Running Head: Self and Learning

Sticky Me:

Self-Relevance Slows Reinforcement Learning

Marius Golubickis[1] and C. Neil Macrae[2]

[1]School of Psychology, University of Plymouth, Plymouth, England, UK
[2]School of Psychology, University of Aberdeen, Aberdeen, Scotland, UK

Address Correspondence to:

Marius Golubickis
School of Psychology
University of Plymouth
Drake Circus
Plymouth PL4 8AA
England, UK

Email: marius.golubickis@plymouth.ac.uk

Abstract

A prominent facet of social-cognitive functioning is that self-relevant information is prioritized in perception, attention, and memory. What is not yet understood, however, is whether similar effects arise during learning. In particular, compared to other people (e.g., best friend), is information about the self acquired more rapidly? To explore this matter, here we used a probabilistic selection task in combination with computational modeling (i.e., Reinforcement Learning Drift Diffusion Model analysis) to establish how self-relevance influences learning under conditions of uncertainty (i.e., choices are based on the perceived likelihood of positive and negative outcomes). Across two experiments, a consistent pattern of effects was observed. First, learning rates for both positive and negative prediction errors were slower for self-relevant compared to friend-relevant associations. Second, self-relevant (vs. friend-relevant) learning was characterized by the exploitation (vs. exploration) of choice selections. That is, in a complex (i.e., probabilistic) decision-making environment, previously rewarded self-related outcomes were selected more often than novel — but potentially riskier — alternatives. The implications of these findings for accounts of self-function are considered.

Keywords: self, learning, self-prioritization, probabilistic selection task, reinforcement learning drift diffusion model.

Sticky Me:

Self-Relevance Slows Reinforcement Learning

1. Introduction

The self is an indispensable psychological construct, providing coherence and continuity to the narrative that underpins a personal sense of being (Baars, 1988; Baumeister, 1998; Conway, 2005; Conway & Pleydell-Pearce, 2000; Gallagher, 2000; James, 1890; Markus & Nurius, 1986; Markus & Wurf, 1987; Oakley & Halligan, 2017). As Markus and Wurf (1987, pp. 299-300) reported, "the self-concept…interprets and organizes self-relevant actions and experiences, it has motivational consequences, providing the incentives, standards, plans, rules, and scripts for behavior; and it adjusts in response to challenges from the social environment." In other words, the self serves as a basic processing hub around which social-cognitive functioning unfolds (Humphreys & Sui, 2016; Sui & Humphreys, 2015).

In documenting how the self influences thinking and doing, a common theme runs through decades of research. Regardless of the specific outcome under investigation (e.g., attributions, memories), personal relevance biases information-processing and response selection in self-enhancing and self-prioritizing ways (e.g., Conway, 2005; Mezulis et al., 2004; Sedikides & Alicke, 2012; Sui & Humphreys, 2015; Sui & Rothstein, 2019; Symons & Johnson, 1997). Most strongly associated with the self-reference effect (SRE) in memory (Kelley et al., 2002; Maki & McCaul, 1985; Rogers et al., 1977) — whereby material enjoys a recollective benefit when processed in the context of the self compared to other people (e.g., family members, friends, celebrities) — comparable advantages also emerge when attention and decision-making are probed (e.g., Alexopoulos et al., 2012; Bargh & Pratto, 1986; Constable et al., 2019; Falbén et al., 2020; Gray et al., 2004; Golubickis et al., 2018; Shapiro et al., 1997; Sui et al., 2012, 2015). Indeed, such is the potency of self-prioritization (i.e., the self-prioritization effect [SPE], Sui et al., 2012), benefits accrue even when the stimuli paired with the self (vs. other people) are arbitrary and meaningless, such as geometric shapes, abstract symbols, and colors/sounds (Golubickis et al., 2017, 2020; Schäfer et al., 2015, 2016; Sui et al., 2012; Wang et al., 2016; Woźniak & Knoblich, 2019; Yin et al., 2019).

Despite an extensive literature cataloguing the effects of self-relevance on core components of social cognition, important issues nevertheless remain. In particular, aside from a few notable exceptions, research has largely overlooked the extent to which the personal significance of stimuli influences a fundamental and crucial facet of daily life, the rate at which information is learned (Liao et al., 2021; Lockwood et al., 2018). That is, just as self-relevance facilitates the detection, appraisal, and memorability of stimuli (Humphreys & Sui, 2016; Sui & Humphreys, 2015, 2017; Symons & Johnson, 1997), so too it may enhance how rapidly this material is acquired. In one of the few studies to explore this matter, Lockwood et al. (2018) adopted a deterministic associative-learning task in which participants had to learn, from a pool of fractals (i.e., abstract, unfamiliar stimuli), which items belonged to various social targets (Brovelli et al., 2008; Schultz et al., 1997).[1] Specifically, a single fractal appeared on each experimental trial and participants had to report (i.e., learn) whether the stimulus was owned by the self, a friend, or a stranger. Feedback was then provided indicating if the response was correct or incorrect, and the task was deterministic in that participants were told each target always possessed the same fractals. To establish the respective target-related learning rates, data were submitted to an associative learning (AL) algorithm (Sutton & Barto, 1998).

Lockwood et al.'s (2018) findings were revealing. Reflecting the operation of an egocentric decisional strategy (Epley & Gilovich, 2004; Golubickis et al., 2018, 2019), participants tended to report that the fractal presented on the first trial belonged to them, when in reality it was just as likely to be owned by either of the other targets. In addition, responses were faster and more accurate when learning about fractals owned-by-self compared to those that belonged to others.

[1] Forming (and probing) target-object associations through ownership is a common methodology to explore self-prioritization (Constable et al., 2011, 2014; 2019; Falbén et al., 2019, 2020; Golubickis et al., 2018, 2019, 2021).

Finally, learning rates were higher when acquiring knowledge about the self, although this effect was only significant when stranger comprised the target of comparison — learning rates for self and friend were comparable. The absence of a reliable difference in learning rates between self and friend is interesting as while a self-advantage has frequently been reported when the target of comparison is best friend (e.g., Ma & Han, 2010; Sui & Han, 2007; Sui et al., 2012, 2013; Zhu et al., 2007), some research has indicated that the benefits of personal-relevance can be attenuated, or even eliminated, when the self is compared with an intimate (i.e., highly familiar) other (Bower & Gilligan, 1979; Kuiper & Rogers, 1979; Symons & Johnson, 1997). Notwithstanding this observation, Lockwood et al. (2018) provided initial evidence for the biasing effects of self-relevance on aspects of associative learning.[2]

Building upon and extending prior research, here we also explored the extent to which the personal relevance (or otherwise) of material impacts learning. Our overarching objectives were to probe the characteristics of self-learning effects in a different task context (i.e., learning environment) and to establish the pathway through which these effects arise. In so doing, rather than adopting a deterministic learning paradigm, a probabilistic selection task (PST) was employed (Frank et al., 2004, 2007). We used this task for a couple of reasons. First, the PST explores reinforcement learning (RL) in uncertain (vs. certain) task environments (cf. Lockwood et al., 2018), thus examines the impact of self-relevance when knowledge is acquired under demanding decision-making conditions. It is possible, for example, that basic components of self-representation and self-function may prompt learning effects to diverge when studied in uncertain (i.e., probabilistic) compared to certain (i.e., deterministic) task settings (Gershman & Daw, 2017). Second, in combination with recent developments in computational modeling (i.e., Reinforcement Learning Drift Diffusion Model (RL-DDM) analysis), adoption of the PST enables identification of

[2] As Lockwood et al.'s (2018) neural findings are beyond the scope of the current investigation, here we focus only on their behavioral results.

the latent psychological processes that underpin RL (Fontanesi et al., 2019; Pedersen & Frank, 2020; Pedersen et al., 2017).

In the current PST, participants were presented with three different stimulus pairs (i.e., AB, CD, EF) — comprising symbols (i.e., Japanese Hiragana characters; see Frank et al., 2004, 2007) with an item in each pairing (i.e., A, C, E) representing either the self or a friend — and they were required to learn, following a series of choice selections, which of the symbols was most likely to denote each target based on feedback that was provided (see Figure 1). Critically, the feedback was probabilistic and varied for each stimulus pair (i.e., AB = 80% - 20%, CD = 70% - 30%, EF = 60% - 40%). For example, in AB trials, a choice of stimulus A led to positive feedback on 80% of the trials, whereas selecting stimulus B resulted in positive reinforcement on 20% of the trials. Thus, in this PST, learning was accomplished via choice-related feedback. Over numerous choice selections, participants learned which item in each pairing was more likely to be correct (i.e., represent self or friend; A, C, E rather than B, D, F) and the task was completed when sufficient levels of accuracy were achieved for each stimulus pair (Frank et al., 2004, 2007).

To identify the mechanisms underpinning learning, computational modeling was undertaken on the data. Specifically, based on recent developments, a Reinforcement Learning Drift Diffusion Model (RL-DDM) analysis was adopted (Fontanesi et al., 2019; Pedersen & Frank, 2020; Pedersen et al., 2017). Integrating sequential sampling and RL models, the RL-DDM pinpoints the psychological operations that underpin decision-making (i.e., choice selection) and how these are adjusted as learning progresses (Miletić et al., 2020; Pedersen & Frank, 2020; Ratcliff et al., 2016). This is realized through the simultaneous hierarchical Bayesian modeling of response time (RT) and choice data. A drift rate scaling parameter (v_{scaling}) measures sensitivity to feedback and the exploration-exploitation trade-off (Cohen et al., 2007), such that higher values indicate more confident learning based on current knowledge (Pedersen et al., 2017). A learning rate parameter (η) — ranging from zero to one — quantifies how quickly individuals learn, with larger values indicating utilization of current feedback (i.e., fast learning), and smaller values reflecting reduced updating from recently experienced outcomes (i.e., slow learning). In this respect, either a single learning rate (η) that captures all learning, or separate learning rates for negative and positive prediction errors (η^- & η^+ respectively) can be estimated (Miletić et al., 2020; Pedersen et al., 2017). Finally, the model also establishes how much evidence is needed to make a decision (i.e., threshold separation, *a*) and the efficiency of non-decisional processes (e.g., stimulus encoding, response execution, t_0).

Central to the current inquiry is the classic exploration-exploitation trade-off that underlies learning (Cohen et al., 2007; Daw et al., 2006; Sutton & Barto, 1998). Confronted with a decision-making dilemma, learning can entail either the exploitation of options that have been optimal in the past or the exploration of alternatives that, in the long run, may prove to be more rewarding (Cohen et al., 2007). That is, one can either stick with existing knowledge or try something new. Critically, whereas exploration generally facilitates the acquisition of information, exploitation yields immediate decisional rewards, but it may impair learning (Sutton & Barto, 1998). As such, whether self-relevance enhances or reduces learning relative to a target of comparison (e.g., friend) should be reflected in decisions to explore or exploit the choice selections during RL. In this regard, an interesting possibility is that, in complex (i.e., probabilistic) task settings, people may prefer to stick (i.e., exploit) rather than switch (i.e., explore) when to-be-learned material is self-relevant, thereby prompting a slower learning rate for information pertaining to the self compared to others (cf. Lockwood et al., 2018). Several strands of evidence suggest such an outcome.

According to Humphreys and Sui (2015), via enhanced binding, self-reference serves as a form of associative glue for perception, attention, and memory (Cunningham et al., 2008; Rogers et al., 1977; Sui et al., 2012; Wang et al., 2016). While generally facilitating information processing and response selection, these potent self-object associations can also impede performance in certain task contexts. For example, participants find it difficult to overcome prior self-shape (vs. friend-shape) associations when given the task of forming new relations (Wang et al., 2016) and display a stubborn preference for self-relevant (vs. other-relevant) items during decision-making (Constable et al., 2019; Golubickis et al., 2018, 2019; Lockwood et al., 2018). Although such sticky learning undoubtedly supports the maintenance of a stable self-concept — an essential component of social-cognitive functioning (Greenwald, 1980; Markus, 1977) — it also suggests that exploitation rather than exploration may be the preferred strategy when acquiring information pertaining to the self in uncertain (i.e., probabilistic) learning environments. That is, previously rewarded self-object associations may be selected more often than novel (but riskier) options, thereby reducing the learning rate for the acquisition of personally meaningful material. Accordingly, using a PST in conjunction with computational modeling, here we explored the possibility that self-relevance may slow RL relative to an optimal target of comparison (e.g., best friend).

2. Experiment 1

2.1 Method

2.1.1 Participants and Design

Fifty participants (33 females, 17 males, 3 others; M_{age} = 23.04, *SD* = 3.06), with normal or corrected-to-normal visual acuity, took part in the research. Data collection was conducted online using Prolific Academic (www.prolific.co), with each participant receiving compensation at the rate of £7.50 (~$10) per hour. Informed consent was obtained from participants prior to the commencement of the experiment and the protocol was reviewed and approved by the Ethics Committee at the School of Psychology, University of Plymouth. The experiment had a single factor (Correct Symbol: self or friend) repeated-measures design. To detect a significant effect, a sample of fifty participants afforded 92% power for a large effect size (i.e., *d* = .80; PANGEA, v .0.2).

2.1.2. Stimulus Materials and Procedure

Participants performed two versions of a PST (Frank et al., 2004, 2007), with each comprising a learning phase in which three pairs of symbols (denoted as AB, CD, and EF, see

Figure 1) were presented. Participants were instructed they were required to learn, based on feedback provided, which symbol in each pair was most likely to represent them (i.e., self) or their best friend. Following previous research, prior to the task, participants were requested to bring their best friend (i.e., target of comparison) to mind (Golubickis et al., 2018). After each choice selection, participants were informed that onscreen information would indicate whether their response was correct or incorrect. Half of the participants were randomly assigned to perform a version of the PST in which self-related symbols were more likely to be correct, followed by another version of the task in which friend-related items were more likely to comprise the correct response. That is, trial type (i.e., learning) was blocked by target. The order of the PSTs was reversed for the remaining participants.

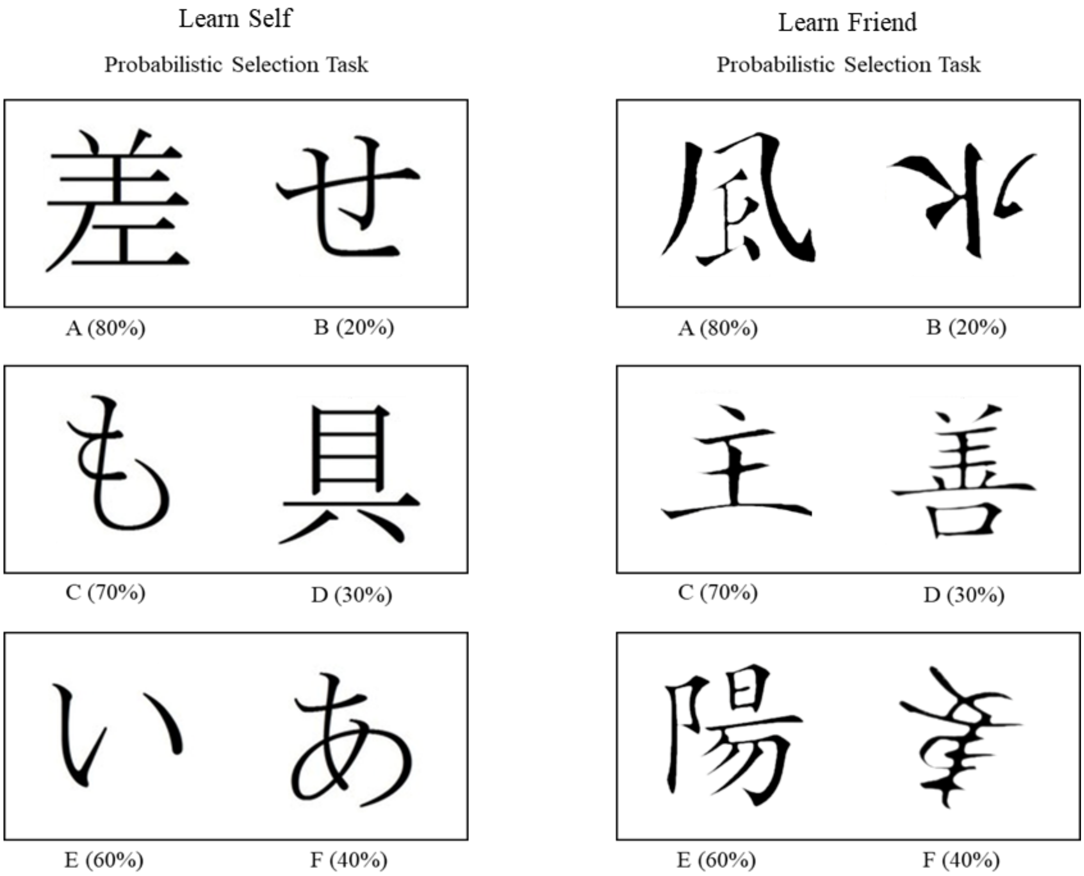

Figure 1. Example of the stimulus pairs (i.e., Japanese Hiragana characters) and the probabilities of correct responses during the probabilistic selection task.

The probabilities indicating which symbol was more likely to be correct followed the standard version of the PST (Frank et al., 2004, 2007). Specifically, for the AB pair, A was 80% likely to be correct (20% for B), for the CD pair, C was 70% likely to be correct (30% for D), and finally, for the EF pair, E was 60% likely to be correct (40% for F). Over numerous choice selections, participants learned which item in each pairing was more likely to be correct (i.e., A, C, E rather than B, D, F) based on the feedback provided. The task was completed when participants reached sufficient levels of accuracy for each pairing (i.e., AB, 60% or above; CD, 55% or above; EF, 50% or above; Frank et al., 2004, 2007).

Each trial began with the presentation of a pair of symbols that remained on the screen until the participant made a response. After the participant selected one of the symbols, feedback (i.e., the word ‘Correct’ in green or ‘Incorrect’ in red) was presented for 1000 ms, followed by a blank screen for 500 ms, after which the next trial commenced. Participants had to select a symbol by pressing the appropriate button on the keyboard (i.e., A for the symbol on the left side of the screen, L for the symbol on the right side of the screen). The symbols in each pair were equally likely to be presented on the left or right side of the screen. The experiment was conducted using Inquisit Web. Participants completed blocks of 60 trials in which each of the three stimulus pairs appeared randomly, equally often, until accuracy reached a satisfactory level. The maximum number of learning blocks was set to six (i.e., 360 trials in total) if the participant did not reach satisfactory levels of accuracy earlier in the task (Frank et al., 2007). On completion of the experiment, participants were debriefed and thanked.

2.2 Results and Discussion

2.2.1. Behavioral Analysis

The mean latency and accuracy of choice selections were submitted to a paired-sample (Correct Symbol: self or friend) *t*-test (two-tailed). No significant difference emerged on either dependent measure (i.e., decision time: M_{self} = 1203 ms vs. M_{friend} = 1148 ms; learning performance: M_{self} = 68% vs. M_{friend} = 66%).

2.1.2. Modeling Analysis

To identify the processes underpinning learning, data were submitted to a RL-DDM analysis (Frank et al., 2015; Pedersen & Frank, 2020; Pedersen et al., 2017). This analysis combines the strengths of RL and sequential-sampling models (SSMs) to elucidate the operations that support task performance. Specifically, although RL models account for changes in the relative proportion of choice probabilities over the course of learning, they do not speak to concurrent differences in response latencies, a fundamental and important dimension of the available data (e.g., as learning takes place, decision times decrease). In this respect, SSMs (e.g., drift diffusion model; Ratcliff et al., 2016; Smith & Radcliff, 2004) are useful as they provide a mechanistic account of binary decision-making by explaining how choice accuracy and response latencies collectively arise from a common set of latent cognitive processes (e.g., rate of evidence accumulation, response caution). Thus, crucially, the RL-DDM extends standard RL models by explicating the processes through which learning unfolds over time (Fontanesi et al., 2019; Miletic et al., 2020; Pedersen & Frank, 2020; Pedersen et al., 2017).

Two significant modifications characterize the RL-DDM. First, the typical choice rule for reinforcement learning (i.e., softmax) is replaced by the drift diffusion model (i.e., Wiener process, see Miletić et al. 2020; Pedersen et al., 2017). This change is important as it affords the possibility to model choice and RT data simultaneously. Second, the algorithm that captures the learning of subjective expectation values from stimuli and actions (i.e., value-based approach) is integrated into the process of evidence accumulation (i.e., drift rate). Thus, applying the delta learning rule, the model initially describes the updating of the expected Q-value for a chosen option (e.g., positively reinforced symbol A) based on the scaled by learning rate (α) reward prediction error (i.e., the difference between observed and expected feedback) in the previous trial (Rescorla & Wagner 1972; Watkins & Dayan 1992, see Eq. 1):

$$Q_{\text{chosen-option}}\,(t) = Q_{\text{chosen-option}}\,(t\text{-}1) + \alpha\,(\text{Reward}\,(t\text{-}1) - Q_{\text{chosen-option}}\,(t\text{-}1)) \tag{1}$$

Subsequently, the RL-DDM formulates the drift rate (*v*) during reinforced decisions based on the difference between the expected value of positively ($Q_{\text{positively-reinforced}}$) and negatively ($Q_{\textit{negatively-reinforced}}$) reinforced choices. To accommodate the manner in which this knowledge is used, the RL-DDM allows an additional free scaling parameter to be estimated (i.e., drift rate scaling, v_{scaling}). This scaling parameter is similar to inverse temperature in the softmax choice rule and reflects the level of exploration/exploitation during learning (Pedersen & Frank, 2020), such that larger values reflect stronger exploitation of the option with the highest expected value (see Eq. 2).

$$v\,(t) = (Q_{\text{positively-reinforced}}\,(t) - Q_{\textit{negatively-reinforced}}\,(t)) * v_{\text{scaling}} \tag{2}$$

Thus, in essence, the RL-DDM assumes that evidence is gathered for each choice option (e.g., symbol A vs. symbol B) until a critical evidential threshold is reached, at which point a response is made. This response threshold is captured by the boundary separation (*a*) parameter, and it reflects speed-accuracy trade-offs during decision-making. For example, if a conservative (vs. liberal) decision-making style (i.e., higher evidential requirements) is adopted, this would yield slower but more accurate responses. At the start of the PST, participants make slow guesses as the stimuli have not yet been reinforced, thus the difference in expected values between symbol pairings is extremely low (i.e., slow evidence accumulation due to high uncertainty). As participants start to receive feedback, via application of the delta learning rule (Rescorla & Wagner, 1972), the subjective Q-values of positively/negatively reinforced stimuli increase/decrease. The speed at which participants update the expected values is described by the learning rate (η) parameter. On a trial-by-trial basis, this knowledge (i.e., learning which symbol is correct, Q-value) is integrated into the drift rate such that over time the difference in expected values between reinforced options (ACE vs. BDF symbol pairings) increases. The larger the difference between positively and negatively reinforced options, the easier (i.e., faster and more accurate) choice selection becomes (i.e., fast information sampling).

To estimate model parameters, an extension of the Bayesian hierarchical drift diffusion toolbox was adopted (Wiecki et al., 2013). Models were response-coded, such that the upper threshold corresponded to responses to stimuli that were positively reinforced (i.e., symbols corresponding to the letters A, C, & E) and the lower threshold to stimuli that were negatively reinforced (i.e., symbols corresponding to the letters B, D, & F; Pedersen & Frank, 2020). Bayesian posterior distributions were modeled using a Markov chain Monte Carlo (MCMC) with 10,000 samples (including 1,000 burn), with outliers (5% of the trials) removed by the HDDM software (Ratcliff & Tuerlinckx, 2002; Wiecki et al., 2013). Two RL-DDM models were estimated for comparison (i.e., single vs. dual learning rate model). In the first model, only a single learning rate (η) was allowed to vary across Correct Symbol (i.e., self vs. friend). This model examined whether there were differences in the speed of learning across the experimental conditions without taking the potential influence of different types of prediction error into consideration. In contrast, in the second model, learning rates for negative and positive prediction errors (η^- & η^+, respectively) were allowed to vary by Correct Symbol. As such, this model considered whether learning self-related or friend-related stimuli was accelerated following negative or positive prediction errors. In both models, drift rate scaling (v_{scaling}) and boundary separation (a) varied across Correct Symbol.

Model comparison was performed using the Deviance Information Criterion (DIC) as this approach is routinely adopted when comparing hierarchical Bayesian models (Spiegelhalter et al., 1998, 2002). Lower DIC values favor models with the highest likelihood and least number of parameters. This revealed better fit for the dual (DIC: 60999) compared to the single (DIC: 61059) learning rate model. Examination of the posterior distributions (see Figure 2) revealed differences in learning rates for negative and positive prediction errors (η^-& η^+), drift rate scaling (v_{scaling}), and threshold separation (a). Specifically, comparisons yielded very strong evidence that learning rates were faster for friend compared to self, both for negative (p_{Bayes}(self < friend) = .032, BF_{10} = 30) and positive (p_{Bayes}(self < friend) < .001, BF_{10} > 1000) prediction errors.[3] In addition, participants integrated information more efficiently from negative than positive prediction errors, an effect that was larger for self ($p_{\text{Bayes}}(\eta^+ < \eta^-)$ = .008, BF_{10} = 125) than friend ($p_{\text{Bayes}}(\eta^+ < \eta^-)$ = .162, BF_{10} = 6). There was also very strong evidence that drift rate scaling (v_{scaling}) was larger for self- than friend-related symbols (p_{Bayes}(self > friend) = .019, BF_{10} = 52). Finally, for boundary separation (a), there was extremely strong evidence that more decisional information was required when selecting self- compared to friend-related responses (p_{Bayes}(self > friend) < .001, BF_{10} > 1000).

[3] Bayes Factors were transformed from Bayesian *p*-values (for details see Marsman & Wagenmakers, 2017).

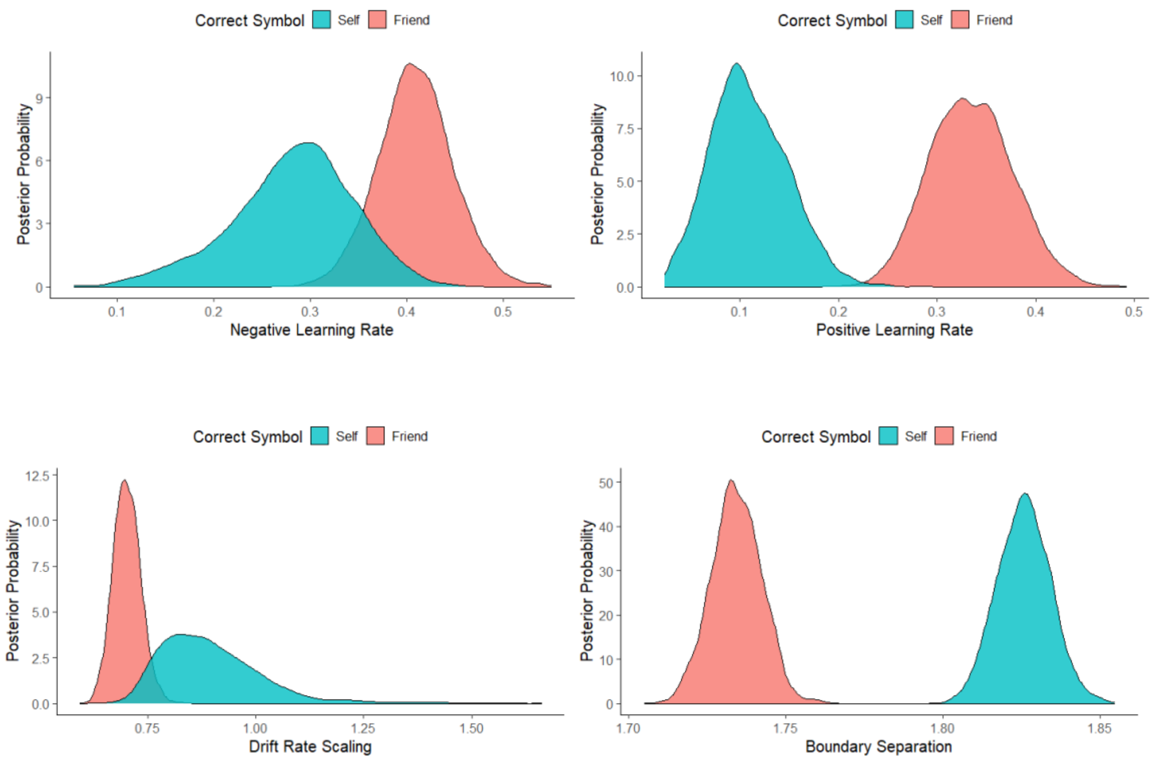

Figure 2. Mean posterior parameter distributions as a function of Correct Symbol for negative (η^-) and positive (η^+) learning rates, drift rate scaling (v_{scaling}) and boundary separation (a).

These findings reveal that, in a probabilistic task context (Frank et al., 2004, 2007), self-relevance (vs. friend-relevance) reduced the rate of learning. In addition, the RL-DDM analysis also indicated a difference in the balance between the strategies that drive learning — exploration and exploitation (Cohen et al., 2007; Sutton & Barto, 1998). Specifically, as indexed by the drift rate scaling parameter (v_{scaling}), self-relevant (vs. friend-relevant) trials were characterized by the tendency to exploit previously rewarded outcomes rather than explore new alternatives. In other words, self-relevance elicited a greater sensitivity to current outcomes (i.e., existing knowledge) during learning (Pedersen et al., 2017).

To probe the reproducibility of these effects, in our next experiment we also explored how self-relevance influenced learning in a PST (Frank et al., 2004, 2007), but with an important methodological modification. Rather than blocking the PST (i.e., learning) by target, participants simultaneously learned about self and friend in an intermixed design as previous research has demonstrated that self-relevance exerts a greater influence on decisional processing under these conditions (Golubickis & Macrae, 2021). Replicating Experiment 1, we expected self-relevance (vs. friend-relevance) to reduce the rate of learning and favor exploitation (vs. exploration) of the choice selections.

3. Experiment 2

3.1. Method

3.1.1. Participants and Design

Thirty-four participants (22 females, 10 males, 2 others; M_{age} = 22.97, *SD* = 2.62), with normal or corrected-to-normal visual acuity, took part in the research. Data collection was conducted online using Prolific Academic (www.prolific.co), with each participant receiving compensation at the rate of £7.50 (~$10) per hour. Informed consent was obtained from participants prior to the commencement of the experiment and the protocol was reviewed and approved by the Ethics Committee at the School of Psychology, University of Plymouth. The experiment had a single factor (Correct Symbol: self or friend) repeated-measures design. To detect a significant effect, a sample of thirty-four participants afforded 80% power for a large effect size (i.e., *d* = .80; PANGEA, v .0.2).

3.1.2. Stimulus Materials and Procedure

A modified version of the PST from Experiment 1 was adopted. Specifically, on a trial-by-trial basis, participants were required to learn which symbol in each pairing was more likely to represent self or best friend. Before the presentation of each stimulus pair, a cue (i.e., the labels "YOU" or "FRIEND") appeared on the screen indicating the target to which the symbols pertained (see Figure 3). The cue appeared 500 ms before the symbols and remained on the screen, above the stimuli, until a response was made. Participants completed blocks of 120 trials (i.e., 60 self and 60 friend) in which each stimulus pair appeared randomly, equally often, until accuracy reached a satisfactory level. The maximum number of learning blocks was set to three (i.e., 360 trials in total) if the participant did not reach satisfactory levels of accuracy earlier in the task (Frank et al., 2007). In all other respects, the procedure was identical to Experiment 1.

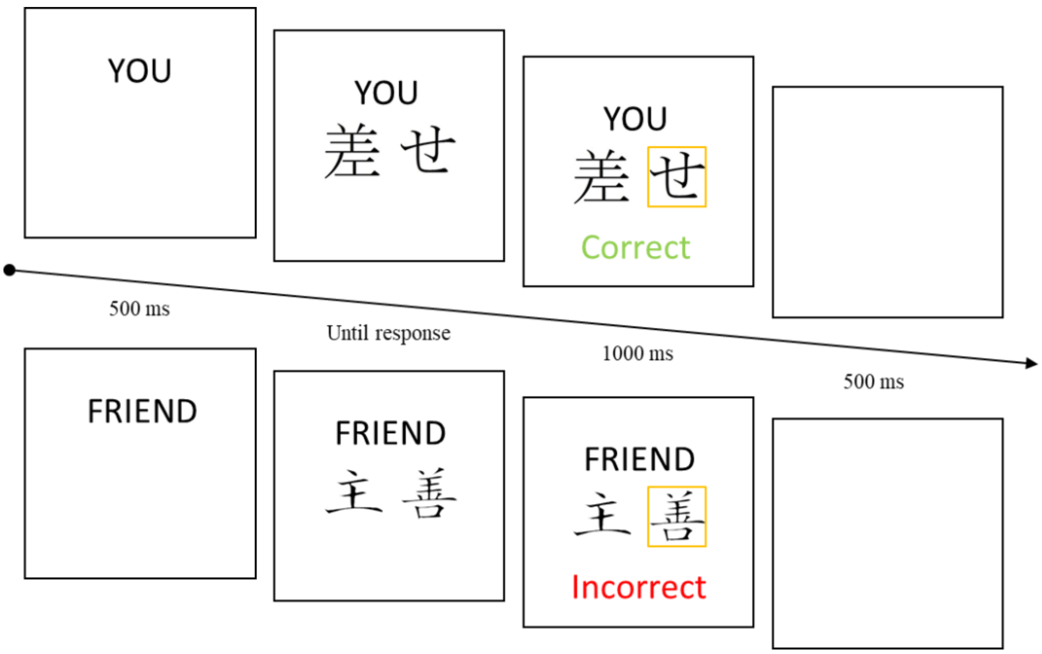

Figure 3. Examples of the experimental trials.

3.2. Results and Discussion

3.2.1. Behavioral Analysis

Four participants (3 females) failed to learn the probabilities associated with the symbols, thus were excluded from the analyses. The mean latency and accuracy of choice selections were submitted to a paired-sample (Correct Symbol: self or friend) *t*-test (two-tailed). The analysis of choice latencies revealed faster responses to self-related compared to friend-related symbols, $t(29)$ =

2.77, $p = .010$, $d = .51$; respective *M*s: 1546 ms vs. 1689 ms). In addition, accuracy was greater for self-related than friend-related stimuli, $t(29) = 3.39$, $p = .002$, $d = .62$; respective *M*s: 70% vs. 63%).

3.2.2. Modeling Analysis

To identify the processes underpinning learning, data were submitted to a RL-DDM analysis following the same modeling procedure as Experiment 1. As previously, fit was better for the dual (DIC: 43524) compared to the single (DIC: 43541) learning rate model. Examination of the posterior distributions (see Figure 4) revealed differences in learning rates for negative and positive prediction errors (η^-& η^+), drift rate scaling (v_{scaling}), and threshold separation (a). Specifically, comparisons yielded very strong evidence that learning rates were faster for friend compared to self, both for negative (p_{Bayes}(self < friend) = .011, $BF_{10} = 90$) and positive (p_{Bayes}(self < friend) = .005, $BF_{10} = 199$) prediction errors. As in Experiment 1, participants integrated information more efficiently from negative than positive prediction errors, an effect that was larger for self ($p_{\text{Bayes}}(\eta^+ < \eta^-) = .03$, $BF_{10} = 33$) than friend ($p_{\text{Bayes}}(\eta^+ < \eta^-) = .10$, $BF_{10} = 10$). There was also extremely strong evidence that drift rate scaling (v_{scaling}) was larger for self-related than friend-related symbols (p_{Bayes}(self > friend) < .001, $BF_{10} > 1000$). Finally, for boundary separation (a), there was extremely strong evidence that more decisional information was required when selecting friend- compared to self-related responses (p_{Bayes}(self < friend) < .001, $BF_{10} > 1000$).

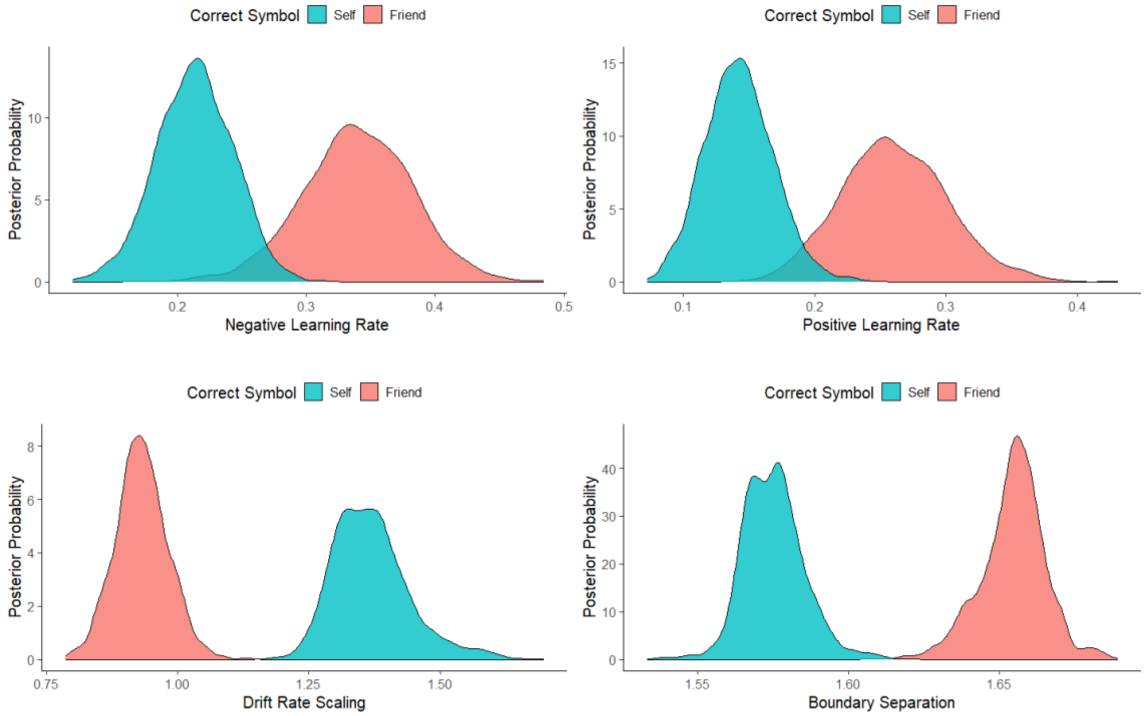

Figure 4. Mean posterior parameter distributions as a function of Correct Symbol for negative (η^{-}) and positive (η^{+}) learning rates, drift rate scaling (v_{scaling}) and boundary separation (*a*).

Using a different experimental design, these findings replicated the effects observed in Experiment 1. First, for both negative and positive prediction errors, learning rates were slower for self-related compared to friend-related symbols. Second, reflecting a greater reliance on existing knowledge (i.e., sensitivity to current outcomes), self-relevant (vs. friend-relevant) trials were characterized by the tendency to exploit previously rewarded outcomes rather than explore new choice selections (Pedersen et al., 2017). Interestingly, unlike Experiment 1 in which response caution was greater for self-relevant compared to friend-relevant symbols, this effect was reversed in the current experiment. This reversal can likely be traced to task-specific differences in the presentation of the stimulus trials during the PST (i.e., Expt. 1 - blocked by target; Expt. 2 - intermixed; Golubickis & Macrae, 2021).

4. General Discussion

Notwithstanding the acknowledged benefits that self-relevance exerts on information processing and response selection (Humphreys & Sui, 2016; Sui & Humphreys, 2015, 2017; Symons & Johnson, 1997), here we demonstrated a quite different effect. In the context of a PST, self-relevance (vs. friend-relevance) reduced the rate at which information was acquired. Specifically, whether stimuli were blocked by target (Expt. 1) or intermixed (Expt. 2), learning rates were slower for self-related compared to friend-related associations. In addition, self-relevant (vs. friend-relevant) learning was characterized by the tendency to exploit rather than explore the choice selections during the task (Cohen et al., 2007; Sutton & Barto, 1998). This indicates that, in a complex (i.e., probabilistic) decision-making setting, previously rewarded self-related outcomes were chosen more often than novel — but potentially riskier — choice selections. In other words, when learning about the self (vs. friend), participants tended to rely on their existing knowledge, thereby trading enhanced future learning for guaranteed current rewards (Pedersen et al., 2017).

That self-relevance has the capacity to impair performance in certain task contexts is unsurprising. Forging immediate and powerful target-object associations in working memory, personal-relevance (vs. friend-relevance) yields substantial processing benefits when responding is driven by the enhanced accessibility of these relations (Humphreys & Sui, 2016; Sui & Humphreys, 2015, 2017). That is, highly accessible self-object associations — even when the stimuli in question are unfamiliar and trivial — give rise to rapid and accurate responses (e.g., Golubickis et al., 2017, 2020; Schäffer et al., 2016, 2017; Stein et al., 2016; Sui et al., 2012, 2013; Woźniak & Knoblich, 2019). The strength of these sticky associations, however, can also hinder performance, particularly when participants must override previous learning experiences and acquire new target-object relations (Constable & Knoblich, 2020; Wang et al., 2016). For example, Wang et al. (2016) reported that, once self-shape associations were formed, participants found it difficult to break (i.e., undo) these relations and associate the shapes with a new target (e.g., friend). As they reported (p.

255), "…self-association can either enhance or disrupt processing, depending on whether new associations are assessed or whether old associations have to be discarded."

By enhancing the binding of target-object relations, self-relevance has obvious implications for decision-making and learning, at least in settings in which these associations are a task-relevant component of the methodology (Caughey et al., 2021; Constable et al., 2019; Falbén et al., 2019; Woźniak & Knoblich, 2021). As demonstrated here, in a PST (Frank et al., 2004, 2007), learning rates were slower when material was self-relevant (vs. friend-relevant). Several factors probably contributed to the emergence of this effect. Most notably, by shifting the balance toward exploitation rather than exploration during RL, choice selections served both to bolster the stability of the self-concept and optimize response-related rewards. A basic component of social-cognitive functioning is the possession (and maintenance) of a stable self-concept (Greenwald, 1980; Markus, 1977). In this respect, favoring choice selections that previously were (correctly) associated with the self would unquestionably service this objective.

In addition, the reward value of self-relevant (vs. friend-relevant) outcomes would similarly encourage exploitation over exploration (Cohen et al., 2007). According to Northhoff and Hayes (2011), self-referential processing is underpinned by the intrinsic reward-related properties of self-relevant stimuli (Northhoff & Hayes, 2011). Given the pivotal role of reward value during learning (Dayan & Belleine, 2002; Schultz, 1998; Sutton & Barto, 1998), exploiting formerly successful self-related outcomes would be particularly appealing (i.e., dopamine uptake), much more so than comparable friend-related responses or the exploration of novel choice selections. As such, although the precise relationship between self and reward remains a matter of continued scrutiny and debate (Sui et al., 2015; Stolte et al., 2015), during probabilistic learning this connection is likely intimate. Interestingly, in each of the reported experiments, learning was more effective following negative than positive prediction errors, an effect that was most pronounced for the self (vs. friend). It is possible that the tendency to exploit rather explore choice selections during self-related learning (i.e., sticky self-symbol associations) may underpin this asymmetry. Future research should explore this possibility.

Although, in the current investigation, the rate of learning was slower for self-relevant compared to friend-relevant stimuli, it is unlikely this effect is immutable. Indeed, as noted earlier, Lockwood and colleagues (2018) reported that, during deterministic learning, personal (vs. other) associations were formed most rapidly, albeit only when stranger comprised the target of comparison. For a familiar target of comparison (i.e., friend), self-other learning rates did not differ significantly. These inconsistent findings potentially derive from differences in self-function across probabilistic and deterministic learning environments (Gershman & Daw, 2017). In a fully certain (i.e., deterministic) world, exploration is not a viable strategy as pursuing new choice selections following positive feedback would impair performance. In contrast, in probabilistic settings (e.g., PSTs) feedback is accompanied by uncertainty (Frank et al., 2004, 2007), thereby moderating the balance between the competing strategies that drive choice selections (i.e., exploration-exploitation trade-off). As was observed in the current experiments, self-relevant (vs. friend-relevant) learning was characterized by the tendency to exploit rather than explore the response-related outcomes, such that potentially enhanced knowledge acquisition was traded for the certainty of immediate rewards (Cohen et al., 2007). This suggests that, depending on the characteristics of the learning environment (i.e., deterministic vs. probabilistic), self-relevance can exert quite different effects on RL.

Operating in this flexible way, learning mirrors the other domains in which the effects of self-relevance have been explored (e.g., attention, memory, decision-making). Inspection of a rapidly developing literature reveals the inherent malleability of self-prioritization and the divergent cognitive origins of self-bias. Specifically, whether self-prioritization facilitates or impedes performance — or indeed arises at all — is highly contingent upon the way in which self-object associations are operationalized, established, and probed (Caughey et al., 2021; Constable et al., 2019; Falbén et al., 2019, 2020; Golubickis et al., 2020, 2021; Macrae et al., 2017, 2018; Siebold et al., 2015; Stein et al., 2016; Svensson et al., 2021; Wang et al., 2016; Woźniak & Knoblich, 2021). Moreover, whereas in some task contexts self-relevance influences the efficiency of stimulus processing (Golubickis et al., 2017, 2020), in others it impacts response-related operations (Constable et al., 2019; Falbén et al., 2020; Golubickis et al., 2018, 2019). A useful task for future research will therefore be to establish how this contextual-dependency modulates the acquisition of self-knowledge across learning environments that vary in important ways; including the identity and number of targets of comparison, the characteristics of the to-be-learned material, and the distribution of rewards (Haruno & Kawato, 2006; Lockwood et al., 2018; Knowlton et al., 1994).

Attention should also be directed to the task context in which information pertaining to the self and others is encountered. Here differences in response caution were observed across two instrumental learning experiments that differed in task structure. Specifically, whereas response caution was greater on self-relevant compared to friend-relevant trials when stimuli were blocked by target (i.e., Experiment 1), this effect was reversed when the trial types were intermixed (i.e., Experiment 2). Relatedly, both Golubickis and Macrae (2021) and Desebrock et al. (in press) have similarly demonstrated the sensitivity of self-referential processing to the characteristics of the task environment. For example, using a shape-label matching task, Golubickis and Macrae (2021) observed a reduction in self-prioritization when stimuli were intermixed compared to blocked by target. Extending this finding, again in a shape-label matching task but using unisensory and multisensory stimuli, Desebrock et al. (in press) found that self-prioritization was greatest when trials were blocked by sensory modality. Collectively, these findings highlight the contextual dependence of self-bias, a factor that has largely been overlooked in research to date.

Consideration should also be given to the neural mechanisms that support the learning of material pertaining to the self and others. For example, is the acquisition of person-related knowledge underpinned by the same associative operations that drive reward-based learning in non-social contexts? Given the established role of the pre-frontal cortex (PFC) during self-referential processing (Kelley et al., 2002; Mitchell et al., 2002, 2006; Sui et al., 2013), it is interesting to note that resolution of the exploration-exploitation dilemma is also associated with activation in this region (Blanchard & Gershman, 2018; Domenech et al., 2020). Specifically, whereas activity in the ventromedial PFC (vmPFC) indexes the subjective value of outcomes given the action plan that is currently in place, modulation in dorsomedial PFC (dmPFC) reflects a reduction in these values and the generation of new response-related strategies (Donoso et al., 2014). In their investigation of the neural correlates of self-learning, Lockwood et al. (2018) reported that no brain area tracked exclusively with self-bias (i.e., self-ownership effect) during a deterministic learning task. Nevertheless, vmPFC responded more strongly to self- compared to stranger-related (but not friend-related) associations. As the current experiments yielded differences in both learning rates and the drift-rate scaling parameter (i.e., exploration-exploitation trade-off) for self and friend, it would therefore be interesting to explore the neural mechanisms that underlie self/other learning during a PST. In such a task setting, distinct patterns of activation may emerge in the mPFC and other cortical regions that support learning (e.g., anterior cingulate cortex [ACC]; Kennerley et al., 2006; Holroyd & McClure, 2015).

5. Conclusion

Using a PST in combination with a RL-DDM analysis, here we considered how self-relevance influences instrumental learning. Across two experiments, learning rates were slower for self-related compared to friend-related associations and self-relevant (vs. friend-relevant) learning was characterized by exploitation (vs. exploration) of the choice selections. Together with related research (Lockwood et al., 2018), these findings affirm the utility of computational approaches in the investigation of core social-cognitive topics (Hackel & Amodio, 2018; Lockwood & Klein-Flügge, 2020). Continuing in this way, further research should clarify exactly when, how, and for whom self-relevance influences associative learning.

References

Alexopoulos, T., Muller, D., Ric, F., & Marendaz, C. (2012). I, me, mine: Automatic attentional capture by self-related stimuli, *European Journal of Social Psychology*, *42*, 770-779.

Baars, B. J. (1988). *A cognitive theory of consciousness*. New York: Cambridge University Press.

Bargh, J. A., & Pratto, F. (1986). Individual construct accessibility and perceptual selection. *Journal of Experimental Social Psychology*, *22*, 293-311.

Baumeister, R. F. (1998). The self. In D. T. Gilbert, S. T. Fiske, & G. Lindzey (Eds.), *Handbook of social psychology* (4th ed., pp. 680-740). New York: McGraw-Hill.

Blanchard, T. C., & Gershman, S. J. (2018). Pure correlates of exploration and exploitation in the human brain. *Cognitive, Affective, & Behavioral Neuroscience*, *18*, 117-126.

Bower, G. H., & Gilligan, S. G. (1979). Remembering information related to one's self. *Journal of Research in Personality*, *13*, 420-432.

Brovelli, A., Laksiri, N., Nazarian, B., Meunier, M., & Boussaoud, D. (2008). Understanding the neural computations of arbitrary visuomotor learning through fMRI and associative learning theory. *Cerebral Cortex*, *18*, 1485-1495.

Caughey, S., Falbén, J. K., Tsamadi, D., Persson, L. M., Golubickis, M., & Macrae, C. N. (2021). Self-prioritization during stimulus processing is not obligatory. *Psychological Research*, *85*, 503-508.

Cohen, J. D., McClure, S. M., & Yu, A. J. (2007). Should I stay or should I go? How the human brain manages the trade-off between exploitation and exploration. *Philosophical Transactions of the Royal Society B*, *362*, 933-942.

Constable, M. D., & Knoblich, G. (2020). Sticking together? Re-binding previous other-associated stimuli interferes with self-verification but not partner-verification. *Acta Psychologica*, *210*, 103167.

Constable, M. D., Kritikos, A., & Bayliss, A. P. (2011). Grasping the concept of personal property. *Cognition*, *119*, 430-437.

Constable, M. D., Kritikos, A., Lipp, O. V., & Bayliss, A. P. (2014). Object ownership and action: The influence of social context and choice on the physical manipulation of personal property. *Experimental Brain Research*, *232*, 3749-3761.

Constable, M. D., Welsh, T. N., Huffman, G., & Pratt, J. (2019). I before U: Temporal order judgements reveal bias for self-owned objects. *Quarterly Journal of Experimental Psychology*, *72*, 589-598.

Conway, M. A. (2005). Memory and the self. *Journal of Memory and Language*, *53*, 594-628.

Conway, M. A., & Pleydell-Pearce, C. W. (2000). The construction of autobiographical memories in the self-memory system. *Psychological Review*, *107*, 261-288.

Cunningham, S. J., Turk, D. J., Macdonald, L. M., & Macrae, C. N. (2008). Yours or mine? Ownership and memory. *Consciousness and Cognition*, *17*, 312-318.

Daw, N. D., O'Doherty, J. P., Dayan, P., Seymour, B., & Dolan, R. J. (2006). Cortical substrates for exploratory decisions in humans. *Nature*, *441*, 876-879.

Dayan, P., & Balleine, B. W. (2002). Reward, motivation, and reinforcement learning. *Neuron*, *36*, 285-298.

Domenech, P., Rheims, S., & Koechlin, E. (2020). Neural mechanisms resolving exploitation-exploration dilemmas in the medial prefrontal cortex. *Scienc*e, *369*, eabb0184.

Donoso, M., Collins, A. G., & Koechlin, E. (2014). Foundations of human reasoning in the prefrontal cortex. *Science*, *344*, 1481-1486.

Epley, N., & Gilovich, T. (2004). Are adjustments insufficient? *Personality and Social Psychology Bulletin*, *30*, 447-460.

Falbén, J. K., Golubickis, M., Balseryte, R., Persson, L. M., Tsamadi, D., Caughey, S., & Macrae, C. N. (2019). How prioritized is self-prioritization during stimulus processing? *Visual Cognition*, *27*, 46-51.

Falbén, J. K., Golubickis, M., Wischerath, D., Tsamadi, D., Persson, L. M., Caughey, S., Svensson, S., & Macrae, C. N. (2020). It's not always about me: The effects of prior beliefs and stimulus prevalence on self-other prioritization. *Quarterly Journal of Experimental Psychology*, *73*, 1466-1480.

Fontanesi, L., Gluth, S., Spektor, M. S., & Rieskamp, J. (2019). A reinforcement learning diffusion decision model for value-based decisions. *Psychonomic Bulletin and Review, 26,* 1099-1121.

Frank, M. J., Moustafa, A. A., Haughey, H. M., Curran, T., & Hutchinson, K. E. (2007). Genetic triple dissociation reveals multiple roles for dopamine in reinforcement learning. *Proceedings in the National Academy of Sciences, 104,* 16311-16316.

Frank, M. J., Seeberger, L. C., & O'Reilly, R. C. (2004). By carrot or by stick: Cognitive reinforcement learning in Parkinsonism. *Science, 306,* 1940-1943.

Gallagher, S. (2000). Philosophical conceptions of the self: Implications for cognitive science. *Trends in Cognitive Sciences*, *4*, 14-21.

Gershman, S. J., & Daw, N. D. (2017). Reinforcement learning and episodic memory in humans and animals: An integrative framework. *Annual Review of Psychology*, *68*, 101-128.

Golubickis, M., & Macrae, C. N. (2021). Judging me and you: Task design modulates self-prioritization. *Acta Psychologica*, *218*, 103350.

Golubickis, M., Falbén, J. K., Sahraie, A., Visokomogilski, A., Cunningham, W. A., Sui, J., & Macrae, C. N. (2017). Self-prioritization and perceptual matching: The effects of temporal construal. *Memory & Cognition*, *45*, 1223-1239.

Golubickis, M., Falbén, J. K., Ho, N. S., Sui, J., Cunningham, W. A., & Macrae, C. N. (2020). Parts of me: Identity-relevance moderates self-prioritization. *Consciousness and Cognition*, *77*, 102848.

Golubickis, M., Falbén, J. K., Cunningham, W. A., & Macrae, C. N. (2018). Exploring the self-ownership effect: Separating stimulus and response biases. *Journal of Experimental Psychology: Learning, Memory and Cognition*, *44*, 295-306.

Golubickis, M., Ho, N. S. P., Falbén, J. K., Mackenzie, K. M., Boschetti, A., Cunningham, W. A., & Macrae, C. N. (2019). Mine or mother's? Exploring the self-ownership effect across cultures. *Culture and Brain*, *7*, 1-25.

Golubickis, M., Ho, N. S. P., Falbén, J. K., Schwertel, C. L., Maiuri, A., Dublas, D., Cunningham, W. A., & Macrae, C. N. (2021). Valence and ownership: Object desirability influences self-prioritization. *Psychological Research*, 85, 91-100.

Gray, H. M., Ambady, N., Lowenthal, W. T., & Deldin, P. (2004). P300 as an index of attention to self-relevant stimuli. *Journal of Experimental Social Psychology*, *40*, 216-224.

Greenwald, A. G. (1980). The totalitarian ego: Fabrication and revision of personal identity. *American Psychologist*, *35*, 603-618.

Hackel, L. M., & Amodio, D. M. (2018). Computational neuroscience approaches to social cognition. *Current Opinion in Psychology*, *24*, 92-97.

Haruno, M., & Kawato, M. (2006). Different neural correlates of reward expectation and reward expectation error in the putamen and caudate nucleus during stimulus-action-reward association learning. *Journal of Neurophysiology*, *95*, 948-959.

Holroyd, C. B., & McClure, S. M. (2015). Hierarchical control over effortful behavior by rodent medial frontal cortex: A computational model. *Psychological Review*, *122*, 54-83.

Humphreys, G. W., & Sui, J. (2016). Attentional control and the self: The self-attention network (SAN). *Cognitive Neuroscience*, *7*, 5-17.

James, W. (1890). *The principles of psychology*. New York: Henry-Holt & Co.

Kelley, W. M., Macrae, C. N., Wyland, C. L., Caglar, S., Inati, S., & Heatherton, T. F. (2002). Finding the self? An event-related fMRI study. *Journal of Cognitive Neuroscience*, *14*, 785-794.

Kennerley, S. W., Walton, M. E., Behrens, T. E., Buckley, M. J., & Rushworth, M. F. (2006). Optimal decision making and the anterior cingulate cortex. *Nature Neuroscience*, *9*, 940-947.

Knowlton, B. J., Squire, L. R., & Gluck, M. A. (1994). Probabilistic classification learning in amnesia. *Learning & Memory*, *1*, 106-120.

Kuiper, N. A., & Rogers, T. B. (1979). Encoding of personal information: Self–other differences. *Journal of Personality and Social Psychology*, *37*, 499-514.

Liao, Z., Huang, L., & Luo, S. (2021). Intranasal oxytocin decreases self-oriented learning. *Psychopharmacology*, *238*, 461-474.

Lockwood, P. L., & Klein-Flugge, M. C. (2020). Computational modelling of social cognition and behaviour: A reinforcement learning primer. *Social Cognitive and Affective Neuroscience, 16,* 761-771.

Lockwood, P. L., Wittmann, M. K., Apps, M. A., Klein-Flügge, M. C., Crockett, M. J., Humphreys, G. W., & Rushworth, M. F. (2018). Neural mechanisms for learning self and other ownership. *Nature Communications*, *9*, 1-11.

Ma, Y., & Han, S. (2010). Why we respond faster to the self than to others? An implicit positive association theory of self-advantage during implicit face perception. *Journal of Experimental Psychology: Human Perception and Performance*, *36*, 619-633.

Macrae, C. N., Visokomogilski, A., Golubickis, M., Cunningham, W. A., & Sahraie, A. (2017). Self-relevance prioritizes access to visual awareness. *Journal of Experimental Psychology: Human Perception and Performance*, *43*, 438-443.

Macrae, C. N., Visokomogilski, A., Golubickis, M., & Sahraie, A. (2018). Self-relevance enhances the benefits of attention on perception. *Visual Cognition*, *26*, 475-481.

Maki, R. H., & McCaul, K. D. (1985). The effects of self-reference versus other-reference on the recall of traits and nouns. *Bulletin of the Psychonomic Society*, *23*, 169-172.

Markus, H. (1977). Self-schemata and processing information about the self. *Journal of Personality and Social Psychology*, *35*, 63-78.

Markus, H. R., & Nurius, P. (1986). Possible selves. *American Psychologist*, *41*, 954-969

Markus, H., & Wurf, E. (1987). The dynamic self-concept: A social psychological perspective. *Annual Review of Psychology*, *38*, 299-337.

Marsman, M., & Wagenmakers, E. -J. (2017). Three insights from a Bayesian interpretation of the one-sided *p* value. *Educational and Psychological Measurement*, *77*, 529-539.

Mezulis, A. H., Abramson, L. Y., Hyde, J. S., & Hankin, B. L. (2004). Is there a universal positivity bias in attributions? A meta-analytic review of individual, developmental, and cultural differences in the self-serving attributional bias. *Psychological Bulletin*, *130*, 711-747.

Miletić, S., Boag, R. J., & Forstmann, B. U. (2020). Mutual benefits: Combining reinforcement learning with sequential sampling models. *Neuropsychologia*, *136*, 107261.

Mitchell, J. P., Heatherton, T. F., & Macrae, C. N. (2002). Distinct neural systems subserve person and object knowledge. *Proceedings of the National Academy of Sciences*, *99*, 15238-15243.

Mitchell, J. P., Macrae, C. N., & Banaji, M. R. (2006). Dissociable medial prefrontal contributions to judgments of similar and dissimilar others. *Neuron*, *50*, 655-663.

Northoff, G., & Hayes, D. J. (2011). Is our self nothing but reward? *Biological Psychiatry*, *69*, 1019-1025.

Oakley, D. A., & Halligan, P. W. (2017). Chasing the rainbow: the non-conscious nature of being. *Frontiers in Psychology*, *8*, 1924.

Pedersen, M. L., & Frank, M. J. (2020). Simultaneous hierarchical Bayesian parameter estimation for reinforcement learning and drift diffusion models: A tutorial and links to neural data. *Computational Brain & Behavior*, *3*, 458–471.

Pedersen, M. L., Frank, M. J., & Biele, G. (2017). The drift diffusion model as the choice rule in reinforcement learning. *Psychonomic Bulletin & Review*, *24*, 1234-1251.

Ratcliff, R., & Tuerlinckx, F. (2002). Estimating parameters of the diffusion model: Approaches dealing with contaminant reaction times and parameter variability. *Psychonomic Bulletin & Review*, *9*, 438-481.

Ratcliff, R., Smith, P. L., Brown, S. D., & McKoon, G. (2016). Diffusion decision model: Current issues and history. *Trends in Cognitive Sciences*, *20*, 260-281.

Rescorla, R. A., & Wagner, A. R. (1972). A theory of Pavlovian conditioning: Variations in the effectiveness of reinforcement and nonreinforcement. In A. H. Black & W. F. Prokasy (Eds.), *Classical conditioning II: Current research and theory* (pp. 64-99). Appleton Century Crofts.

Rogers, T. B., Kuiper, N. A., & Kirker, W. S. (1977). Self-reference and the encoding of personal information. *Journal of Personality and Social Psychology, 35*, 677-688.

Schäfer, S., Wentura, D., & Frings, C. (2015). Self-prioritization beyond perception. *Experimental Psychology*, *62*, 415-425.

Schäfer, S., Frings, C., & Wentura, D. (2016). About the composition of self-relevance: Conjunctions not features are bound to the self. *Psychonomic Bulletin and Review*, *23*, 887-982.

Schäfer, S., Wentura, D., & Frings, C. (2017). Distinctiveness effects in self-prioritization. *Visual Cognition*, *25*, 1-13.

Schultz, W., Dayan, P., & Montague, P. R. (1997). A neural substrate of prediction and reward. *Science*, *275*, 1593-1599.

Schultz, W. (1998). Predictive reward signal of dopamine neurons. *Journal of Neurophysiology*, *80*, 1-27.

Sedikides, C., & Alicke, M. D. (2012). Self-enhancement and self-protection motives. In R. M. Ryan (Ed.), *The Oxford handbook of human motivation* (pp. 303–322). Oxford University Press.

Shapiro, K. L., Caldwell, J., & Sorensen, R. E. (1997). Personal names and the attentional blink: A visual “cocktail party” effect. *Journal of Experimental Psychology: Human Perception and Performance, 23*, 504-514.

Siebold, A., Weaver, M. D., Donk, M., & van Zoest, W. (2015). Social salience does not transfer to oculomotor visual search. *Visual Cognition, 23*, 989-1019.

Spiegelhalter, D. J., Best, N. G., Carlin, B. P., & van der Linde, A. (1998). *Bayesian deviance, the effective number of parameters, and the comparison of arbitrarily complex models*. Research Report, 98-1009.

Spiegelhalter, D. J., Best, N. G., Carlin, B. P., & van der Linde, A. (2002). Bayesian measures of model complexity and fit. *Journal of the Royal Statistical Society: Series B, 64,* 583-639.

Stein, T., Siebold, A., & van Zoest, W. (2016). Testing the idea of privileged awareness of self-relevant information. *Journal of Experimental Psychology: Human Perception and Performance*, *42,* 303-307.

Stolte, M., Humphreys, G., Yankouskaya, A., & Sui, J. (2017). Dissociating biases towards the self and positive emotion. *Quarterly Journal of Experimental Psychology*, *70*, 1011-1022.

Sui, J., & Han, S. (2007). Self-construal priming modulates neural substrates of self-awareness. *Psychological Science*, *18*, 861-866.

Sui, J., He, X., & Humphreys, G. W. (2012). Perceptual effects of social salience: Evidence from self-prioritization effects on perceptual matching. *Journal of Experimental Psychology: Human Perception and Performance*, *38*, 1105-1117.

Sui, J., & Humphreys, G. W. (2015). The integrative self: How self-reference integrates perception and memory. *Trends in Cognitive Sciences*, *19*, 719-728.

Sui, J., & Humphreys, G. W. (2017). The ubiquitous self: What the properties of self-bias tell us about the self. *Annals of the New York Academy of Sciences*, *1396*, 222-235.

Sui, J., Liu, M., Mevorach, C., & Humphreys, G. W. (2015). The salient self: The left intra-parietal sulcus responds to social as well as perceptual salience after self-association. *Cerebral Cortex, 25*, 1060-1068.

Sui, J., & Rotshtein, P. (2019). Self-prioritization and the attentional systems. *Current Opinion in Psychology, 29*, 148-152.

Sui, J., Rotshtein, P., & Humphreys, G. W. (2013). Coupling social attention to the self forms a network for personal significance. *Proceedings of the National Academy of Sciences*, *110,* 7607-7612.

Sutton, R. S., & Barto, A. G. (1998). *Reinforcement learning: An introduction* (Vol. 135). Cambridge: MIT press.

Svensson, S. L., Golubickis, M., Maclean, H., Falbén, J. K., Persson, L. M., Tsamadi, D., ... & Macrae, C. N. (2021). More or less of me and you: Self-relevance augments the effects of item probability on stimulus prioritization. *Psychological Research*, 10.1007/s00426-021-01562-x. Advance online publication.

Symons, C. S., & Johnson, B. T. (1997). The self-reference effect in memory: A meta-analysis. *Psychological Bulletin*, *121*, 371-394.

Wang, H., Humphreys, G., & Sui, J. (2016). Expanding and retracting from the self: Gains and costs in switching self-associations. *Journal of Experimental Psychology: Human Perception and Performance*, *42*, 247-256.

Watkins, C., & Dayan, P. (1992). Q-learning. *Machine Learning*, *8*, 279-292.

Wiecki, T. V., Sofer, I., & Frank, M. J. (2013). HDDM: Hierarchical Bayesian estimation of the drift-diffusion model in Python. *Frontiers in Neuroinformatics*, *7*, 1-10.

Woźniak, M., & Knoblich, G. (2019). Self-prioritization of fully unfamiliar stimuli. *Quarterly Journal of Experimental Psychology*, *72*, 2110-2120.

Woźniak, M., & Knoblich, G. (2021). Self-prioritization depends on assumed task-relevance of self-association. *Psychological Research*. 10.1007/s00426-021-01584-5. Advance online publication.

Yin, S., Sui, J., Chiu Y-C, Chen, A., & Egner, T. (2019). Automatic prioritisation of self-referential stimuli in working memory. *Psychological Science*, *30*, 415-423.

Zhu, Y., Zhang, L., Fan, J., & Han, S. (2007). Neural basis of cultural influence on self-representation. *Neuroimage*, *34*, 1310-1316.