3. Average accuracy across participants for value learning task. Error bars represent +/- SEM.

Conceptual Space Organization

First, we examined which sources of information were related to the organization of faces in similarity space. Distances between each pair of faces were calculated (pixels), separately for the pre- and post-learning spaces, for each participant separately. Next, differences between the post-learning rankings of each pair of players were

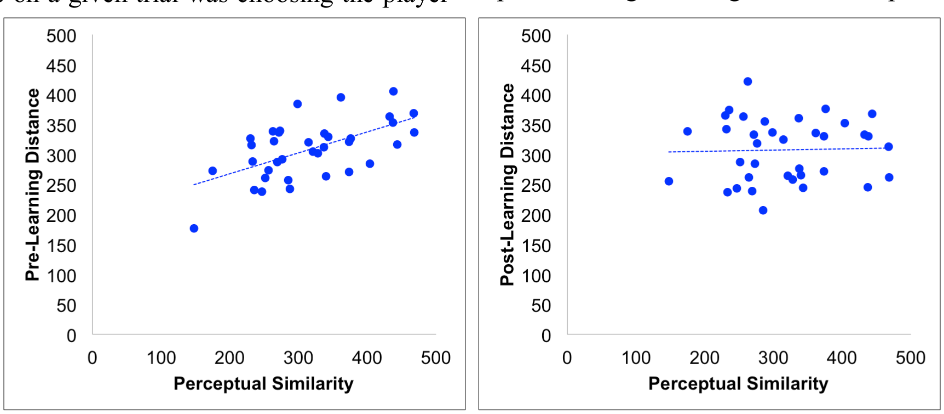


Figure 4. Correlations between perceptual similarity distances and pre-learning distances (left), and post-learning distances (right). Pairwise distances between faces were related to perceptual similarity (distances derived from a separate group of participants) before value learning, but not after. One data point represents one pair of faces.

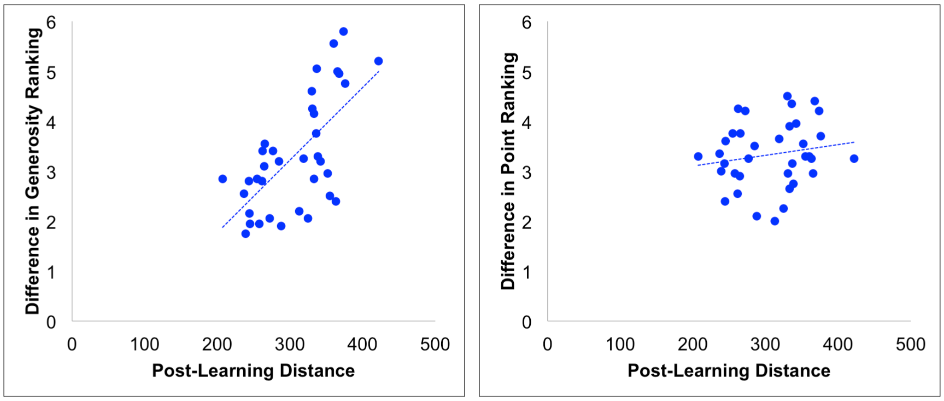


Figure 5. Correlations between post-learning distances and generosity ranking differences (left), and point ranking differences (right). After value learning, pairwise distances between faces were related to pairwise differences in generosity ranking, but not point ranking. One data point represents one pair of faces.

calculated for each ranking (social value and reward value) separately. For instance, if one player was listed as the third most generous player, and another was the seventh most generous player, their pairwise generosity difference would be four. Resulting values were averaged across participants, for each pair of faces separately, and correlated with the average pairwise distances, separately for the pre- and post-learning similarity spaces.

Perceptual Similarity Perceptual similarity was defined as the average pairwise distances between faces in the perceptual similarity spaces across participants (from a separate group), for each pair of faces. Pre-learning distances correlated positively with perceptual similarity distances (Fig. 4; Pearson R = 0.59, p < 0.001). After learning, distances between faces did not correlate with perceptual similarity (R = 0.03, p = 0.853), and the difference between these correlations was significant (Z = 2.94, p = 0.003). This suggests that perceptual similarity drives initial organization of the facial identities in conceptual space, however this influence is attenuated after more information about the person is learned.

Social Value Similarity Differences in post-learning social value ranking of face pairs correlated positively with the pairwise post-learning distances (Fig. 5; R = 0.65, p < 0.001), but not with the pre-learning distances (R = 0.18, p = 0.282), and the difference between these correlations was significant (Z = 2.62, p = 0.009), indicating that after learning, social value similarity influenced the organization of faces in conceptual similarity space.

Reward Value Similarity Reward similarity, or differences in the post-learning point value rankings of face pairs, did not correlate with the post-learning distances (Fig. 5; R = 0.17, p = 0.322), or the pre-learning distances (R = 0.06, p = 0.746). This indicates that reward value similarity is not likely related to distances between faces in similarity space.

Together, these results confirm that before learning, the spatial organization was related to perceptual similarity, while after value learning, social values were related to the organization of faces in similarity space.

Relationship Between Conceptual Similarity and Prospective Social Behavior

In order to compare the post-learning measure of future social preferences with similarity space organization, we calculated pairwise differences between the preference ratings and then correlated these results with the pairwise distances between faces. This analysis showed that before learning, spatial organization was not related to preferences

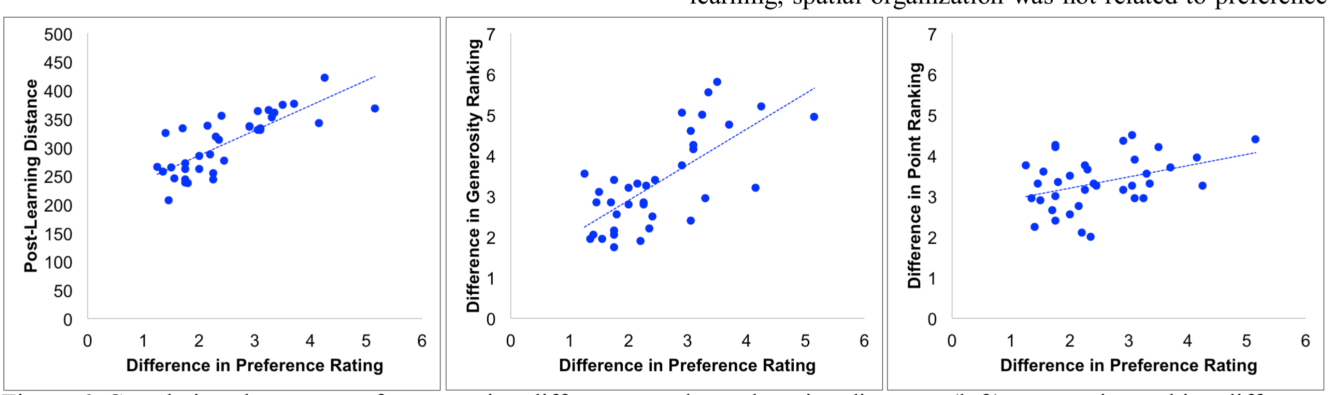


Figure 6. Correlations between preference rating differences and post-learning distances (left), generosity ranking differences (middle), and point ranking differences (right). After value learning, pairwise differences in future preference ratings were related to pairwise distances between faces and pairwise differences in generosity and point ranking. One data points represents one pair of faces.

(R = 0.13, p = 0.466), however after learning the preference ratings correlated positively with distance in similarity space (Fig. 6; R = 0.79, $p < 0.001$). Moreover, preference rating differences correlated with generosity ranking (R = 0.71, $p < 0.001$) and point ranking (R = 0.39, p = 0.018) differences. Preference ratings were not related to perceptual similarity (R = 0.03, p = 0.865).

These results confirm that after value learning, both social values and reward values were related to propensities in future social behavior, even in a non-economic setting. As both value types seemed to influence preferences for social interactions, we sought to further test whether there were individual differences between participants in tendencies to use either or both sources of value information.

Relationship Between Individual Choices and Prospective Social Behavior

To examine individual differences in reliance on social and/or reward values, we looked for a relationship between choices made in the learning task and the future preference ratings. Specifically, we examined whether participants tended to use different ratios of value information, and whether these tendencies generalized across tasks. To derive a measure of individual sensitivity to generosity and point information in the preference ratings, we performed the following for each participant separately. Following Hackel, Doll, & Amodio (2015), we computed generosity sensitivity as the difference between average ratings for high generosity players and average ratings for low generosity players. Then, we computed reward sensitivity as the difference between the average ratings for high reward targets and low reward targets. Differential sensitivity was calculated as the reward sensitivity subtracted by the generosity sensitivity.

Next, we examined individual differences in choices made during the last session of the learning task. The number of trials in which a participant chose the player with a higher average generosity value was divided by the number of trials in which the participant chose the player with the higher average reward value, and then log transformed. This resulted in a choice ratio that quantified a participant's tendency to choose players based on generosity or point information.

The differential sensitivity and choice ratio measures correlated positively (Fig. 7; R = 0.64, p = 0.002), indicating that the extent to which a participant used social and/or reward value information in making their choices during the learning task was related to their sensitivity to social and reward values in their future preference ratings. This finding confirms that people differentially weighed learned social and reward values to guide decisions, and did so consistently across different social contexts.

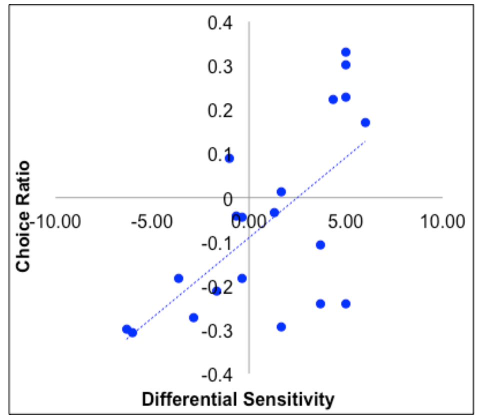


Figure 7. Sensitivity to generosity (+) or point (-) information in the future preference ratings (differential sensitivity) as a function of reliance on generosity (+) or point (-) information in choices on the last day of learning (choice ratio). Participants differentially weighed learned generosity and point values, but did so consistently across tasks. One point represents one participant.

Discussion

In this study, we examined how learned social values influence conceptual representations of facial identities. While previous studies had established that knowledge about specific people influenced neural processing of their faces, it remained to be established how exactly this learned information influences underlying perceptual representations, and how such changes are related to behavior.

We had participants learn social (generosity) and reward (point) values associated with different people, and perform a free sorting task before and after learning in order to quantify conceptual similarity. We found that before learning, organization of the similarity spaces was related to the perceptual similarity of the faces, such that faces that were more perceptually similar were closer together. After value learning, social values influenced the spatial organization, such that faces of more similar generosities were closer together. These results show how learned social information about a person is integrated with representations of their facial identity.

It could be argued that the similarity space of the faces has not been reorganized, but rather that participants are accessing another representation of the social value of the faces in response to the demands of the experimental task. We believe this is unlikely, as if our results were primarily driven by features of the task, post-learning spaces would incorporate both social and reward values, especially given that people were sensitive to both sources of information in choices on the last day of learning and in the post-learning preference ratings. However, the possibility that the experimental context is influencing similarity judgements cannot be ruled out based on these data.

Importantly, the re-organization of similarity space was related to prospective social behavior. Specifically, distances between faces correlated positively with future preferences of interacting with each person, as well as with perceived generosity. In other words, faces with higher social values (more generous people) were more preferred than those with lower values, and this was reflected in the spatial organization of post-learning similarity spaces. This suggests that such spatial re-organization operates in a manner that can guide expectations of future behavior based on acquired knowledge, at least in a social context.

Moreover, participants' tendencies to rely on generosity and/or point information during the later stages of value learning were related to their sensitivity to this information in future social preferences. For example, participants who chose players with higher social values more often than those with higher reward values during the learning task preferred to be paired with the high social value, low reward players more than the low social value, high reward players. This shows how individual differences in weighing of social and reward value information to make choices in an economic task generalizes to prospective social behavior in a non-economic setting.

Presumably, the integration of social value and facial identity information allows for rapid and successful recognition of faces, and can be used to guide future social behaviors and decisions. It remains to be established whether associating information with people, such as social values, influences recognition performance or sensitivity to such information on other social decision tasks.

Another open question is where the neural correlates of such representational changes are located. Facial information is processed in the fusiform face area (FFA), occipital face area (OFA), and the posterior superior temporal sulcus (pSTS). It is possible that activity in the FFA and OFA related to facial identity processing (Grill-Spector, Knouf, & Kanwisher, 2004) is modulated by learned information about a person. That being said, areas of the ventral anterior temporal lobe (ATL) have been found to contain neurons that encode paired associations between facial identity and abstract semantic knowledge (Eifuku, Nakata, Sugimori, Ono, & Tamura, 2010), thus the ATL may integrate facial identity with person-specific conceptual knowledge (Olsen, McCoy, Klobusicky, & Ross, 2013). No work to our knowledge has examined representational changes of faces in such areas as a result of learned associations.

While the present study establishes an influence of learned social information associated with different facial identities on conceptual representations, more work is needed to further establish the behavioral consequences of such conceptual space changes, as well as the underlying cognitive and neural mechanisms of such representational re-organization.

Acknowledgments

This research was supported by a grant from the US National Institutes of Health awarded to Dr. Thompson-Schill (R01 DC015359).

References

Behrens, T. E., Hunt, L. T., Woolrich, M. W., & Rushworth, M. F. (2008). Associative learning of social value. *Nature*, *456*(7219), 245-249.

Collins, J. A., & Olson, I. R. (2014). Knowledge is power: How conceptual knowledge transforms visual cognition. *Psychonomic Bulletin & Review*, *21*(4), 843-860.

De Baene, W., Ons, B., Wagemans, J., & Vogels, R. (2008). Effects of category learning on the stimulus selectivity of macaque inferior temporal neurons. *Learning & Memory*, *15*(9), 717-727.

Eifuku, S., Nakata, R., Sugimori, M., Ono, T., & Tamura, R. (2010). Neural correlates of associative face memory in the anterior inferior temporal cortex of monkeys. *Journal of Neuroscience*, *30*(45), 15085-15096.

Folstein, J. R., Palmeri, T. J., & Gauthier, I. (2012). Category learning increases discriminability of relevant object dimensions in visual cortex. *Cerebral Cortex*, *23*(4), 814-823.

Grill-Spector, K., Knouf, N., & Kanwisher, N. (2004). The fusiform face area subserves face perception, not generic within-category identification. *Nature Neuroscience*, *7*(5), 555-562.

Hackel, L. M., Doll, B. B., & Amodio, D. M. (2015). Instrumental learning of traits versus rewards: dissociable neural correlates and effects on choice. *Nature Neuroscience*, *18*(9), 1233-1235.

Heisz, J. J., & Shedden, J. M. (2009). Semantic learning modifies perceptual face processing. *Journal of Cognitive Neuroscience*, *21*(6), 1127-1134.

Herzmann, G., & Sommer, W. (2010). Effects of previous experience and associated knowledge on retrieval processes of faces: An ERP investigation of newly learned faces. *Brain research*, *1356*, 54-72.

Izuma, K., Saito, D. N., & Sadato, N. (2008). Processing of social and monetary rewards in the human striatum. *Neuron*, *58*(2), 284-294.

Lee, D. (2008). Game theory and neural basis of social decision making. *Nature Neuroscience*, *11*(4), 404-409.

Marini, F., Marzi, T., & Viggiano, M. P. (2011). "Wanted!" The effects of reward on face recognition: electrophysiological correlates. *Cognitive, Affective, & Behavioral Neuroscience*, *11*(4), 627-643.

Olson, I. R., McCoy, D., Klobusicky, E., & Ross, L. A. (2013). Social cognition and the anterior temporal lobes: a review and theoretical framework. *Social cognitive and affective neuroscience*, *8*(2), 123-133.

Stoet, G. (2017). PsyToolkit: A novel web-based method for running online questionnaires and reaction-time experiments. *Teaching of Psychology*, *44*(1), 24-31.

Willenbockel, V., Sadr, J., Fiset, D., Horne, G. O., Gosselin, F., & Tanaka, J. W. (2010). Controlling low-level image properties: The SHINE toolbox. *Behavior Research Methods*, 42(3), 671-684.