Acta Psychologica 139 (2012) 343–349

Contents lists available at SciVerse ScienceDirect

Acta Psychologica

journal homepage: www.elsevier.com/ locate/actpsy

Is the enhancement of memory due to reward driven by value or salience?

Christopher R. Madan *, Marcia L. Spetch

Department of Psychology, University of Alberta, Canada

ARTICLE INFO

Article history:
Received 3 May 2011
Received in revised form 2 December 2011
Accepted 21 December 2011
Available online 20 January 2012

PsycINFO classification:
2343
2360

Keywords:
Reward
Memory
Value-learning
Free recall
Reward salience

ABSTRACT

Past research using two levels of reward has shown that the higher-value items are remembered better than lower-value items and this enhancement is assumed to be driven by an effect of reward value. In the present study, multiple levels of reward were used to test the influence of reward salience on memory. Using a value-learning procedure, words were associated with reward values, and then memory for these words was later tested with free recall. Critically, multiple reward levels were used, allowing us to test two specific hypotheses whereby rewards can influence memory: (a) higher value items are remembered better than lower value items (reward value hypothesis), and (b) highest and lowest value items are remembered best and intermediate-value items are remembered worst (following a U-shaped relationship between value and memory; reward salience hypothesis). In two experiments we observed a U-shaped relationship between reward value and memory, supporting the notion that memory is enhanced due to reward salience, and not purely through reward value.

1. Introduction

In day-to-day life, people often remember stimuli associated with rewarding experiences better than those associated with less rewarding experiences. For example, you are likely to remember a person you enjoyed talking to at a party better than someone you found less interesting. Thus, experiences associated with higher reward values are likely to be remembered better than those with lower reward values and, given the choice, one would choose to repeat a more rewarding experience over a less rewarding experience. Thus, in the party example, you would be more likely to remember the person you enjoyed talking to, and at a future party you would likely choose to talk to that person over the less interesting one.

The enhancement of memory due to reward suggested by this anecdotal example has been the subject of a recent flurry of studies (e.g., Adcock, Thangavel, Whitfield-Gabrieli, Knutson, & Gabrieli, 2006; Castel, Benjamin, Craik, & Watkins, 2002; Gruber & Otten, 2010; Shigemune et al., 2010; Soderstrom & McCabe, 2011; Wittmann et al., 2005). This research has confirmed the effects of reward value on memory with two levels of reward (high-value versus low-value rewards, or reward versus no reward), and evidence suggests that reward enhances memory via dopaminergic modulation (reviewed in Shohamy & Adcock, 2010).

A recent study by Madan, Fujiwara, & Caplan, (under review) trained items to have reward values and later tested memory for these items in two unrewarded memory tasks: lexical decision (implicit memory) and free recall (explicit memory). They found enhanced memory for high-value items in *both* memory tasks, as well as a negative correlation between each task. Reward values were trained through a value-learning procedure (see Pessiglione, Seymour, Flandin, Dolan, & Frith, 2006), in which participants are presented with two items at a time and asked to choose one. At the beginning of training, performance is at chance, as item value is not yet known; however, through feedback, participants learn to choose higher-value items over lower-value items. This value-learning procedure is essentially a reinforcement learning procedure (Ludvig, Bellemare, & Pearson, 2011) and the effects of value on memory can be viewed in terms of prediction error (Rescorla & Wagner, 1972; Sutton & Barto, 1998). Specifically, during the learning phase, the higher the reward value, the bigger the prediction error between what is expected for that item and the reward that occurs. Higher prediction errors trigger more activation of the dopamine system (Lisman & Grace, 2005; Shohamy & Adcock, 2010) and should make the higher-valued items more memorable. In the present study we closely follow the design of Madan et al. (under review) with one critical change: the use of *multiple* reward levels. While Madan et al. trained words to be either high- or low-value, we also include one (Experiment 1) or several (Experiment 2) intermediate values.

Prior studies investigating the enhancement of memory due to reward often explicitly instructed participants of an item's value when the item was being intentionally studied. For example, Adcock

* Corresponding author at: Psychology Department, Biological Sciences Building, University of Alberta, Edmonton, Alberta, Canada T6G 2E9. Tel.: +1 780 492 7872; fax: +1 780 492 0023.
E-mail address: cmadan@ualberta.ca (C.R. Madan).

0001-6918/$ – see front matter © 2011 Elsevier B.V. All rights reserved.
doi:10.1016/j.actpsy.2011.12.010

et al. (2006) presented a reward value cue (e.g., “$5”) to participants just prior to presenting the to-be-remembered item. Participants would only earn the rewarded amount for successfully remembering the item during the subsequent memory test. Since these studies only provided rewards for correctly remembering an item, and the reward given was based on an item's value, participants should deliberately prioritize their memory for the high-value items at the cost of the low-value items. Castel et al. (2002) tested this directly by presenting participants with twelve items, one-at-a-time, along with a unique point value ranging from 1 to 12. Points would only be earned for the successful recall of the associated item. Here the researchers also found better memory for the higher-value items relative to the low-value items. Studies using rewarded memory tests suggest that participants deliberately prioritize their memory for the higher-value items and suggest that memory performance should increase monotonically with reward value. In our procedure, participants are not told that they will need to recall the items and therefore our task, unlike those used in these prior studies, does not necessitate that participants deliberately prioritize their memory. Nevertheless, it is still possible that participants may attend more to the higher-value items at the cost of the lower-value items. Thus, one hypothesis is that reward value will monotonically predict memory performance, which we term the reward value hypothesis.

A second hypothesis regarding the relationship between reward value and memory is suggested by recent evidence of neural representations for reward salience, in addition to that for reward value (Cooper & Knutson, 2008; Jensen et al., 2007; Litt, Plassmann, Shiv, & Rangel, 2011; Zink, Pagnoni, Martin-Skurski, Chappelow, & Berns, 2004). This neural evidence suggests that both the most positive and the most negative values are more salient than intermediate values. Although our study includes only positive values, it is possible that the extremes of the range of values experienced are more salient than those closer to the center of the range. Thus a second hypothesis is that memory will show a U-shaped relationship to reward value, with both the highest- and the lowest-valued items being remembered better than the intermediate valued items, which we term the reward salience hypothesis. We should note that this hypothesis could also be consistent with a prediction error interpretation, because the results of Jensen et al. (2007) suggest that some brain regions may in fact represent prediction error based on reward salience.

Our experiment was designed to test the role of both reward value and reward salience on memory, using multiple reward values and an unrewarded memory task. Through the inclusion of multiple reward values, as well as the use of an unrewarded memory task, our results will be able to determine if memory is enhanced due to reward value or reward salience. If there is a U-shaped (quadratic) relationship between reward value and memory as suggested by the reward salience hypothesis, then people may remember both the highest- and the lowest-valued items better than the intermediate-valued items.

The inclusion of multiple reward values also allows an examination of the relationships between choice and memory. The reward value hypothesis predicts a monotonic relationship between choice and memory because higher valued items should be chosen more and remembered better. On the other hand, the reward salience hypothesis predicts that choice frequency and memory for items will not be monotonically related. Instead, the lowest valued item will be chosen the least but will be remembered better than intermediate valued items that were chosen more frequently.

Thus, our experiments were designed to test whether the reward salience hypothesis fares better than the reward value hypothesis in predicting the relationship between reward value and memory and between choice and memory. With our multiple level value-learning procedure, the reward salience hypothesis predicts that people will learn to choose items in direct relation to their value (i.e., choice will be monotonically related to reward value) but that the lowest and the highest valued items will be remembered best (memory will show a U-shaped relationship with reward value).

2. Experiment 1

2.1. Methods

2.1.1. Participants

71 introductory psychology students (40 females) at the University of Alberta participated for partial fulfillment of course credit. All participants were required to have learned English before the age of six and were required to be comfortable typing. Participants gave written informed consent prior to beginning the study, which was approved by a University of Alberta Research Ethics Board.

2.1.2. Materials

All of the words and non-words used in this study have been previously used in Madan et al. (under review). Words were selected from the MRC Psycholinguistic database (Wilson, 1988). Imageability and word frequency were all held at mid-levels and all words had six to seven letters and exactly two syllables. Words were additionally controlled to be of neutral emotional valence and low arousal using the Affective Norms for English Words (ANEW; Bradley & Lang, 1999). 21 words were removed manually due to possible confounding effects (e.g., ‘reward’, ‘defeat’, ‘profit’) or because they were deemed by the authors to be emotional in nature but were not included in ANEW (e.g., ‘terror’, ‘regret’). The final word pool consisted of 160 words.

160 non-words were generated with the LINGUA non-word generator (Westbury, Hollis, & Shaoul, 2007). Non-words were generated using a Markov chaining length of three. Half of the non-words were generated to be six letters in length, with the remaining half being seven letters in length, in order to match the length of the non-words to the words.

2.1.3. Procedure

Prior to the experiment, participants were informed that the experiment was a ‘word choice task,’ and that they would receive an honorarium proportional to the total points earned in the value-learning phase of the experiment, in addition to their partial course credit.

The experiment consisted of four sequential tasks: value-learning, lexical decision, free recall, and a value judgment task. Participants were not provided with details of the subsequent task until the current task was completed.

2.1.3.1. Value learning. Participants were shown two words on the computer screen simultaneously. Participants were instructed to choose one of the two words in each choice set. Participants pressed the ‘Z’ key of a computer keyboard to choose the word presented on the left side of the computer screen; to choose the word on the right side of the screen they were instructed to press the ‘/’ key.

For each participant, 36 words were randomly selected from our pool of 160 words and randomly assigned to one of three reward levels: 2, 7, or 12 points. Thus, assignment of words to reward values varied across participants. Sets were pseudorandomly generated each round to never pair two words of the same reward level and to present pairings of all possible reward levels an equal number of times. This constraint was not revealed to the participant. Each choice by the participant immediately resulted in earning the respective reward. When a choice was made by the participant, an image of a pile of coins was shown to the participant in the center of the computer screen for 2000 ms, where the number of coins in the image was directly proportional to the number of points earned (e.g., if the participant earned 12 points, the image displayed a pile of 12 coins). The participant's current point balance was continually presented at the bottom of the computer screen throughout the duration of the reward training phase. There was no time limit on how quickly the choices had to be made and participants were given a 1000 ms delay before the next choice was presented.

Training consisted of 18 choice sets per round for 13 rounds. At the end of the session, participants were paid $1.00 for every 400 points earned during the value-learning task, rounded up to the nearest 25 cent amount. All participants earned between $3.25 and $4.50.

2.1.3.2. Lexical decision. 12 additional words, selected at random from the same pool as the previously rewarded words, were included as 'new' words. Participants were then asked to judge the lexical status of 96 items: 36 previously rewarded words, 12 new words and 48 non-words. Each item was presented for up to 10,000 ms, and the participant pressed either 'Z' on the computer keyboard to indicate that the item was a word, or '/' to indicate that the item was a non-word. A fixation cross ('+') was presented for 1000 ms to separate each decision prompt.

The 96 items were preceded by an additional 8 items (four words/four non-words). These 8 items were presented prior to the 96 items to attenuate a possible recency effect over the last words from the preceding value-learning task. These four words were not presented in the value-learning task and performance on these 8 items was not included in the analyses.

2.1.3.3. Free recall. In a surprise free recall task, participants were given 5 min to recall all of the words they could remember from the task. Participants were given a maximum of 45 s for each response. After each response, a blank screen was presented for 500 ms. Misspellings or variants of the correct word were scored as incorrect responses. Repetitions of correct responses were ignored.

2.1.3.4. Value judgment. To obtain an objective measure of a participant's explicitly learned reward value information, we included a value judgment task. At the end of the experiment, participants were presented with each of the words previously shown in the reward training phase, one at a time, and asked to judge how many points each word was worth from the initial value-learning task. Participants were reminded of the three possible reward levels and asked to type the number of points they thought the presented word was worth.

2.2. Data analysis

All analyses are reported with Greenhouse–Geisser correction for non-sphericity where appropriate. Effects were considered significant based on an alpha level of 0.05 and post-hoc pairwise comparisons were always Bonferroni-corrected. Non-significant 'trend' effects ($p<.1$) are also reported.

For response time analyses, only correct responses were analyzed. In the lexical decision phase, only responses made between 200 ms and the individual participant's mean plus three standard deviations were included in the analysis. As response time distributions were not normally distributed, response times were log-transformed to accommodate parametric statistics.

As our hypotheses concern the relationship between reward level and memory, we tested for significant linear and quadratic effects of reward level. The reward value hypothesis would suggest a significant linear component. On the other hand, the reward salience hypothesis would suggest a significant quadratic component, corresponding to a U-shaped relationship.

2.3. Results

2.3.1. Value learning

Accuracy in the value-learning task was measured as how often the participant chose the higher-value word. Performance in the training task began at chance, as the participant could not know which was the higher-value word. In the last round of the training task, accuracy was significantly greater than chance [$M=77.9\%$ correct; $t(70)=15.20$, $p<.001$]. In choice sets, the difference in values between the items within the set (V_{diff}) could be either 5 points (choice between a 2-point item and a 7-point item, or a 7-point item and a 12-point item) or 10 points (choice between a 2-point item and a 12-point item). At both value differences, participants were more likely than chance to choose the higher-value item [$V_{diff}(5)$: $M=74.4\%$, $t(70)=12.00$, $p<.001$; $V_{diff}(10)$: $M=85.0\%$, $t(70)=16.72$, $p<.001$] (see Fig. 1a).

To directly examine if higher-value items were chosen more, we also tested for a relationship between choice frequency (the number of times an item was chosen) and reward level. A one-factor repeated-measures ANOVA, revealed a main effect of reward level on choice [$F(2,140)=204.05$, $p<.001$]. Pairwise comparisons found all differences to be significant, such that 2-point words<7-point words<12-point words [all p's < .001] (Fig. 2a).

To keep analyses comparable between our value-learning and memory measures, we also tested for significant linear and quadratic components in the relationship between choice and reward level. We found a significant linear component of reward level [$F(1,70)=321.37$, $p<.001$]. The quadratic component was not significant [$F(1,70)=1.51$].

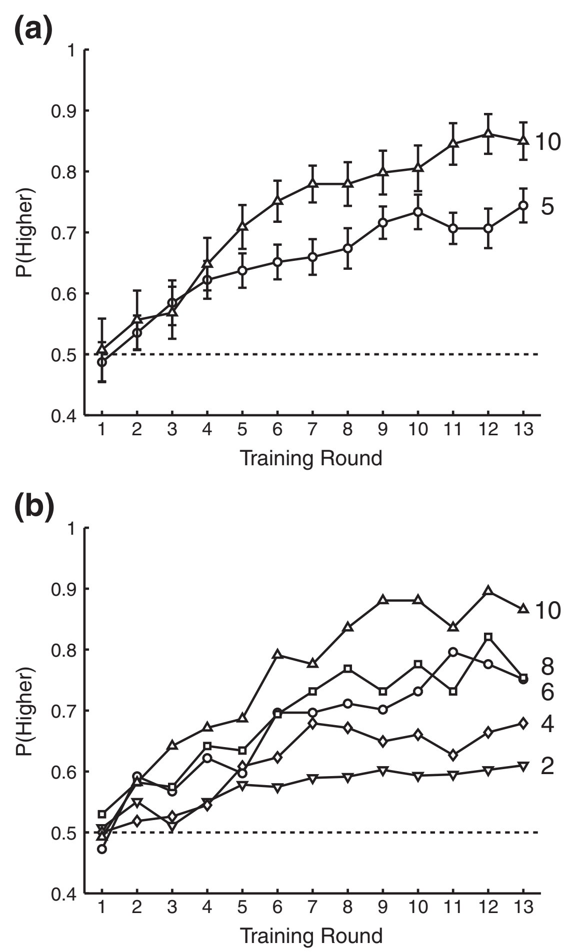

Fig. 1. Percentage of trials in which the higher-value item was chosen in each round of the value-learning task, separated based on the difference in reward value between the two items, for (a) Experiment 1 and (b) Experiment 2. Error bars are 95% confidence intervals, corrected for inter-individual differences. Due to the number of reward levels, error bars were omitted from panel (b) for visual clarity.

2.3.2. Lexical decision

Lexical decision accuracy was near ceiling, as expected, and was not significantly different between value conditions [$F(2,140) = 0.11$]. Accuracy was higher for the previously rewarded 'old' words than for the untrained, 'new' words ['old' words: 98.2% correct; 'new' words: 91.3% correct; $t(70) = 5.11$, $p < .001$].

Participants also identified the 'old' words significantly faster than the 'new' words [$t(70) = 5.75$, $p < .001$]. No significant differences were found between lexical decision response times across the three reward levels [$F(2,140) = 1.97$], although we found a trend quadratic effect of reward level [$F(1,70) = 3.05$, $p < .10$]. The linear component was not significant [$F(1,70) = 0.93$].

2.3.3. Free recall

As expected, participants recalled fewer of the untrained, new words than the previously rewarded words (Fig. 2b). For the previously rewarded words we conducted a one-factor repeated-measures ANOVA (reward level) and found a main effect [$F(2,140) = 11.87$, $p < .001$]. Pairwise comparisons found that significantly more 12-point words were remembered than 2-point and 7-point words [both $ps < .01$]. However, there was no significant difference between recall performance for 2-point and 7-point words [$t(70) = 1.44$]. We also found a significant linear effect of reward level [$F(1,70) = 9.80$, $p < .01$], as well as a significant quadratic effect [$F(1,70) = 14.21$, $p < .001$].

2.3.4. Value judgment

Participants correctly identified the value of the previously rewarded words at levels better than chance (33.3%) [$M = 61.4\%$ correct; $t(70) = 15.03$, $p < .001$]. To illustrate the responses in this task, we plotted the proportion of value judgment responses for each reward level, separated based on the actual reward level of the item (Fig. 3a).

To further examine the relation between memory and value, we compared accuracy on the value judgment task for words that were recalled in free recall relative to words that were not recalled. Value was found to be judged better for words that were recalled than for words that were not recalled [$M_{recalled} = 64.4\%$ correct; $M_{not\text{-}recalled} = 57.8\%$ correct; $t(70) = 3.00$, $p < .01$]. This supports the notion that value is learned as a property of the words themselves, and memory for value is less accurate when the items are also not remembered as well.

2.4. Discussion

Our finding of a significant quadratic component in the effects of reward level on free recall performance is consistent with the reward salience hypothesis. However, with only three reward levels, we are unable to statistically differentiate between linear and quadratic effects and conclusively state that one effect is more prominent than the other. Experiment 2 addressed this issue by using *six* different reward levels (2, 4, 6, 8, 10, and 12 points). The greater number of separable reward levels provided better resolution of reward as a continuous measure and allowed better comparison of the relative fits of both linear and quadratic models.

3. Experiment 2

3.1. Methods

3.1.1. Participants

67 introductory psychology students (48 females) at the University of Alberta participated for partial fulfillment of course credit. Restrictions for participating and informed consent were identical to

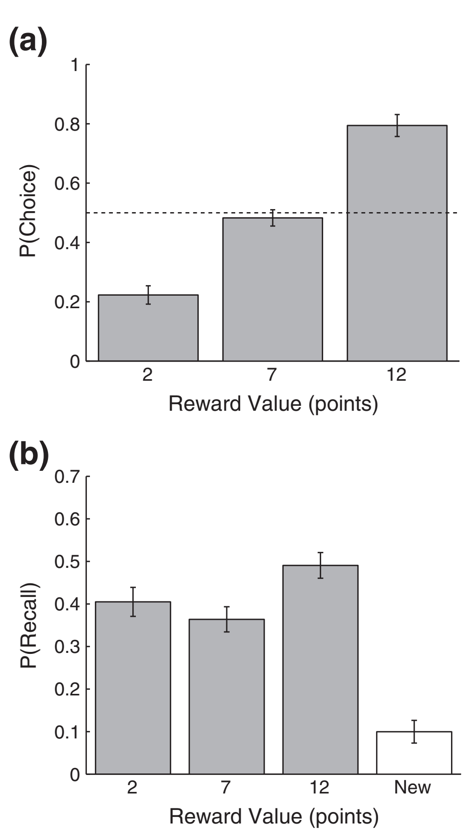

Fig. 2. Performance as a function of reward value for (a) choice frequency in the last three rounds of value-learning and (b) recall rates in free recall, in Experiment 1. Error bars are 95% confidence intervals, corrected for inter-individual differences.

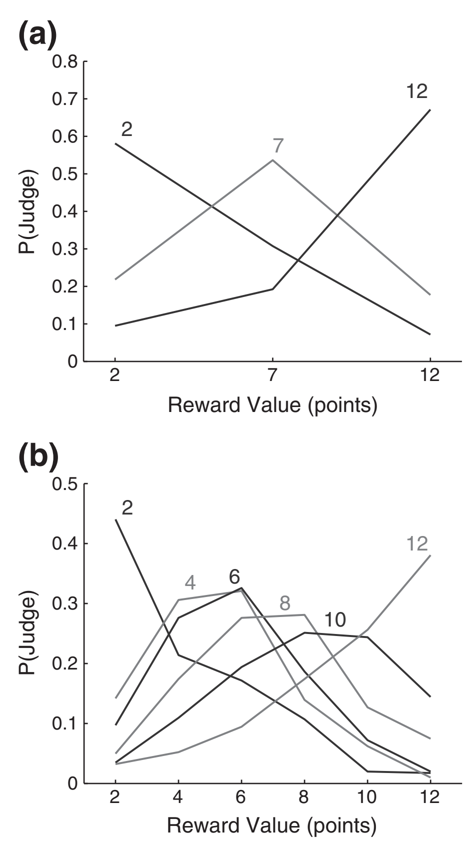

Fig. 3. Proportion of value judgment responses for each reward level, separated based on the actual reward level of the item, for (a) Experiment 1 and (b) Experiment 2. Error bars were omitted from for visual clarity.

Experiment 1. None of the participants from Experiment 1 participated in Experiment 2.

3.1.2. Materials

The same word and non-word pools were used as in Experiment 1.

3.1.3. Procedure

Instead of the three different reward levels used in Experiment 1, here we used *six* reward levels: 2, 4, 6, 8, 10, and 12 points. To keep the total number of previously rewarded words at 36, only six words were assigned to each reward level. All participants earned between $4.00 and $5.00 in the value-learning task. Additionally, only 6 new words were used in lexical decision task, rather than 12, to equate the number of new words with the words in any single reward level. The rest of the procedure remained the same as in Experiment 1.

3.2. Results

3.2.1. Value learning

In the last round of the training task, accuracy was significantly greater than chance [$M = 67.9\%$ correct; $t(66) = 9.46$, $p < .001$]. In choice sets, the difference in values for each item (V_{diff}) could be 2, 4, 6, 8, or 10 points. At all value differences, participants were more likely than chance to choose the higher-value item [$V_{diff}(2)$: $M = 61.0\%$, $t(66) = 5.32$, $p < .001$; $V_{diff}(4)$: $M = 67.9\%$, $t(66) = 5.79$, $p < .001$; $V_{diff}(6)$: $M = 75.1\%$, $t(66) = 7.86$, $p < .001$; $V_{diff}(8)$: $M = 75.4\%$, $t(66) = 7.08$, $p < .001$; $V_{diff}(10)$: $M = 86.6\%$, $t(66) = 8.71$, $p < .001$] (see Fig. 1b).

A one-factor repeated-measures ANOVA revealed a main effect of reward level on choice [$F(4,264) = 59.39$, $p < .001$]. Pairwise comparisons found all differences to be significant [all $ps < .001$], except for the difference between 4-point and 6-point words (Fig. 4a). Similar to Experiment 1, we found a significant linear component of reward level [$F(1,66) = 158.62$, $p < .001$]. The quadratic component was again not significant [$F(1,66) = 2.48$].

3.2.2. Lexical decision

Lexical decision accuracy was near ceiling, as expected, and did not differ significantly between value conditions [$F(4,269) = 0.56$]. Accuracy was higher for the previously rewarded 'old' words than for the untrained, 'new' words ['old' words: 98.8% correct; 'new' words: 94.0% correct; $t(66) = 4.21$, $p < .001$].

Participants also identified the 'old' words significantly faster than the 'new' words [$t(66) = 6.73$, $p < .001$]. No significant differences were found between lexical decision response times across the six reward levels [$F(4,250) = 1.37$]. We found neither a significant linear effect of reward level [$F(1,66) = 0.55$] nor a quadratic effect [$F(1,66) = 0.75$].

3.2.3. Free recall

Recall was again higher for the previously rewarded words than for the untrained, new words (Fig. 4b). For the previously rewarded words, we again conducted a one-factor repeated-measures ANOVA (reward level) and found a main effect [$F(5,305) = 8.42$, $p < .001$]. Pairwise comparisons revealed no significant differences between 2-, 10-, and 12-point words [$p > .5$]. Importantly, more 2-point words were recalled than 4-point words [$p < .05$], and more 12-point words were recalled than 4-, 6-, or 8-point words [all $ps < .001$]. We also found a significant linear effect of reward level [$F(1,66) = 12.27$, $p < .01$], as well as a significant quadratic effect [$F(1,66) = 27.37$, $p < .001$].

To better test our hypotheses, we fit constant, reward value (linear), reward salience (quadratic), and reward value + salience (linear and quadratic) models to our free recall data. The constant model assumed that reward had no effect on memory performance and thus that recall should be equivalent regardless of reward level and only contained one parameter. The reward value model assumed that recall performance is monotonically related to reward value (e.g., higher-value items are remembered better than lower-value items). This model contained two parameters: a term that varied monotonically with reward value and a constant. The reward salience model assumed that recall performance follows a U-shaped (quadratic) function with relation to reward value. In other words, high-salience items (those at the extremes of the range of values experienced) are remembered better than low-salience items (those in the middle of the range). This model contained two free parameters: a term that varied quadratically with reward value and a constant. This function was shifted along the x-axis such that the minima of the function occurred at the middle of the range of values experienced (i.e., 7 points). The reward value + salience model was based on the reward salience model, but additionally incorporated a parameter that shifted the function to best overlap with the data (i.e., the reward value where function's minima occurred). The reward value + salience model therefore allows reward value, in addition to reward salience, to be a predictor of item-memory.

To compare the different models, we calculated the *BIC* (Bayesian Information Criterion) as a measure of model fitness that takes into account the number of free parameters (see Table 1). By convention, if the difference between two model fits is less than two, neither of the models' fit to the data is significantly better—thus we report all scores as *ΔBIC* relative to the best-fitting model (Burnham & Anderson, 2002). We also report R^2 and *RMSD* (root mean squared deviation) as additional measures of model fitness. Note that a best fitting model would be characterized by low *ΔBIC*, low *RMSD*, and high R^2 values. As evident in Fig. 5, the reward value + salience model is the best-fitting model, a conclusion that was supported by all three of our quantitative model fitting measures ($R^2 = .93$). More specifically, it appears that this model was primarily characterized by reward salience ($R^2 = .70$), along with a much smaller contribution from reward value ($R^2 = .22$).

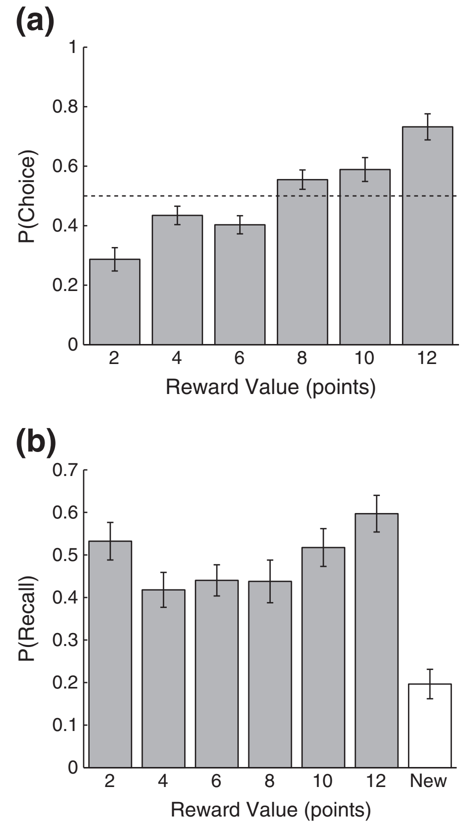

Fig. 4. Performance as a function of reward value for (a) choice frequency in the last three rounds of value-learning and (b) recall rates in free recall, in Experiment 2. Error bars are 95% confidence intervals, corrected for inter-individual differences.

Table 1
Fitness measures for the model fits to the free recall data of Experiment 2.

Model	*df*	*ΔBIC*	R^2	*RMSD*
Constant	1	4.25	.00	.064
Reward value	2	5.28	.22	.056
Reward salience	2	2.41	.70	.035
Reward value + salience	3	0	.93	.017

3.2.4. Value judgment

Participants correctly identified the value of the previously rewarded words at levels better than chance (16.7%) [$M = 33.0\%$ correct; $t(66) = 11.84$, $p < .001$]. As in Experiment 1, we plotted the proportion of value judgment responses for each reward level, separated based on the actual reward level of the item (Fig. 3b). Here it is evident that even when participants were incorrect at judging the value, they often responded with an adjacent value (e.g., judging a 6-point item as being worth either 4- or 8-points). We also observe a large degree of overlap in value judgment responses for items worth 4-, 6-, and 8-points, converging with the clustering observed in memory performance (Fig. 4b).

As in Experiment 1, we compared accuracy on the value judgment task for words that were recalled in free recall relative to words that were not recalled. Value was found to be judged better for words that were recalled than for words that were not recalled [$M_{recalled} = 37.0\%$ correct; $M_{not\text{-}recalled} = 27.5\%$ correct; $t(66) = 5.28$, $p < .001$]. Here we again find evidence that our procedure may cause value to be learned as an attribute of the word itself.

3.2.5. Ruling out simple decision heuristics

One possible explanation for a U-shaped function is that participants relied on simple decision heuristics of choosing the highest-valued item and avoiding the lowest-valued item. However, two sources of evidence argue against the possibility that participants relied primarily on these simple heuristics. First, the value judgment results presented in Fig. 3b show that even for the intermediate items, the value was discriminated from some of the other intermediate items and not just from the two extreme items (i.e., participants rarely judged a 4-point item as having a value of 10 points or vice versa). Second, for the 2-point difference decisions, reliance on a simple choose/avoid heuristic would result in highly accurate choices for 2-point versus 4-point and for 10-point versus 12-point choice sets, but much lower accuracy for the intermediate choice sets, such as 6-point versus 8-point. To test this empirically, we conducted a one-factor repeated-measures ANOVA on the compared choice sets for the last four rounds of training (i.e., when participants had largely learned the values). Our dependent measure was the proportion of choices for which participants correctly chose the higher-value item when the difference in value of the two items was only 2-points. We did not find a main effect, suggesting that there were no significant differences in accuracy between choice sets [$F(3,229) = 2.42$] (Fig. 6). Thus, even though the words associated with extreme values were remembered better, this enhancement did not appear to be caused by a simple decision heuristic based on only the extreme values.

4. General discussion

In the present study we provide clear evidence that memory for items does not correspond monotonically with increases in reward value. Instead, we find that the most and least rewarding items are remembered best and that memory follows a U-shaped function (see Fig. 5), suggesting an enhancement of memory driven by reward salience. Thus, we provide evidence that free recall performance is not driven *solely* by reward value or the number of times an item was chosen (choice frequency), as these accounts would both correspond to a monotonic (linear) relationship between reward value and memory.

Studies have shown that positive and negative values are learned through different neural substrates (e.g., Yacubian et al., 2006). In the present study, however, we found an enhancement of memory for extreme values that did *not* necessitate the presence of both positive and negative values, as all of our values were positive. That is, the lowest value items were remembered better than intermediate-level items, even though our lowest value items were near zero in absolute value. Our results therefore suggest that reward salience is relative to the *range of values experienced*, and is not necessarily driven by the use of positive and negative values (e.g., Cooper & Knutson, 2008) or appetitive and aversive stimuli (e.g., Litt et al., 2011). Although the significant linear effect of reward value on free recall performance suggests that increases in value do enhance memory, this effect was much weaker than the quadratic component and explained less of the variance. Thus, memory in free recall was explained better by the reward salience of the stimuli than by their reward value alone.

Our evidence suggesting that reward salience is driven by the range of values experienced, may also be related to another psychological mechanism: the anchoring effect. The anchoring effect suggests that the ends of the stimulus continuum play an important role in judgments of absolute value (e.g., Eriksen & Hake, 1957) and that the anchoring effect is driven by memory for the end-points (Petrov & Anderson, 2005; Weber & Johnson, 2006). Anchoring effects could also be at play in our value judgment task (see Fig. 3). While it can be argued that temporal effects (i.e., primacy and recency)

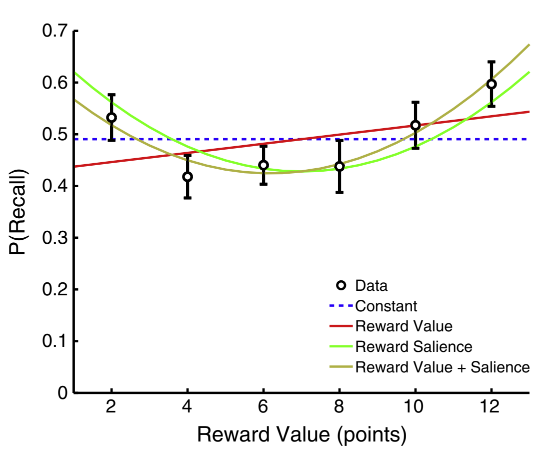

Fig. 5. Model fits to the free recall data from Experiment 2 (see Table 1 for model fitness measures). Error bars are 95% confidence intervals, corrected for inter-individual differences.

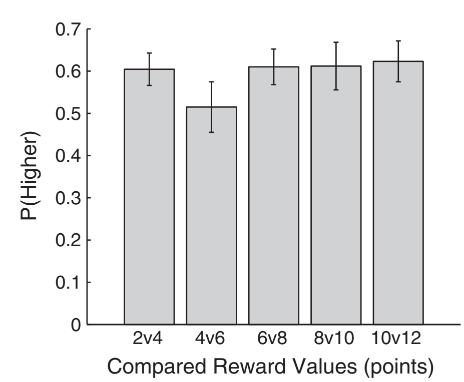

Fig. 6. Accuracy across the last four rounds of the value-learning task in Experiment 2, for choice sets where the difference in value was only 2-points. Error bars are 95% confidence intervals, corrected for inter-individual differences.

in memory are the result of anchoring effects, they have not previously been considered within a value-learning procedure. Thus, it is possible that reward salience is intertwined with the anchoring effect.

Though prior studies have utilized fMRI in conjunction with reward-based memory paradigms, none of these studies utilized multiple levels of reward and are thus unable to test a reward saliency hypothesis. These studies have implicated the activation of several reward-related brain regions as predictors of which items will be later remembered, including the ventral tegmental area, striatum, substantia nigra, and orbitofrontal cortex (Adcock et al., 2006; Shigemune et al., 2010; Wittmann et al., 2005). Of these regions, the orbitofrontal cortex has been related to reward value (Jensen et al., 2007; Litt et al., 2011). However, activation in the striatum has been associated with either reward salience (Cooper & Knutson, 2008; Jensen et al., 2007; Zink et al., 2004) or both reward value and reward salience (Litt et al., 2011). Currently, none of the fMRI studies investigating reward salience mention the substantia nigra or the ventral tegmental area. Based on the results of studies investigating reward salience, the striatum appears to be the reward-related brain region most likely to be responsible for the memory results in the present study. Specifically, activations in the striatum correspond to a combination of reward value and reward salience (e.g., see Fig. 4c of Litt et al., 2011), with a stronger contribution of reward salience, similar to the results of our model fit. Nonetheless, the neural underpinnings of our result are open to debate and will likely be the focus of future research.

Additionally, although we found an effect of reward on explicit memory in our free recall task, we did not observe an effect of reward on implicit memory in our lexical decision task. Although an enhancement of memory due to reward in lexical decision has been reported previously (Madan et al., under review), the current study, which had fewer data points per reward level for each participant, may have been less sensitive to subtle enhancements of implicit priming due to reward. Nonetheless, we did observe a trend quadratic effect of reward on lexical decision response times in Experiment 1, suggesting that the enhancement of implicit memory due to reward seen in the previous study may also be driven in part by reward salience.

5. Conclusion

In two experiments we demonstrate that the enhancement of memory due to reward is driven not only by reward value, but also by reward salience. Most previous studies that suggested a monotonic influence of reward value on memory used only two levels of reward and thus were unable to capture the full relationship between reward and memory. Through the addition of intermediate reward levels we are able to determine that memory is enhanced for both the highest- and lowest-value items. This U-shaped (quadratic) relationship between value and memory occurred even though all reward values were positive and is best characterized as an enhancement of memory due to reward salience.

Acknowledgments

We would like to thank Elliot Ludvig and Eric Legge for insightful feedback in the design and analysis of this study, as well as Esther Fujiwara and Jeremy Caplan for feedback on an earlier draft of the manuscript. This research was partly funded by a Discovery grant from the National Science and Engineering Research Council of Canada held by MLS and by a scholarship from the Canadian Institutes of Health Research to CRM.

References

Adcock, R. A., Thangavel, A., Whitfield-Gabrieli, S., Knutson, B., & Gabrieli, J. D. E. (2006). Reward-motivated learning: Mesolimbic activation precedes memory formation. *Neuron*, *50*, 507–517.

Bradley, M. M., & Lang, P. J. (1999). Affective norms for English words (ANEW): Stimuli, instruction manual and affective ratings (Technical report No. C-1). In F. L. Gainesville (Ed.), *The Center for Research in Psychophysiology*. University of Florida.

Burnham, K. E., & Anderson, D. R. (2002). *Model selection and multimodel inference* (2nd ed.). New York: Springer-Verlag.

Castel, A. D., Benjamin, A. S., Craik, F. I. M., & Watkins, M. J. (2002). The effects of aging on selectivity and control in short-term recall. *Memory & Cognition*, *30*, 1078–1085.

Cooper, J. C., & Knutson, B. (2008). Valence and salience contribute to nucleus accumbens activation. *NeuroImage*, *39*, 538–547.

Eriksen, C. W., & Hake, H. W. (1957). Anchor effects in absolute judgments. *Journal of Experimental Psychology*, *53*, 132–138.

Gruber, M. J., & Otten, L. J. (2010). Voluntary control over prestimulus activity related to encoding. *Journal of Neuroscience*, *30*, 9793–9800.

Jensen, J., Smith, A. J., Willeit, M., Crawley, A. P., Mikulis, D. J., Vitcu, I., et al. (2007). Separate brain regions code for salience vs. valence during reward prediction in humans. *Human Brain Mapping*, *28*, 294–302.

Lisman, J. E., & Grace, A. A. (2005). The hippocampal-VTA loop: Controlling the entry of information into long-term memory. *Neuron*, *46*, 703–713.

Litt, A., Plassmann, H., Shiv, B., & Rangel, A. (2011). Dissociating valuation and saliency signals during decision-making. *Cerebral Cortex*, *21*, 95–102.

Ludvig, E. A., Bellemare, M. G., & Pearson, K. G. (2011). A primer on reinforcement learning in the brain: Psychological, computational, and neural perspectives. In E. Alonso, & E. Mondragon (Eds.), *Computational neuroscience for advancing artificial intelligence: Models, methods and applications* (pp. 111–144). Hershey, PA: IGI Global.

Madan, C. R., Fujiwara, E., & Caplan, J. B. (under review). Reward value separately enhances implicit and explicit memory.

Pessiglione, M., Seymour, B., Flandin, G., Dolan, R. J., & Frith, C. D. (2006). Dopamine-dependent prediction errors underpin reward-seeking behaviour in humans. *Nature*, *442*, 1042–1045.

Petrov, A. A., & Anderson, J. R. (2005). The dynamics of scaling: A memory-based anchor model of category rating and absolute identification. *Psychological Review*, *112*, 383–416.

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* (pp. 64–99). New York: Appleton-Century-Crofts.

Shigemune, Y., Abe, N., Suzuki, M., Ueno, A., Mori, E., Tashiro, M., et al. (2010). Effects of emotion and reward motivation on neural correlates of episodic memory encoding: A PET study. *Neuroscience Research*, *67*, 72–79.

Shohamy, D., & Adcock, R. A. (2010). Dopamine and adaptive memory. *Trends in Cognitive Science*, *14*, 464–472.

Soderstrom, N. C., & McCabe, D. P. (2011). The interplay between value and relatedness as bases for metacognitive monitoring and control: Evidence for agenda-based monitoring. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *37*, 1236–1242.

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

Weber, E. U., & Johnson, E. J. (2006). Constructing preferences from memory. In S. Lichtenstein, & P. Slovic (Eds.), *The construction of preference* (pp. 397–410). New York: Cambridge University Press.

Westbury, C., Hollis, G., & Shaoul, C. (2007). LINGUA: The Language-Independent Neighbourhood Generator of the University of Alberta. *The Mental Lexicon*, *2*, 273–286.

Wilson, M. D. (1988). The MRC psycholinguistic database: Machine readable dictionary, version 2. *Behavioural Research Methods, Instruments and Computers*, *20*, 6–11.

Wittmann, B. C., Schott, B. H., Guderian, S., Frey, J. U., Heinze, H. J., & Düzel, E. (2005). Reward-related fMRI activation of dopaminergic midbrain is associated with enhanced hippocampus-dependent long-term memory formation. *Neuron*, *45*, 459–467.

Yacubian, J., Gläscher, J., Schroeder, K., Sommer, T., Braus, D. F., & Büchel, C. (2006). Dissociable systems for gain and loss related value predictions and errors of prediction in the human brain. *Journal of Neuroscience*, *26*, 9530–9537.

Zink, C. F., Pagnoni, G., Martin-Skurski, M. E., Chappelow, J. C., & Berns, G. S. (2004). Human striatal response to monetary reward depend on saliency. *Neuron*, *42*, 509–517.