



Sluggish retrieval of positive memories in depressed adults

Arkadiy L. Maksimovskiy^{1,2} · Christopher Okine³ · Andrea M. Cataldo^{1,2} · Daniel G. Dillon^{1,2}

Accepted: 28 April 2022 / Published online: 12 May 2022
© The Psychonomic Society, Inc. 2022

Abstract

Although depression is associated with poor memory for positive material, the underlying mechanisms remain unclear. We used the Hierarchical Drift Diffusion Model (HDDM) to determine whether slow evidence accumulation at retrieval contributes to depressed individuals' difficulty remembering positive events. Participants completed the Beck Depression Inventory-II and were stratified into High BDI (HBDI; BDI-II > 20, $n = 49$) and Low BDI (LBDI; BDI-II < 6, $n = 46$) groups. Next, participants completed an oddball task in which neutral, negative, and positive pictures served as rare targets. One day later, recognition memory was tested by presenting the encoded ("old") pictures along with closely matched ("new") lures. Recognition accuracy was analyzed with a generalized linear model, and choice and response time data were analyzed with the HDDM. Recognition accuracy for old positive pictures was lower in HBDI versus LBDI participants, and the HDDM highlighted slow evidence accumulation during positive memory retrieval in the HBDI group. Impaired memory for positive material in depressed adults was related to slow evidence accumulation at retrieval. Because oddballs should elicit prediction errors that normally strengthen memory formation, these retrieval findings may reflect weak positive prediction errors, at encoding, in depressed adults.

Keywords Depression · Memory · Emotion · Oddball · Drift Diffusion Model

Introduction

Depression affects more than 264 million adults annually (James et al., 2018), and it is associated with a broad range of cognitive and emotional problems (Gotlib & Joormann, 2010). In particular, depressed adults often present with impaired episodic memory (Rock et al., 2014; Zakzanis et al., 1998). For instance, relative to nondepressed individuals, depressed adults have displayed overgeneral autobiographical memory retrieval (Dalgleish et al., 2007; Hallford et al., 2021) and poorer recollection (MacQueen et al., 2002; Ramponi, Barnard, and Nimmo-Smith, 2004). Most importantly for the current study, meta-analyses indicate that depressed adults often show relatively poor memory for positive material, whereas memory for negative material

is unaffected or even enhanced (Burt et al., 1995; Matt et al., 1992).

Difficulty retrieving positive memories is clinically relevant, because it may weaken an individual's ability to regulate their emotions and interrupt ongoing depressive states (Dalgleish & Werner-Seidler, 2014; Joormann & Siemer, 2004). If the mechanisms underlying the positive memory deficit were well-characterized, they could be targeted for treatment. Unfortunately, however, the relevant psychological and neural processes remain poorly understood.

We proposed that poor memory for positive events in depression reflects suboptimal encoding (Dillon, 2015; Dillon & Pizzagalli, 2018). This hypothesis rests on several observations, the first of which is of growing behavioral evidence, which shows that depressed adults perform poorly on reward-processing tasks compared with their nondepressed counterparts (Halahakoon et al., 2020 review and a meta-analysis). Recent studies also show that this reward task performance deficit is driven by dysfunctional dopaminergic circuits (Pizzagalli, 2014; Proudfit, 2015; Treadway and Zald, 2011). Furthermore, there is strong evidence that dopamine (DA) neurons within dopaminergic circuits signal positive prediction errors (PPEs)—that is, they fire strongly when an event is better than expected (Schultz,

✉ Arkadiy L. Maksimovskiy
amaksimovskiy@mclean.harvard.edu

¹ Center for Depression, Anxiety and Stress Research, McLean Hospital, Belmont, MA, USA

² Department of Psychiatry, Harvard Medical School, Boston, MA, USA

³ Harvard University, Cambridge, MA, USA

2016). Findings from animal (Bethus et al., 2010; Lisman and Grace, 2005) and healthy adult (Jang et al., 2019) studies suggest that PPEs strengthen memory formation for positive events. Taken together, these observations form the cornerstone of our hypothesis; namely, that depressed adults would generate weak dopaminergic PPEs in response to positive events, which would result in poor encoding of those events into long-term memory.

Supporting this hypothesis, we used functional magnetic resonance imaging (fMRI) to show that responses within the ventral tegmental area/substantia nigra (VTA/SN)—where DA neurons are most densely concentrated (Haber & Knutson, 2010)—positively predicted memory for rewarded stimuli in nondepressed controls but not in adults with Major Depressive Disorder (MDD) (Dillon et al., 2014). The fact that VTA/SN activation was weaker in participants with MDD, and that this effect helped explain the group difference (controls > MDD) in memory for rewarded stimuli, supports the hypothesis that weak PPEs, signaled by DA neurons, play a causal role in the positive memory deficit seen in MDD.

However, this prior study (Dillon et al., 2014) had two key limitations. First, a reward or nonreward outcome was delivered on every trial, and the two kinds of feedback were equiprobable (i.e., half of the encoding trials were rewarded). This is not ideal because PPE magnitude depends partly on surprise; it would be better to present rewarding stimuli less frequently, as rare deviations from a predictable sequence of stimuli. Such an arrangement should elicit robust PPEs and enhance memory in nondepressed controls, facilitating detection of the negative effects of blunted PPEs on memory in depressed adults. Second, participants in our prior study knew that their memories would be tested, raising the possibility of group differences in encoding strategies. Administering a surprise memory test would avoid this problem.

To address these concerns and provide additional insight into mechanisms underlying poor memory for positive material in depression, we designed an oddball task followed by a recognition memory test and administered it to adults with high versus low depressive symptoms. At encoding, participants viewed an equal number of positive, negative, and neutral pictures presented as rare targets (“oddballs”) amidst a stream of rectangles (“standards”). Because the pictures appeared infrequently and unpredictably, we expected them to elicit prediction errors (PEs) whose signs should depend on the picture type: positive pictures should elicit PPEs, and negative pictures should elicit negative PEs (because of their neutral content, neutral pictures should elicit smaller and more weakly signed PEs than valenced pictures). Because depression weakens PPEs, we expected that positive oddballs—but not negative or neutral oddballs—would be poorly encoded in depressed adults. To test this prediction, we administered a recognition memory test for the oddballs

one day later. Participants were not forewarned about the test, which was delayed by 24 hours, because the effects of DA release on memory may be consolidation-dependent (Bethus et al., 2010; but see Jang et al., 2019). Our primary hypothesis was that, relative to participants with low depressive symptoms, the more severely depressed adults would show poorer memory for positive pictures.

Our intention was to acquire electroencephalography (EEG) data at encoding so that we could use the P300 event-related potential as a marker of PE strength (Polich, 2007; Donchin and Coles, 1988). Unfortunately, however, the COVID19 pandemic precluded this plan. We thus conducted the study online, which presented a problem: the oddball task is very simple, and without EEG recordings there was no way to quantify PPEs at encoding. Therefore, we focused on analysis of the recognition memory data. Our rationale was that if PPEs elicited in the oddball task enhanced the encoding of positive pictures more strongly in nondepressed versus depressed adults, then it should be possible to detect the downstream consequences of these effects in participants’ recognition memory performance.

Specifically, we applied the Drift Diffusion Model (DDM) (Ratcliff and McKoon, 2008) to the recognition data. The DDM is a useful tool for breaking retrieval down into distinct cognitive processes (Ratcliff, 1978), and it has successfully detected negative effects of depression on cognition in prior studies (Lawlor et al., 2020; White et al., 2009, 2010). As illustrated in Fig. 1, the DDM quantifies cognitive processes that support memory retrieval and that cannot be detected in standard analysis of mean accuracy and response times (RTs) (Ratcliff, 1978; Wiecki et al., 2013). The model is based on the assumption that the retrieval process consists of accumulating evidence (e.g., from memory networks in the brain) that is necessary to decide whether an item was seen before (is “old,” from encoding) or not (is “new,” a lure). Individuals arrive at their decision by accumulating evidence over time from an initial starting point, z , towards one of two response boundaries that represent possible responses (“old” or “new”). The starting point, z , ranges from 0–1; values <0.5 indicate a bias towards the lower boundary (“new” in Fig. 1) and values >0.5 reflect a bias towards the upper boundary (“old” in Fig. 1). The distance between the two boundaries is represented by the parameter a , which indicates the amount of evidence that needs to accumulate before reaching a decision. Although evidence is assumed to accumulate noisily on a trial-by-trial basis (e.g., the green or blue lines in Fig. 1), the average rate of accumulation across trials (the “drift rate,” e.g., the solid black line in Fig. 1) is represented by the parameter v . The magnitude of the drift rate indicates the speed with which evidence is accumulated, such that higher absolute values reflect faster and more efficient evidence accumulation. The sign

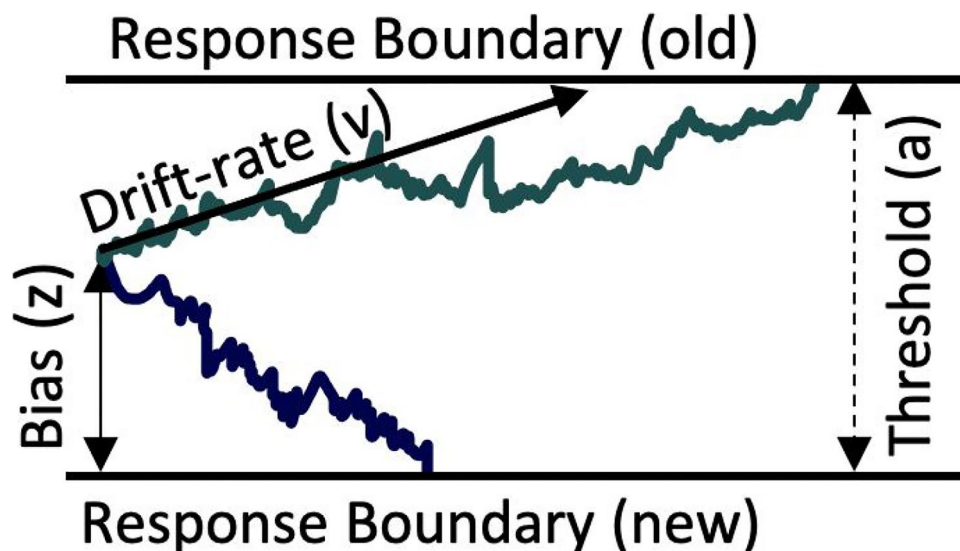


Fig. 1 Drift diffusion model. *Note.* Drift Diffusion Model (DDM) used to model recognition data. The model includes two response boundaries: an upper boundary corresponding to “old” responses and a lower boundary for “new” responses. Starting bias (z) quantifies whether a participant is more likely to respond “old” or “new” before the beginning of each trial. Drift-rate (v) captures the speed of evi-

dence accumulation: the average time it takes to reach a boundary on each trial. This example shows relatively fast evidence accumulation towards the new boundary (dark blue line), and relatively slow evidence accumulation towards the old boundary (green line). The decision threshold (a) quantifies the distance between the two boundaries

of v (negative or positive) indicates the direction towards which evidence is being accumulated (e.g., either negatively towards “new” or positively towards “old” in Fig. 1). Finally, the t_0 parameter represents the time allotted to all “nondecision” processes in which participants engage outside of accumulating evidence for a given choice (e.g., stimulus perception and motor functions). By modeling the decision process dynamically, the DDM jointly estimates both choice proportions (represented by the proportion of trials in which the upper vs. lower boundary was reached) and response times (represented by the amount of time taken to reach a boundary), thus providing a well-constrained model of decision-making as it unfolds over time.

The DDM is helpful in the current context because one model parameter is especially sensitive to differences in encoding quality. Specifically, manipulations that influence encoding strength primarily affect the drift rate at recognition while leaving the other DDM parameters unaffected (Criss, 2010). Consequently, if depression blunts the PPEs elicited by positive oddballs at encoding, then that should result in a lower drift rate when those stimuli are presented again at recognition. Consistent with this argument, a prior study reported that nondysphoric students showed higher drift rates for positive vs. negative words presented in a recognition test, but this effect was not present in dysphoric students (White et al., 2009). This prior study (White et al., 2009) is thus encouraging, and we sought to build upon it by testing our hypothesis about prediction errors in samples drawn from the adult population.

To this end, we used the Beck Depressive Inventory-II (BDI-II; Beck et al., 1996) to identify 49 adults with high BDI-II scores (HBDI group) and 46 with low BDI-II scores (LBDI group). These two groups then completed the oddball task and the memory test using an online platform. We predicted that memory for positive oddballs would be worse in the HBDI versus LBDI group. We used the Hierarchical Drift Diffusion Model (HDDM; Wiecki et al., 2013), a Bayesian implementation of the DDM, to test the prediction that that this deficit would be driven by slow accumulation of positive memory evidence (i.e., lower drift-rates for “old” positive images) in the HBDI group at retrieval. As described above, we hypothesized that this pattern would emerge as a downstream consequence of weak PPEs in the HBDI group during encoding.

Materials and methods

Participants

Participants were recruited through the Harvard SONA system ($n = 95$; Table 1). Prospective participants were eligible if they were fluent English speakers between ages 18 and 55 years, with BDI-II scores less than 6 (LBDI) or greater than 20 (HBDI). We chose 20 as the high cutoff score, because this is the recommended threshold for identifying adults with at least moderate depression (Beck et al., 1996). We chose 6 as the low cutoff, because to meet this criterion,

Table 1 Demographic and self-report data

Group	BDI-II (mean \pm SD)	Gender	Age (mean \pm SD)	Education (mean \pm SD)
HBDI (N = 49)	29.92 \pm 6.55	Female: 39 Male: 8 Other: 2	23.69 \pm 6.27	12.83 \pm 5.38
LBDI (N = 46)	2.3 \pm 1.88	Female: 36 Male: 10 Other: 0	27.17 \pm 8.91	15.98 \pm 3.21

BDI-II: Beck Depression Inventory-II (Beck et al., 1996); *HBDI*: High BDI-II (>20) group; *LBDI*: Low BDI-II (<6) group

adults would have to have minimal depressive symptoms such that the two groups would be distinct (i.e., no participants with mild symptoms). While it would be useful to study behavior across the full range of depressive symptoms, we used an “extreme groups” design (HBDI vs. LBDI) to maximize our chances of detecting the expected effect with relatively few participants (Preacher, 2015). All participants were compensated with one study pool credit or \$15 per hour and assented to a protocol approved by the Mass General Brigham Institutional Review Board. No formal power analysis was conducted for this study. Instead, sample sizes were selected to be similar to or larger than those used in prior work (e.g., 21 subjects per group in Dillon et al., 2014; 17–24 subjects per group in White et al., 2009).

Self-report measures

Participants provided basic demographic information and completed the BDI-II questionnaire using Research Electronic Data Capture (REDCap) software (Harris et al., 2019). Consistent with the IRB protocol, participants were not presented with BDI-II question 9, which assesses suicidality.

Encoding and retrieval tasks

Stimuli The pictures used in the oddball task were selected from a separate picture norming study, which was used to generate pairs of similar images such that every picture shown in the oddball task would have a closely matched lure at recognition; commonly used image sets, such as the IAPS system (Lang et al., 1997), are not ideal for this purpose. See the Supplemental Methods for details of the picture norming study. Once that study was complete, 240 pictures were selected for the oddball task. This final set consisted of 120 picture pairs (40 positive, 40 negative, and 40 neutral). As described in the Supplemental Methods, the positive and negative images differed on valence but not arousal, and the neutral images differed from emotional images on both valence and arousal. One image from each pair was randomly selected to serve as a target for each participant, with the other image serving as a lure. Four

hundred rectangles of the same color for each participant and the same size as the pictures, served as standards (Goldstein et al., 2002); these rectangles constituted 77% of the stimuli. Each picture type—positive, negative, neutral—comprised 7.7% of the stimuli.

Procedure The experimental tasks included encoding and retrieval sessions, spaced 24 hours apart. The encoding session consisted of instructions, nine practice trials, and the oddball task (~25 minutes). Participants were instructed to attend to the rapidly changing stimuli and to press the left arrow key in response to rectangles and the right arrow key in response to pictures. Some participants also were asked to make predictions about upcoming stimuli, but this manipulation did not affect the results; see Supplemental Methods for details. Stimuli were centrally presented (1,800 ms) and separated by a jittered interstimulus interval (250–1,000 ms). The task included 40 blocks with 10 trials per block and 1 stimulus per trial. Stimuli were presented in a semirandom order with the following constraints: (1) each block included 7 rectangles and 3 pictures (one per valence), (2) all pictures were separated by 1–7 rectangles, and (3) each block ended with a rectangle so two pictures never appeared sequentially across blocks. See Fig. 2A for a depiction of the oddball task.

Participants were not informed that their memory would be tested until they began the recognition test (Fig. 2B) during the retrieval session. The test was self-paced (~20 minutes; up to 10 seconds per response) and included 120 old pictures from encoding plus 120 closely matched lures. All pictures were presented in a unique, randomized order for each participant. Participants were instructed to identify each image as either old or new, and to rate their confidence (high, medium, low) in each response by using a scale shown below each picture.

Analyses Data were analyzed using Python 3.8 (van Rossum, 1995) and R software version 3.6.1 (R Core Team, 2019). Computational modeling was conducted using HDDM version 0.6.0 (Wiecki et al., 2013) in Jupyter Notebooks (Kluyver et al., 2016).

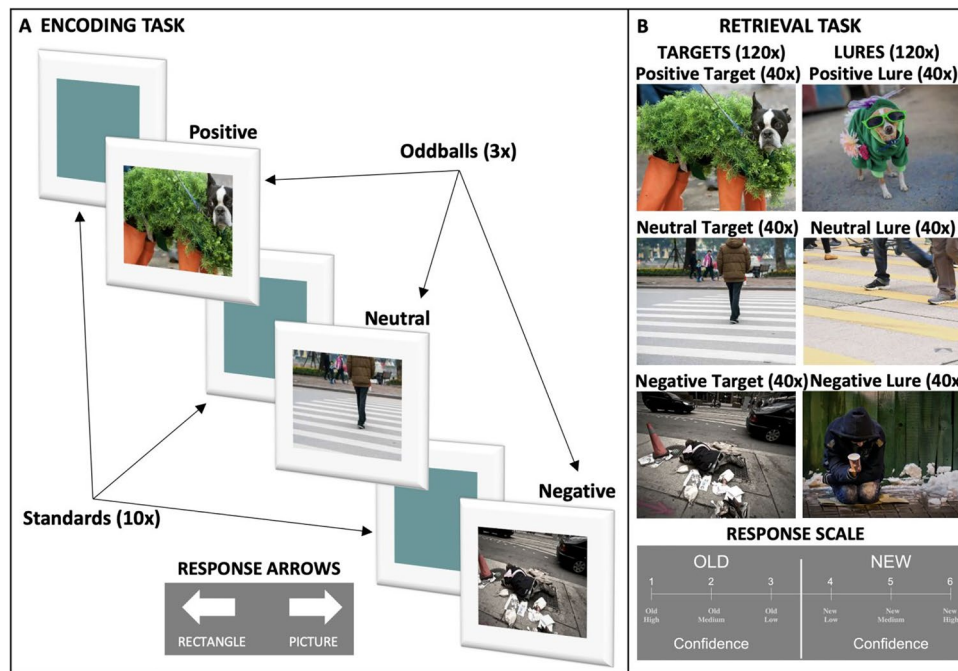


Fig. 2 Experimental tasks. *Note.* **A** Encoding trial and **(B)** graphical overview of the retrieval task. The response arrows as well as the response scales were presented below the stimuli

Continuous self-report measures (BDI-II scores, age, and years of education) were analyzed using two-group *t*-tests. Gender differences between groups were analyzed using a chi-square test. All other measures were analyzed using linear models. To account for idiosyncratic stimulus effects and individual differences in memory ability, all linear models contained *Picture Name* (unique for each image) and *Subject ID* as random intercepts (Bates et al., 2015).

For analysis of response accuracy at encoding and retrieval, trial-level accuracy (0 or 1) served as the dependent variable with *Picture Type*, *Group*, and *Memory Status* (old or new, for retrieval only), as independent variables. Covariates included *Age*, *Trial Number*, and *Years of Education*. A sensitivity analysis was conducted to test the efficacy of this model (i.e., R^2 deviation from zero) using the G*Power 3.1 module for the “Linear multiple regression: Fixed model, R^2 deviation from zero” (Faul et al., 2007). The alpha was set at 0.05 with 0.8 power, 95 sample size, and 7 predictors, to match the details of our data. Results of this sensitivity analysis indicate that the minimum effect size that can be detected with these parameters is $f^2 = 0.16$ ($F = 2.117$), which corresponds to a medium effect size (Cohen, 1988). To further characterize group differences at retrieval, we compared the estimated marginal means for recognition accuracy sorted by picture type. Finally, response times (RTs) at encoding and recognition were

(natural) log-transformed and filtered: we removed raw RTs that exceeded 10,000 ms at retrieval, as well as (log-transformed) RTs that deviated from the participant’s mean ± 3 SD. This resulted in the removal of 3.3% of encoding trials and 0.22% of retrieval trials. Filtered RTs were entered as dependent measures in models in which *Picture Type*, *Memory Status* (old vs. new; for retrieval only), *Pair Number* (for encoding only) and *Group* were independent variables. Covariates were *Age* and *Years of Education*.

HDDM The HDDM is a Bayesian implementation of the DDM that has been validated relative to other software packages (Ratcliff and Childers, 2015) and that has uncovered deficits in depressed adults in prior work (Lawlor et al., 2020). It was fit to trial-level choice and untransformed RT data at retrieval following recommendations by model developers (Wiecki et al., 2013). All parameters [v , a , z , t_0] were allowed to vary by *Group* and *Picture Type*. A total of 10,000 samples, thinned by a factor of 5, were drawn from the posterior distribution, with the first 2,000 discarded as “burn-in” samples. Convergence was confirmed via diagnostic plotting and by running the model five times to obtain the Gelman–Rubin statistic ($\hat{R} < 1.1$ for all parameters). To assess group differences, the between-group overlap of the posterior distributions for each parameter was plotted for each picture type (presented as *q-values*; Lawlor et al., 2020).

Results

Self-report

See Table 1 for demographic and self-report data. The subject pool consisted of a broad community sample of adults. By design, the HBDI and LBDI groups differed on BDI-II scores, $t(56.37) = 28.3$, $p < 0.001$, Cohen's $d = 5.66$. The groups also differed on age ($t(80.33) = -2.19$, $p = 0.032$; $d = 0.45$) and years of education ($t(3.48) = 79.41$, $p < 0.001$; $d = 0.71$), as the HBDI group was younger and had completed fewer years of education than the LBDI group. Because of these group differences, *Age* and *Years of Education* were included as covariates in all other linear models. The groups did not differ on gender, $\chi^2(1) = 0.17$, $p = 0.681$; $\Phi = 0.07$.

Table 2 Mean (\pm SD) encoding behavior in the oddball task

Picture type	Group	RTs	%Correct
Positive	HBDI	0.71 \pm 0.26	0.99 \pm 0.09
	LBDI	0.62 \pm 0.19	0.99 \pm 0.10
Neutral	HBDI	0.71 \pm 0.25	0.98 \pm 0.13
	LBDI	0.63 \pm 0.19	0.98 \pm 0.13
Negative	HBDI	0.73 \pm 0.28	0.99 \pm 0.12
	LBDI	0.63 \pm 0.20	0.99 \pm 0.10

BDI -II: Beck Depression Inventory-II (Beck et al., 1996); HBDI: High BDI-II (>20) group; LBDI: Low BDI-II (<6) group

Encoding responses

We assessed attentiveness to the oddball task by analyzing response accuracy and RT at encoding (Table 2). The linear model on response accuracy did not return any significant effects involving *Group* or *Picture Type* ($ps > 0.05$). Analysis of RTs, however, revealed a main effect of *Group*, $t(94.14) = 2.77$, $p = 0.007$, due to slower responses in the HBDI group.

Recognition memory

As displayed in Fig. 3, relative to the LBDI group, individuals with HBDI showed lower recognition accuracy for (old) positive pictures but not neutral or negative pictures; there were no apparent group differences for correct rejections of (new) lures of any type. To statistically test this visual impression, we ran a generalized linear model on recognition accuracy. To highlight the most critical test, we set responses to *Old Positive* pictures in the LBDI group as the reference level for the generalized linear model. Consequently, a significant main effect of *Group* (with a negative sign) would indicate that memory for old positive pictures was worse in the HBDI group, as predicted. As shown in Table 3, the model did indeed return a main effect of *Group* (specifically, HBDI) with a negative sign. This result confirms that recognition accuracy for old positive pictures was lower in the HBDI versus LBDI group. Although not strictly necessary, post-hoc pairwise analyses confirmed that group differences were not present for old neutral (Cohen's $d = 0.059$, $z = 0.6$, $p = 0.55$) or old negative pictures (Cohen's $d = 0.11$, $z = -1.06$, $p = 0.29$).

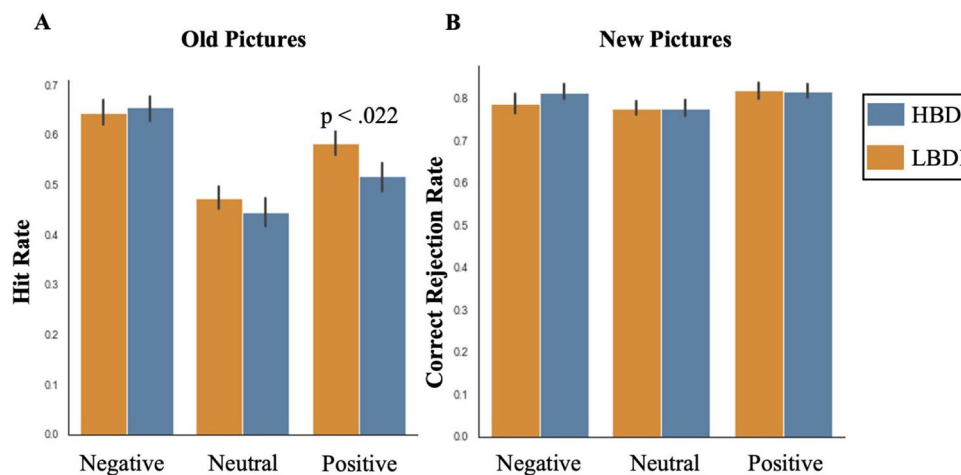


Fig. 3 Recognition accuracy. *Note.* HBDI: High BDI-II group; LBDI: Low BDI-II group. The p value corresponds to the group difference in recognition memory accuracy for old positive pictures, based on

the linear modeling results given in Table 3. The y-axes differ across panels A and B

Table 3 Generalized linear mixed model results for trial-level recognition accuracy

	Estimate	SE	Z-value	p value
Intercept	0.31	0.08	3.86	<0.001
Negative	0.27	0.09	3.06	0.002
Neutral	-0.47	0.09	-5.48	<0.001
New	1.24	0.08	15.65	<0.001
HBDI	-0.23	0.10	-2.29	0.022
Age	-0.03	0.04	-0.74	0.463
Trial number	-0.06	0.02	-3.79	<0.001
Years of education	0.07	0.04	1.77	0.077
Negative x new	-0.27	0.11	-2.37	0.018
Neutral x new	0.22	0.11	2.02	0.043
Negative x HBDI	0.33	0.10	3.42	<0.001
Neutral x HBDI	0.17	0.10	1.76	0.079
New x HBDI	0.28	0.11	2.51	0.012
Negative x New x HBDI	-0.55	0.15	-3.55	<0.001
Neutral x New x HBDI	-0.19	0.15	-1.27	0.206

BDI -II: Beck Depression Inventory-II (Beck et al., 1996); *HBDI*: High BDI-II (>20) group; *LBDI*: Low BDI-II (<6) group. The reference condition for this model was *Old Positive* pictures in the LBDI group, thus the significant *HBDI* effect (in bold) confirms poorer recognition accuracy for old positive pictures in the HBDI group

The model returned three additional interactions involving *Group*. First, there was a *Negative x HBDI* interaction. This indicates that the HBDI group showed a greater increase in memory accuracy for old negative versus old positive pictures than the LBDI group. This effect is driven primarily by the aforementioned group difference for old positive pictures, as accuracy in response to old negative pictures was similar across groups. Second, a *New x HBDI*

interaction indicates that although all participants were better at correctly rejecting new, positive lures than they were at identifying old, positive pictures, this difference was larger in the HBDI group versus the LBDI group. Note that this effect is also driven by poorer memory for old positive pictures in the HBDI group, as the groups correctly rejected new positive pictures at a similar rate. Third, a significant *Negative x New x HBDI* interaction emerged, because the ability to correctly reject negative lures was slightly better in the HBDI versus LBDI group. Precise quantitative interpretations for every term in the model, including ones not detailed here, can be made by examining the regression coefficients presented in Table 3.

HDDM

As shown in the top row of Fig. 4, there was substantial between-group overlap for drift-rates in response to old negative pictures ($q = 0.44$), little overlap for old neutral pictures ($q = 0.1$), and markedly less overlap for old positive pictures ($q = 0.01$). The last two results were driven by lower drift-rates in the HBDI group, and they indicate that memory evidence for old positive pictures (and, to a lesser extent, old neutral pictures) accrued more slowly in depressed adults. The bottom row shows that group differences were not pronounced for new pictures.

The other model parameters revealed limited evidence of group differences. As depicted in Figure S1, decision boundaries were generally wider for the HBDI versus LBDI group but this effect was weakest for old positive pictures ($q = 0.46$), indicating that it did not drive the group difference in recognition accuracy for these stimuli. Figure S2 shows that nondecision times were longer in the HBDI group, and

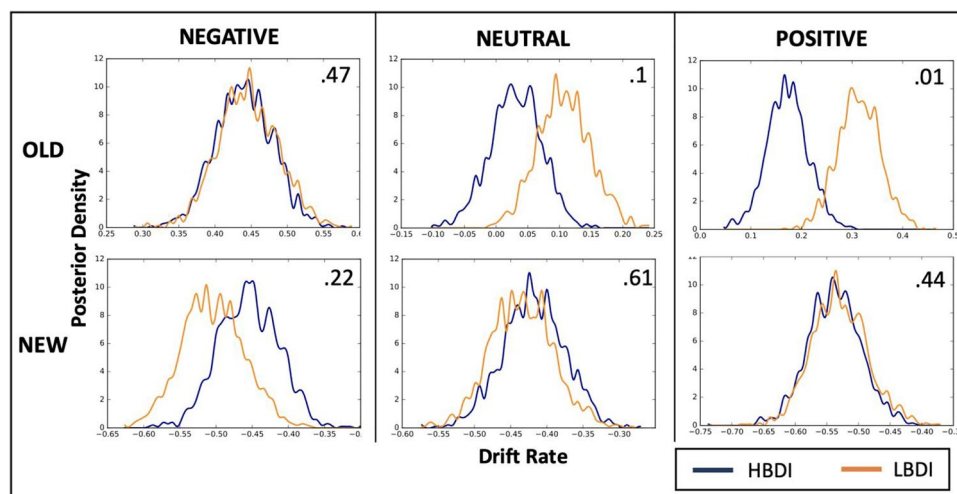


Fig. 4 Posterior distributions for drift-rate. *Note.* Between-group overlap of posterior distributions for drift-rates, by *Picture Types* (columns) and *Old/New* status (rows). *HBDI*: High BDI-II group; *LBDI*: Low BDI-II group

Figure S3 shows that there was also a weaker bias to respond “new” in the HBDI group. In both cases, the between-group overlap was substantial; thus, the group differences are modest. Although we caution that the degree of overlap of posterior probability distributions does not constitute a statistical significance test, no parameter other than the drift rate for old positive pictures showed less than 5% between-group overlap (i.e., no other q -value was below 0.05). Thus, the strongest result from the HDDM was the observation of slower drift-rates in response to old positive pictures in the HBDI group.

Discussion

This study found that, relative to adults with minimal depressive symptoms, adults with elevated symptoms had greater difficulty recognizing positive pictures encoded a day earlier. The HDDM linked this deficit to slow evidence accumulation during memory retrieval. There were no group differences in recognition accuracy for old negative or old neutral pictures, or for correct rejections of new pictures of any valence. Furthermore, analysis of the other HDDM parameters revealed minimal evidence of group differences. In short, this study shows a selective association between elevated depressive symptoms and difficulty retrieving positive memories.

We interpret the retrieval results as reflecting poor encoding of positive oddballs in the HBDI group. Although this interpretation is speculative given the lack of EEG data, we believe it is sensible for two reasons. First, the oddball task has successfully elicited prediction errors in hundreds, if not thousands, of EEG and fMRI studies over the past six decades. Consequently, it is reasonable to assume that it elicited PEs in the current study. Given that the pictures used as oddballs were selected based on their ability to elicit valenced responses in the picture norming study, it is also reasonable to assume that the positive pictures likely elicited positive PEs—PPEs—in the LBDI group. Given that PPEs elicit DA bursts (Schultz, 1998), strengthen memory formation in healthy adults (Jang et al., 2019), and that depressed adults perform poorly on reward-processing tasks (Halahakoon et al., 2020) due to diminished activity within brain-reward networks (Proudfit, 2015; Treadway & Zald, 2011), a parsimonious conclusion is that PPEs at encoding were weaker in the HBDI group and that this contributed directly to the group difference in memory for positive oddballs.

The second point in favor of this interpretation concerns the overall pattern of the recognition memory data. It is noteworthy that although the HBDI group performed more poorly than the LBDI group in response to old positive pictures, there were no group differences in response to new positive pictures. These results cannot be easily interpreted

in terms of either a general insensitivity to positive valence (Levens and Gotlib, 2009) or diminished attention to positive events (Peckham et al., 2010), because the HBDI group did not show lower accuracy in response to old *and* new positive pictures, just old ones. The fact that the deficit was selective to previously studied pictures implies that encoding of positive material was less effective in the HBDI group. We hypothesize that this reflects blunted PPEs due to abnormalities in the DA system (Dillon, 2015; Dillon and Pizzagalli, 2018).

How would such abnormalities lead to slow evidence accumulation at retrieval? The speed of evidence accumulation in recognition memory tests is thought to depend on how well the stimulus presented on a given trial matches its representation in long-term memory (Ratcliff, 1978). For correctly rejected new items, a parallel search process yields minimal evidence for a match and the drift process rapidly reaches the “new” boundary. For well-encoded old items, the match with the existing representation is good and thus the drift process efficiently reaches the “old” boundary. This match is weaker for poorly encoded items and evidence accumulation is slower and more meandering, sometimes terminating (incorrectly) at the “new” boundary and thus leading to lower recognition accuracy. If our interpretation is correct, then weaker PPEs in response to positive oddballs at encoding in the HBDI group ultimately led to poorer quality matches during the recognition test, causing slower drift-rates and lower recognition accuracy.

These findings suggest an interesting hypothesis for future EEG/ERP and fMRI studies. A common result in the control literature is that when successful recognition of old items (hits) is contrasted with correct rejection of lures, robust activation of parietal cortex emerges (Rugg and Curran, 2007; Spaniol et al., 2009). The specific mechanism underlying this activation remains controversial, but one possibility is that it reflects the accumulation of memory evidence supplied by the hippocampus and other cortical areas (Wagner et al., 2005). This idea builds on a large nonhuman primate literature indicating that lateral parietal regions support the accumulation of perceptual evidence (Gold and Shadlen, 2007; Shadlen and Newsome, 2001). The current data indicate that using EEG/ERP and fMRI to study emotional memory retrieval in depressed adults would be worthwhile, because in addition to any effects related to PEs at encoding, depression may disrupt the accumulation of positive—but not negative—memory evidence, and this disruption may be detectable by monitoring parietal activation. Indeed, an earlier study from our lab found lower amplitude, left parietal ERPs during recollection of neutral memories in adults with MDD versus healthy controls (Barrick and Dillon, 2018), but extending this approach to emotional memories and linking EEG/ERP signals to model parameters remains a goal.

Although the current findings are promising, several limitations should be noted. First, participants only completed the BDI-II, so we have no information about other conditions, such as anxiety, that are frequently comorbid with depression and that may have influenced the results. Second, the use of an extreme groups design comes at the cost of generalizability, as we do not know how adults with moderate depressive symptoms would behave. Third, the lack of a neurophysiological measurement at encoding means that while we could make reasonable inferences about PPEs at encoding, we could not observe them directly. Thus, future studies should acquire neurophysiological data from larger, well-characterized samples whose depressive symptoms span the range from mild to severe.

An important focus of future research will be to translate these findings to concrete clinical applications. Recent work in memory therapeutics (Dalgleish and Werner-Seidler, 2014) is promising. For instance, Memory Specificity Training seeks to improve memory by training participants to increase the number of specific details they recall from their pasts. Our study suggests that depressed adults may benefit from focusing specifically on retrieving detailed positive memories, and indeed there is already evidence that this approach is therapeutic (Dalgleish et al., 2013; Hitchcock et al., 2017). It would be valuable to know whether therapy-induced changes in the retrieval of positive, autobiographical memories are related to changes in drift rate.

Given the hypothesis that the positive memory deficit in depression is related to abnormalities in the DA system, it also would be useful to study the memory effects of interventions that stimulate DA production, either behaviorally (i.e., via mirthful laughter induction; Manninen et al., 2017) or pharmacologically (i.e., via bupropion administration; Stahl et al., 2004). Another clinical extension of this work might use noninvasive brain stimulation, which successfully enhanced cortical-hippocampal connectivity and improved memory performance in healthy adults (Wang et al., 2014). Given that evidence accumulation is linked to prefrontal and parietal function (Hanks et al., 2015), using noninvasive methods to stimulate these brain regions during positive memory retrieval might yield benefits for depressed adults.

Conclusions

The current study adds further support to the claim that depression can impair memory for positive material, and points to two relevant underlying mechanisms: blunted PPEs at encoding and slow evidence accumulation at retrieval. This study offers new insights into the harmful effects of depressive symptoms on memory and improves our understanding of why depressed individuals have difficulty remembering positive events. These findings also suggest

new therapeutic interventions that might help alleviate the positive memory deficit in depressed individuals.

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.3758/s13415-022-01010-2>.

Acknowledgments This work was supported by grants from the National Institute of Mental Health of the National Institutes of Health (K01 MH122672, awarded to Dr. Maksimovskiy; R01 MH111676, awarded to Dr. Dillon) and the National Institute on Drug Abuse of the National Institutes of Health (T32 DA015036, awarded to Dr. Scott Lukas). The authors thank Dr. Lukas for his help with stimulus preparation and experimental design.

Declarations

Disclaimer The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

References

- Barrick, E. M., & Dillon, D. G. (2018). An ERP study of multidimensional source retrieval in depression. *Biological Psychology, 132*, 176–191. <https://doi.org/10.1016/j.biopsycho.2018.01.001>
- Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting Linear Mixed-Effects Models Using lme4. *Journal of Statistical Software, 67*(1), 1–48. <https://doi.org/10.18637/jss.v067.i01>
- Beck, A., Steer, R. A., & Brown, G. K. (1996). *Beck Depression Inventory – Second Edition: Manual*. The Psychological Corporation: San Antonio, TX.
- Bethus, I., Tse, D., & Morris, R. G. (2010). Dopamine and memory: modulation of the persistence of memory for novel hippocampal NMDA receptor-dependent paired associates. *Journal of Neuroscience, 30*(5), 1610–1618. <https://doi.org/10.1523/JNEUROSCI.2721-09.2010>
- Burt, D. B., Zembor, M. J., & Niederehe, G. (1995). Depression and memory impairment: a meta-analysis of the association, its pattern, and specificity. *Psychological Bulletin, 117*(2), 285–305. <https://doi.org/10.1037/0033-2909.117.2.285>
- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Erlbaum.
- Criss, A. H. (2010). Differentiation and response bias in episodic memory: Evidence from reaction time distributions. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 36*(2), 484–499. <https://doi.org/10.1037/a0018435>
- Dalgleish, T., Navrady, L., Bird, E., Hill, E., Dunn, B. D., & Golden, A. M. (2013). Method-of-loci as a mnemonic device to facilitate access to self-affirming personal memories for individuals with depression. *Clinical Psychological Science, 1*(2), 156–162. <https://doi.org/10.1177/2F2167702612468111>
- Dalgleish, T., & Werner-Seidler, A. (2014). Disruptions in autobiographical memory processing in depression and the emergence of memory therapeutics. *Trends in Cognitive Sciences, 18*(11), 596–604. <https://doi.org/10.1016/j.tics.2014.06.010>
- Dalgleish, T., Williams, J. M. G., Golden, A. M. J., Perkins, N., Barrett, L. F., Barnard, P. J., et al. (2007). Reduced specificity of autobiographical memory and depression: the role of executive control. *Journal of Experimental Psychology: General, 136*(1), 23–42. <https://doi.org/10.1037/0096-3445.136.1.23>

- Dillon, D. G. (2015). The neuroscience of positive memory deficits in depression. *Frontiers in Psychology*, 6, 1295. <https://doi.org/10.3389/fpsyg.2015.01295>
- Dillon, D. G., Dobbins, I. G., & Pizzagalli, D. A. (2014). Weak reward source memory in depression reflects blunted activation of VTA/SN and parahippocampus. *Social cognitive and affective neuroscience*, 9(10), 1576–1583.
- Dillon, D. G., & Pizzagalli, D. A. (2018). Mechanisms of memory disruption in depression. *Trends in Neurosciences*, 41(3), 137–149. <https://doi.org/10.1016/j.tins.2017.12.006>
- Donchin, E., & Coles, M. G. (1988). Is the P300 component a manifestation of context updating? *Behavioral and brain sciences*, 11(3), 357–374. <https://doi.org/10.1017/S0140525X00058027>
- Faul, F., Erdfelder, E., Lang, A. G., & Buchner, A. (2007). G* Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods*, 39(2), 175–191. <https://doi.org/10.3758/bf03193146>
- Gold, J. I., & Shadlen, M. N. (2007). The neural basis of decision making. *Annual Review of Neuroscience*, 30, 535–574. <https://doi.org/10.1146/annurev.neuro.29.051605.113038>
- Goldstein, A., Spencer, K. M., & Donchin, E. (2002). The influence of stimulus deviance and novelty on the P300 and novelty P3. *Psychophysiology*, 39(6), 781–790. <https://doi.org/10.1017/S004857720201048X>
- Gotlib, I. H., & Joormann, J. (2010). Cognition and depression: current status and future directions. *Annual review of clinical psychology*, 6, 285–312.
- Hanks, T. D., Kopec, C. D., Brunton, B. W., Duan, C. A., Erlich, J. C., & Brody, C. D. (2015). Distinct relationships of parietal and prefrontal cortices to evidence accumulation. *Nature*, 520(7546), 220–223. <https://doi.org/10.1038/nature14066>
- Halahakoon, D. C., Kieslich, K., O'Driscoll, C., Nair, A., Lewis, G., & Roiser, J. P. (2020). Reward-processing behavior in depressed participants relative to healthy volunteers: A systematic review and meta-analysis. *JAMA Psychiatry*, 77(12), 1286–1295. <https://doi.org/10.1001/jamapsychiatry.2020.2139>
- Hallford, D. J., Rusanov, D., Yeow, J. J. E., & Barry, T. J. (2021). Over-general and specific autobiographical memory predict the course of depression: an updated meta-analysis. *Psychological Medicine*, 51(6), 909–926. <https://doi.org/10.1017/S0033291721001343>
- Harris, P. A., Taylor, R., Minor, B. L., Elliott, V., Fernandez, M., O'Neal, L., et al. (2019). The REDCap consortium: Building an international community of software platform partners. *Journal of Biomedical Informatics*, 95, 103208. <https://doi.org/10.1016/j.jbi.2019.103208>
- Hitchcock, C., Werner-Seidler, A., Blackwell, S. E., & Dalgleish, T. (2017). Autobiographical episodic memory-based training for the treatment of mood, anxiety and stress-related disorders: A systematic review and meta-analysis. *Clinical Psychology Review*, 52, 92–107. <https://doi.org/10.1016/j.cpr.2016.12.003>
- James, S. L., Abate, D., Abate, K. H., Abay, S. M., Abbafati, C., Abbasi, N., et al. (2018). Global, regional, and national incidence, prevalence, and years lived with disability for 354 diseases and injuries for 195 countries and territories, 1990–2017: a systematic analysis for the Global Burden of Disease Study 2017. *The Lancet*, 392(10159), 1789–1858. [https://doi.org/10.1016/S0140-6736\(18\)32279-7](https://doi.org/10.1016/S0140-6736(18)32279-7)
- Jang, A. I., Nassar, M. R., Dillon, D. G., & Frank, M. J. (2019). Positive reward prediction errors during decision-making strengthen memory encoding. *Nature Human Behaviour*, 3(7), 719–732. <https://doi.org/10.1038/s41562-019-0597-3>
- Joormann, J., & Siemer, M. (2004). Memory accessibility, mood regulation, and dysphoria: Difficulties in repairing sad mood with happy memories? *Journal of Abnormal Psychology*, 113(2), 179–188. <https://doi.org/10.1037/0021-843X.113.2.179>
- Kluyver, T., Ragan-Kelley, B., Pérez, F., Granger, B. E., Bussonnier, M., Frederic, J., et al. (2016). *Jupyter Notebooks—a publishing format for reproducible computational workflows*, 2016, 87–90. <https://doi.org/10.3233/978-1-61499-649-1-87>
- Haber, S. N., & Knutson, B. (2010). The reward circuit: linking primate anatomy and human imaging. *Neuropsychopharmacology*, 35(1), 4–26. <https://doi.org/10.1038/npp.2009.129>
- Lang, P. J., Bradley, M. M., & Cuthbert, B. N. (1997). International affective picture system (IAPS): Technical manual and affective ratings. *NIMH Center for the Study of Emotion and Attention*, 1(39–58), 3.
- Lawlor, V. M., Webb, C. A., Wiecki, T. V., Frank, M. J., Trivedi, M., Pizzagalli, D. A., & Dillon, D. G. (2020). Dissecting the impact of depression on decision-making. *Psychological Medicine*, 50(10), 1613–1622. <https://doi.org/10.1017/S0033291719001570>
- Levens, S. M., & Gotlib, I. H. (2009). Impaired selection of relevant positive information in depression. *Depression and Anxiety*, 26(5), 403–410. <https://doi.org/10.1002/da.20565>
- Lisman, J. E., & Grace, A. A. (2005). The hippocampal-VTA loop: controlling the entry of information into long-term memory. *Neuron*, 46(5), 703–713. <https://doi.org/10.1016/j.neuron.2005.05.002>
- MacQueen, G. M., Galway, T. M., Hay, J., Young, L. T., & Joffe, R. T. (2002). Recollection memory deficits in patients with major depressive disorder predicted by past depressions but not current mood state or treatment status. *Psychological Medicine*, 32(2), 251–258. <https://doi.org/10.1017/S0033291701004834>
- Matt, G. E., Vázquez, C., & Campbell, W. K. (1992). Mood-congruent recall of affectively toned stimuli: A meta-analytic review. *Clinical Psychology Review*, 12(2), 227–255. [https://doi.org/10.1016/0272-7358\(92\)90116-P](https://doi.org/10.1016/0272-7358(92)90116-P)
- Manninen, S., Tuominen, L., Dunbar, R. I., Karjalainen, T., Hirvonen, J., Arponen, E., et al. (2017). Social laughter triggers endogenous opioid release in humans. *Journal of Neuroscience*, 37(25), 6125–6131. <https://doi.org/10.1523/JNEUROSCI.0688-16.2017>
- Peckham, A. D., McHugh, R. K., & Otto, M. W. (2010). A meta-analysis of the magnitude of biased attention in depression. *Depression and Anxiety*, 27(12), 1135–1142. <https://doi.org/10.1002/da.20755>
- Pizzagalli, D. A. (2014). Depression, stress, and anhedonia: toward a synthesis and integrated model. *Annual Review of Clinical Psychology*, 10(1), 393–423. <https://doi.org/10.1146/annurev-clinpsy-050212-185606>
- Polich, J. (2007). Updating P300: an integrative theory of P3a and P3b. *Clinical Neurophysiology*, 118(10), 2128–2148. <https://doi.org/10.1016/j.clinph.2007.04.019>
- Preacher, K. J. (2015). Extreme groups designs. *The Encyclopedia of Clinical Psychology*, 2, 1189–1192.
- Proudfit, G. H. (2015). The reward positivity: From basic research on reward to a biomarker for depression. *Psychophysiology*, 52(4), 449–459. <https://doi.org/10.1111/psyp.12370>
- R Core Team (2019). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria. URL <https://www.R-project.org/>
- Ramponi, C., Barnard, P., & Nimmo-Smith, I. (2004). Recollection deficits in dysphoric mood: An effect of schematic models and executive mode? *Memory*, 12(5), 655–670. <https://doi.org/10.1080/09658210344000189>
- Ratcliff, R. (1978). A theory of memory retrieval. *Psychological Review*, 85(2), 59–108. <https://doi.org/10.1037/0033-295X.85.2.59>
- Ratcliff, R., & Childers, R. (2015). Individual differences and fitting methods for the two-choice diffusion model of decision making. *Decision*, 2(4), 237. <https://doi.org/10.1037/dec000030>
- Ratcliff, R., & McKoon, G. (2008). The diffusion decision model: theory and data for two-choice decision tasks. *Neural Computation*, 20(4), 873–922. <https://doi.org/10.1162/neco.2008.12-06-420>

- Rock, P. L., Roiser, J. P., Riedel, W. J., & Blackwell, A. D. (2014). Cognitive impairment in depression: a systematic review and meta-analysis. *Psychological Medicine*, *44*(10), 2029–2040. <https://doi.org/10.1017/S0033291713002535>
- Rugg, M. D., & Curran, T. (2007). Event-related potentials and recognition memory. *Trends in Cognitive Sciences*, *11*(6), 251–257. <https://doi.org/10.1016/j.tics.2007.04.004>
- Schultz, W. (1998). Predictive reward signal of dopamine neurons. *Journal of neurophysiology*, *80*(1), 1–27. <https://doi.org/10.1152/jn.1998.80.1.1>
- Schultz, W. (2016). Dopamine reward prediction-error signaling: a two-component response. *Nature Reviews Neuroscience*, *17*(3), 183–195. <https://doi.org/10.1038/nrn.2015.26>
- Shadlen, M. N., & Newsome, W. T. (2001). Neural basis of a perceptual decision in the parietal cortex (area LIP) of the rhesus monkey. *Journal of Neurophysiology*, *86*(4), 1916–1936. <https://doi.org/10.1152/jn.2001.86.4.1916>
- Spaniol, J., Davidson, P. S., Kim, A. S., Han, H., Moscovitch, M., & Grady, C. L. (2009). Event-related fMRI studies of episodic encoding and retrieval: meta-analyses using activation likelihood estimation. *Neuropsychologia*, *47*(8-9), 1765–1779. <https://doi.org/10.1016/j.neuropsychologia.2009.02.028>
- Stahl, S. M., Pradko, J. F., Haight, B. R., Modell, J. G., Rockett, C. B., & Learned-Coughlin, S. (2004). A review of the neuropharmacology of bupropion, a dual norepinephrine and dopamine reuptake inhibitor. *Primary care companion to the Journal of clinical psychiatry*, *6*(4), 159. <https://doi.org/10.4088/pcc.v06n0403>
- Treadway, M. T., & Zald, D. H. (2011). Reconsidering anhedonia in depression: Lessons from translational neuroscience. *Neuroscience and Biobehavioral Reviews*, *35*(3), 537–555. <https://doi.org/10.1016/j.neubiorev.2010.06.006>
- van Rossum, G., 1995. *Python tutorial* (Report No. CS-R9526). Computer Science/Department of Algorithmics and Architecture, National Research Institute for Mathematics and Computer Science. <https://ir.cwi.nl/pub/5007/05007D.pdf>
- Wagner, A. D., Shannon, B. J., Kahn, I., & Buckner, R. L. (2005). Parietal lobe contributions to episodic memory retrieval. *Trends in Cognitive Sciences*, *9*(9), 445–453. <https://doi.org/10.1016/j.tics.2005.07.001>
- Wang, J. X., Rogers, L. M., Gross, E. Z., Ryals, A. J., Dokucu, M. E., Brandstatt, K. L., et al. (2014). Targeted enhancement of cortical-hippocampal brain networks and associative memory. *Science*, *345*(6200), 1054–1057. <https://doi.org/10.1126/science.1252900>
- White, C., Ratcliff, R., Vasey, M., & McKoon, G. (2009). Dysphoria and memory for emotional material: A diffusion-model analysis. *Cognition and Emotion*, *23*(1), 181–205. <https://doi.org/10.1080/02699930801976770>
- White, C. N., Ratcliff, R., Vasey, M. W., & McKoon, G. (2010). Using diffusion models to understand clinical disorders. *Journal of Mathematical Psychology*, *54*(1), 39–52. <https://doi.org/10.1016/j.jmp.2010.01.004>
- Wiecki, T. V., Sofer, I., & Frank, M. J. (2013). HDDM: Hierarchical Bayesian estimation of the Drift-Diffusion Model in Python. *Frontiers in Neuroinformatics*, *7*, 14–14. <https://doi.org/10.3389/fninf.2013.00014>
- Zakzanis, K. K., Leach, L., & Kaplan, E. (1998). On the nature and pattern of neurocognitive function in major depressive disorder. *Neuropsychiatry, Neuropsychology, & Behavioral Neurology*, *11*(3), 111–119. https://hollis.harvard.edu/permalink/f/1mdq5o5/TN_cdi_proquest_miscellaneous_73926383

Open Practice Statement

The data and materials from the experiments reported here is available upon request and upon approval from the Institutional Review Board (IRB). These data were not preregistered.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.