Consciousness and Cognition 63 (2018) 206–217

Contents lists available at ScienceDirect

Consciousness and Cognition

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

Tip-of-the-tongue states predict enhanced feedback processing and subsequent memory

Paul A. Bloom[*], David Friedman, Judy Xu, Matti Vuorre, Janet Metcalfe

Columbia University, Department of Psychology, Schermerhorn Hall #406, 1190 Amsterdam Ave, New York, NY, USA

ARTICLE INFO

Keywords:
Tip-of-the-tongue
ERP
EEG
Learning
Memory

ABSTRACT

This article investigates the relations among the tip-of-the-tongue (TOT) state, event related potentials (ERPs) to correct feedback to questions, and subsequent memory. ERPs were used to investigate neurocognitive responses to feedback to general information questions for which participants had expressed either being or not being in a TOT state. For questions in which participants were unable to answer within 3 s, they indicated whether they were experiencing a TOT state and then were immediately provided with the correct answer. Feedback during a TOT state, as opposed to not knowing the answer, was associated with enhanced positivity over centro-parietal electrodes 250–700 ms post-feedback, and this enhanced positivity mediated a positive relationship between TOTs and later recall. Although effects of increased semantic access during TOT states cannot be ruled out, these results suggest that information received during TOT states elicits enhanced processing—suggestive of curiosity—leading to enhanced learning of studied material.

1. Introduction

The tip-of-the-tongue (TOT) state is often a highly frustrating phenomenon. Described as “a gap that is intensely active” (James, 1890, p. 251) or as “personal introspections of inner turmoil when grappling for an elusive word," (Brown, 1991, p. 205) TOT states are defined as the simultaneous feeling of being unable to recall an item, and confidence that one could recall, given enough time (Schwartz, 2006). According to metacognitive models, the TOT state is analogous to a “low tire pressure” signal in a car in that it alerts one about a problem (failed recall); yet it is not the problem itself (Schwartz & Metcalfe, 2011). While a low tire pressure signal serves to remind car owners to inflate their tires, the function of the TOT state is more opaque. As TOT states have been associated with curiosity for un-retrieved information (Metcalfe, Schwartz, & Bloom, 2017), it may be the case that TOTs function as a signal to amplify processing when feedback, in the form of the target of the failed recall attempt, is provided. In the present research, we investigated whether TOT states function to aide learning, and whether such learning is mediated by enhanced processing of external feedback.

Apart from curiosity specifically, TOTs have been identified in several contexts as motivating cognitive processes. One possible function for TOTs is to drive retrieval attempts, and indeed several studies have found that individuals devote more resources towards retrieving items while they are experiencing TOT states (Kikyo, Ohki, & Sekihara, 2001; Schwartz, 2001). For example, individuals spent more time attempting to recall answers while in TOT states than when they were sure they did not know the answers (Schwartz, 2001). When participants were given a secondary working memory task in addition to TOT-eliciting cues, performance was poorer on the secondary task for trials in which TOT states were experienced, suggesting greater devotion of cognitive resources to memory search processes (Ryan, Petty, & Wenzlaff, 1982).

---

[*] Corresponding author.
E-mail address: pab2163@columbia.edu (P.A. Bloom).

https://doi.org/10.1016/j.concog.2018.05.010
Received 24 January 2018; Received in revised form 30 April 2018; Accepted 23 May 2018
Available online 07 June 2018
1053-8100/ © 2018 Elsevier Inc. All rights reserved.

In addition to motivating retrieval efforts, TOT states have also been suggested to drive curiosity as operationalized by answer-seeking (i.e., either looking up the answer or asking for the answer). Litman, Hutchins, and Russon (2005) gave participants a series of 12 general information questions, and for each asked participants whether they knew the answer, they didn't know the answer, or they were in a TOT state. When participants were later allowed to open envelopes containing answers to these same questions, the researchers reported that participants were more likely to choose questions about which they had earlier been in a TOT state than those about which they were not in a TOT state. These findings were limited by the small number of questions used per participant (and the small number of TOT states) and by the long time period between the experience of being in the TOT state and the choices. In a subsequent study (Metcalfe et al., 2017) using many more questions (82) with very little time to attempt retrieval (resulting in all participants experiencing TOT states), and also requiring participants to choose immediately – while they were still in a TOT state – participants were approximately twice as likely to choose to see the TOT state-related answers than the non-TOT state answers, indicating that TOTs are associated with elevated curiosity for target items.

Indeed, TOTs seem to fit within Berlyne's (1954) theory of epistemic curiosity, which posits that the items that are almost known are the ones which most spark human interest. Epistemic curiosity's function is thought to be a motivator to close the gap between what one knows and what one desires to know (Berlyne, 1954, 1960; Litman & Spielberger, 2003; Loewenstein, 1994), with items that are perceived to be almost known being those most likely to provoke such curiosity. Indeed, when individuals were presented with learning tasks, they chose to spend more time studying items that were within a "region of proximal learning," (RPL) which is defined as consisting of those as-yet-unlearned items with the highest judgments of learning (Metcalfe & Kornell, 2005; Metcalfe, 2002). In addition, individuals mind-wandered less when studying items they judged to be within their RPL, as compared to items that were perceived to be either too easy or too difficult (Xu & Metcalfe, 2016). It may be possible to conceptualize the TOT state as a marker that the unrecalled information resides within an RPL-like domain: between information that is already well known and that which is completely unfamiliar. It thereby elicits a high degree of curiosity.

Curiosity has also been linked to enhanced subsequent memory. In an fMRI study, Kang et al. (2009) administered general information questions to participants and probed for their curiosity on each question. When the experimenters provided participants with correct answer feedback, answers to curiosity-evoking questions were associated with activation in the left inferior frontal gyrus (IFG) and left parahippocampal gyrus (PHG), regions believed to be active in verbal memory encoding (Paller & Wagner, 2002; Wagner et al., 1998). In a follow-up behavioral experiment, feedback following incorrect answers about which the participant was curious was better recalled than feedback about errors about which the participant had not been curious. While the researchers did probe participants as to whether they experienced TOTs, participants did not report enough TOT states for analysis under this paradigm.

If the TOT state is, itself, a marker of curiosity, then memory may also be better for feedback provided when the learner is in a TOT state. Supporting this idea, Gardiner, Craik, and Bleasdale (1973) found improved subsequent memory for items for which participants had been in a TOT, compared to those items for which participants were not in a TOT. They suggested that TOT states enhance processing, either through increasing the amount of attention devoted to feedback or by activating more semantic attributes of the target word. MacKay and Burke (1990) have suggested that TOTs occur when considerable semantic information about a target is available, and yet the phonological information is insufficient to allow production of the name (also see, Burke, MacKay, Worthley, & Wade, 1991). It is well documented that individuals have partial information about the target when they are in a TOT state: they can often recall the first letter (Brown & McNeill, 1966; Koriat & Lieblich, 1974), the number of syllables (Burke et al., 1991), and even the grammatical gender of the word (Miozzo & Caramazza, 1997). Furthermore, given sufficient time they can sometimes recall the target (Cohen & Faulkner, 1986). Thus, as noted by Gardiner et al. (1973), it is likely that this partial information, as well as TOT-related increases in attention, contributes to the high rates of recall for TOT-evoking items. We further explore this relation between TOT states and subsequent memory, and the possibility that people's event-related potentials (ERPs), while receiving feedback when they are either in a TOT state or not, may shed light on this relation.

There is a substantial body of work showing that distinct neural patterns during learning can be found based on whether items are successfully recalled in a subsequent test. In particular, ERPs for later remembered items tend to display heightened late positivity during learning compared to those that are forgotten (Fabiani, Karis, & Donchin, 1990; Friedman & Johnson, 2000; Friedman & Trott, 2000; Friedman, 2003; Friedman, Ritter, & Snodgrass, 1996; Paller, Kutas, & Mayes, 1987; Paller, McCarthy, & Wood, 1988; Sanquist, Rohrbaugh, Syndulko, & Lindsley, 1980). This enhanced processing for subsequently remembered items, named the difference in memory effect (Dm; Paller et al., 1987), has been found with a variety of learning tasks, and with both words (Fabiani & Donchin, 1995; Fabiani, Karis, & Donchin, 1986) and pictures (Friedman & Sutton, 1987) as stimuli.

While Dm effects show a consistent pattern such that subsequently remembered items are associated with more positive ERP amplitudes, the latency and scalp distributions of such effects vary with the task and stimuli such that the effect has been recorded over frontal (Friedman, 1990; Münte, Heinze, Scholz, & Künkel, 1988; Paller & Kutas, 1992; Sanquist et al., 1980), central (Paller, 1990), and parietal electrodes (Paller & Wagner, 2002; Paller et al., 1987; Weyerts, Tendolkar, Smid, & Heinze, 1997). In addition, while some Dm effects resemble the P300 component elicited by novel items (Wagner, Koutstaal, & Schacter, 1999), the heightened positivity for subsequently recalled items in many cases persists much longer, up to 900–1000 ms post-stimulus, than the typical P300 (Friedman & Sutton, 1987; Smith, 1993). Furthermore, as demonstrated by Friedman (1990), the subsequent memory effect may be dissociated from the P300 with which it overlaps by differences in scalp distribution between the two components. The heightened ERP amplitudes for later remembered items are thought to reflect enhanced encoding, as items subject to deep semantic processing display larger Dm effects than items processed in a shallow non-semantic fashion (Paller et al, 1987). Further, Wagner et al. (1999) suggested that in some instances, subsequent memory effects depend on the existence of prior knowledge that individuals can incorporate into their representations of the current stimuli. Because TOT states are thought to involve a high degree of prior knowledge (Schwartz & Metcalfe, 2011) or semantic access (Burke et al., 1991), it may be the case that the differential processing of feedback to TOT items, as compared to no-TOT (and unknown) items, could resemble Dm effects previously found in the literature.

Several neurocognitive studies have found evidence for heightened monitoring, executive control, or continued memory search during TOTs, as compared to instances when a target was either recalled or unknown (for review, see; Díaz, Lindín, Galdo-Álvarez, & Buján, 2014; Schwartz & Díaz, 2014). In particular, fMRI studies indicate heightened TOT-related activation in areas associated with monitoring and executive control such as the dorsolateral prefrontal cortex and anterior cingulate cortex (Maril, Simons, Weaver, & Schacter, 2005; Maril, Wagner, & Schacter, 2001). Additionally, an ERP study by Díaz, Lindín, Galdo-Alvarez, Facal, and Juncos-Rabadán (2007), showed that when people were shown pictures of faces and then tasked with coming up with the names, from 0 to 450 ms the ERP signal was not different between recalled and TOT items. However, there was a difference later, such that the ERP amplitude of the recalled items (and, indeed, non-recalled items) was higher than that of the TOT items – a result that is probably related to differences in response resolution (Burke et al, 1991; Schwartz & Metcalfe, 2011). While these studies allow us to better understand the neural processes during the TOT state itself, to our knowledge no neurocognitive studies have been conducted examining processing of feedback while in the TOT state.

In the current study, we explored, using ERPs, whether the TOT state is associated with heightened processing of feedback to general information questions. We examined whether subsequent memory is enhanced for items that have previously been associated with TOT states. We also asked whether there were distinct ERP signatures for processing of feedback in TOT states. In our experiment, we asked participants a series of questions, probed them as to whether they were in a TOT state, then provided the correct feedback for each question. Afterwards, participants took a memory test on the same questions as before. ERPs time-locked to the correct feedback given immediately after the TOT probe were compared as a function of whether participants reported being in a TOT state or not during that question trial. We hypothesized that memory for the feedback to questions in which the participants had reported TOT states would be better than that for feedback to questions in which the participants had said they were not in a TOT state. Further, we predicted that feedback to TOT items would evoke distinct neural signatures indicative of enhanced processing, similar to Dm effects previously observed in the literature.

2. Method

2.1. Participants

The 30 participants recruited (16 male, 13 female, one who did not report gender) were Columbia University students, ranging in age from 18 to 32 (M = 21.27 years; SD = 4.26). One participant did not report experiencing any TOT states during the TOT/feedback phase and was discarded from analyses prior to EEG preprocessing. According to a criterion of > 5 useable EEG trials each following “TOT” and “no-TOT” responses during the TOT/feedback phase, 3 participants were excluded during EEG preprocessing, leaving 26 total participants for all analyses. Participants were recruited with community flyers and emails from lists of previous study participants. All participants reported themselves to be right-handed, fluent English speakers, in good physical health, and were free from medications known to affect the central nervous system. Participants were paid $15/hr for their time. The Columbia University and New York State Psychiatric Institute institutional review boards approved the project, and all participants provided written, informed consent.

2.2. Materials

The stimuli were 150 general information questions taken from the Nelson and Narens (1980) norms, and updated and corrected (see, Metcalfe, Casal-Roscum, Radin, & Friedman, 2015). For example, one question was “What is the capital of Jamaica?” (Answer: Kingston). The questions were displayed with PsychoPy2 in a large white font with a gray background on a computer monitor (Peirce, 2007).

2.3. Design

The experiment was a within-subjects design, where all participants were presented with a set of general information questions for an initial question presentation/TOT judgment/feedback phase (subsequently referred to as the TOT/feedback phase), then retested on the same questions in a surprise retest phase (test phase). Order of presentation of questions was randomized for each participant for both the TOT/feedback phase and the test phase. EEG was recorded during the TOT/feedback phase only. In conducting our analyses, we treated whether participants experienced TOT states in the TOT/feedback phase as if it were an independent variable with two levels (TOT vs. no-TOT, coded as 1 and 0, respectively).

2.4. Procedure

After the EEG caps were applied, participants were instructed that they would be given a series of general information questions to try to answer verbally and, if they could not answer within the allotted time, to report whether they were in a TOT state or not. Participants were also told that they would get to view the answers to all the questions for which they could not generate an answer in the allotted time. Although participants nearly always knew what a TOT state was (also see Schwartz & Metcalfe, 2011), for each participant, the experimenter explained the TOT state as a phenomenon in which “you feel sure you know the answer and think you can get it—it is imminent—but you cannot think of it at the moment.” Participants completed a practice trial before the experiment began, and were not informed that there would be a test phase at the end of the experiment.

For each trial in the TOT/feedback phase, the general information question first appeared onscreen and was read aloud by the experimenter, who immediately upon finishing hit a key to start the 3-s response period. During this period, if the participant gave any verbal answer, whether correct or not, the experimenter hit a key to move on to the next question immediately without providing feedback. At the end of the 3-s period if no response had been given, participants indicated whether they were in a TOT state or not. Then, a fixation cross appeared on the screen for 500 ms, followed by a 1000 ms duration feedback (correct answer) to the question for that particular trial. A jittered 750–1250 ms blank screen was shown before each subsequent trial.

Following the TOT/feedback phase, participants were given a 2-minute break in which the EEG cap was removed, then given instructions on how to complete the test phase. During the test phase, participants saw the same 150 general information questions for a second time (in a different random order), and were given 15 s per question to type a response. Participants were instructed to type their best guess to each question even if they were not sure of their answers. Feedback was not provided during the test phase, and participants' responses were scored using an approximate pattern matching algorithm (AGREP, generalized Levenshtein edit distance = .1, Ripley & Hornik, 2010), then checked by an experimenter blind to the TOT/feedback phase data. At the end of the experiment, after filling out a demographics sheet, participants were thanked, debriefed, paid for participation, and given the opportunity to see the answers to all of the general information questions if they wished.

2.5. Electroencephalographic (EEG) recording and preprocessing

Brain electrical activity was recorded during the TOT/feedback phase from 62 scalp sites (sintered Ag/AgCl) mounted in an Electrocap (Neuromedical Supplies) and digitized at 500 Hz (DC; high-frequency cut-off of 100-Hz; right-forehead ground). Electrodes were placed on the outer canthus of each eye to record horizontal eye movements, and directly above and below the left eye for vertical movements. Activity was originally referenced to the nose and re-referenced offline to the average of the left and right mastoids. Impedances were maintained below 10kΩ throughout the experiment. Prior to analyses, all recordings were filtered using a 0.1–25 Hz IIR-Butterworth bandpass filter to remove DC drift and muscle movements. Electrodes (between 1 and 4 per subject) were interpolated for 13 of our participants due to abnormal patterns or excessive fluctuation. Interpolation within subjects was done across all trials.

Offline artifact rejection and independent component analysis (ICA, Delorme & Makeig, 2004; Makeig & Onton, 2011) were used to remove eye blinks, eye movements, muscle activity, and other noise from the data. A subsequent round of artifact rejection was also conducted in which single trials containing voltage deviations of over 50 μV from baseline were removed from analysis. As no feedback was presented for trials in which participants verbally recalled the answers during the 3 s response period after question presentation, these trials were not included in the EEG analyses. Our preprocessing pipeline yielded a total of 2022 (648 TOT, 1374 no-TOT) trials for EEG analysis across all 26 participants included (by participant: $M_{\text{TOT Trials}} = 24.9$, SD = 12.6; $M_{\text{NO-TOT Trials}} = 52.8$, SD = 18.6).

Numeric integration was used to calculate the average area under the curve over the measurement time window of interest (i.e., negative amplitudes were subtracted from positive amplitudes and integrated over the measurement interval) using EEGLAB and ERPLAB (Delorme & Makeig, 2004; Lopez-Calderon & Luck, 2014). All ERP analyses were conducted on mean amplitudes equivalent to dividing the integral of the waveform by the length of our time window. A 200 ms pre-stimulus baseline was used. ERPs were time locked to presentation of the correct answer during the TOT/feedback phase of the experiment. Electrodes for our analyses were determined based on scalp topographies, and the most active electrodes from 250 to 700 ms post-onset across all trials were the focus of our analyses (Fig. 2). Analyses focused primarily on this time window, as we were interested in semantic processing, rather than sensory components occurring before 250 ms (Luck, 2014). Based on these criteria, we selected 9 centro-parietal electrodes for analysis: Cz, C1, C2, CPz, CP1, CP2, Pz, P1, & P2[1].

3. Results

For the analyses computed, the criterion for statistical significance was set at $p < .05$. For all ANOVAs, the Greenhouse-Geisser procedure (Jennings & Wood, 1976) was used to correct the degrees of freedom in order to account for violations of the homogeneity of variance assumption.

3.1. Behavioral results

3.1.1. TOT states

Of the 26 participants included for analyses, the average proportion of questions eliciting TOT states was .22 (SD = .08). The average proportion of questions eliciting a 'don't know' response was .46 (SD = .13). All trials during the TOT/feedback phase in which the participant gave an answer (average proportion = .32, SD = .15) were excluded from both behavioral and ERP analyses, as no feedback was presented for these trials.

---

[1] Because the most active centro-parietal region was slightly right-lateralized, we also completed the same analyses with 12 electrodes corresponding to the lateralization (Cz, C1, C2, C4, CPz, CP1, CP2, CP4, Pz, P1, P2, & P4) and our findings remained the same.

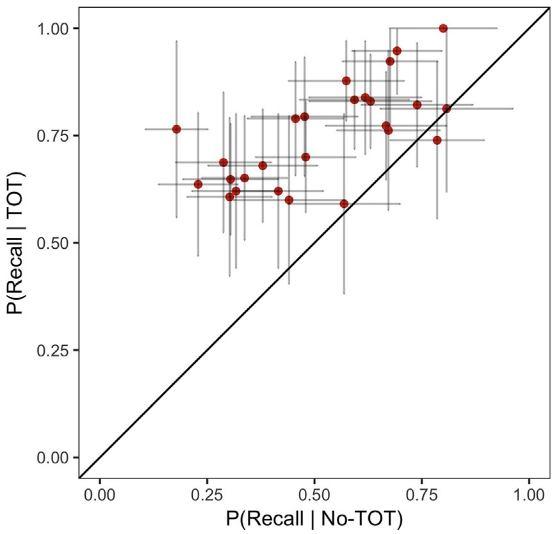

Fig. 1. Within-subjects scatter plot. Dots are individual subject recall probabilities given an initial No-TOT trial (x-axis) vs. an initial TOT trial (y-axis). Error bars are ± 2 standard errors.

3.1.2. Post-test recall

Participants' typed responses were scored for accuracy. Recall probability in the test phase for all items, including questions where participants initially responded verbally during the TOT/feedback phase and saw no feedback, was .58 (SD = .15). Participants' recall was better for trials in which they had been in the TOT state (M = .75, SD = .11) than those in which they had simply not known the answer (M = .52, SD = .19; $t(25) = 8.64$, $p < .0001$, see Fig. 1).

3.2. Electrophysiology

3.2.1. TOT versus no-TOT analysis

Tip-of-the-tongue state was associated with increased ERP amplitudes, at centro-parietal electrodes, 250–700 ms post-feedback-onset. We conducted a 2 × 9 repeated measures ANOVA of ERP amplitude using TOT state (TOT vs. no-TOT) and Electrode (Cz, C1, C2, CPz, CP1, CP2, Pz, P1, & P2) as within-subjects factors. Degrees of freedom were adjusted using the Greenhouse-Geisser procedure. There was a main effect of TOT state, such that participants showed greater amplitude to feedback for TOT trials than for no-TOT trials ($F(1, 25) = 21.77$, $p < .0001$; Fig. 3). Overall ERP amplitudes did not differ significantly across the 9 electrodes chosen for analysis ($F(2.04, 51.1) = 2.12$, $p = .130$), and electrode location did not interact significantly with TOT state ($F(3.17, 79.24) = 1.37$, $p = .257$).

3.2.2. Dm analysis

ERP amplitudes to feedback differed as a function of whether an item was subsequently correctly recalled or unrecalled. Fig. 4 displays the grand average waveforms at cento-parietal electrode sites during feedback in the TOT/feedback phase as a function of subsequent recall in the test phase. An ANOVA (Greenhouse-Geisser adjusted) revealed a main effect[2] for subsequently remembered trials on averaged ERP amplitudes in our time window of 250–700 ms, such that amplitudes were greater for TOT/feedback phase trials in which the correct feedback was subsequently recalled in the test phase ($F(1,25) = 26.31$, $p < .0001$). In this model ERP amplitudes did not differ significantly across the 9 electrodes chosen for analysis ($F(1.57, 39.34) = 2.25$, $p = .129$), and electrode location did not significantly interact with subsequent recall ($F(2.72, 67.99) = 1.41$, $p = .248$).

3.2.3. TOT versus no-TOT conditional upon recall

It is well documented in the literature (and was shown in our data in Fig. 4) that items that are subsequently recalled reveal a higher amplitude late positivity than items that are subsequently forgotten. Furthermore, TOT items are more likely to be recalled than non-TOT items. In an attempt to examine whether it was only this difference in ease of recall, and the increased information associated with the target item at time of encoding, that was responsible for the difference between the TOT and non-TOT ERP patterns, or whether the TOT state itself might also have contributed, we conducted an additional analysis in which we used only the items that were subsequently recalled. These items were divided into those that were associated with TOTs and those that were not. Even when all of the items contributing to the analysis were recalled, there was still a main effect of TOT state, as is shown in Fig. 5: ERP amplitudes in the 250–700 ms window were greater for feedback to TOT trials than for no-TOT trials ($F(1, 25) = 4.84$ $p = .037$).

---

[2] In order to weight individual subjects by their respective number of included trials and create a more easily interpretable mapping between TOT states, ERPs, and learning, we conducted further analysis (see *Single-Trial Analysis* & Fig. 6) using a model in which single-trial ERP amplitude was used as a predictor for subsequent recall.

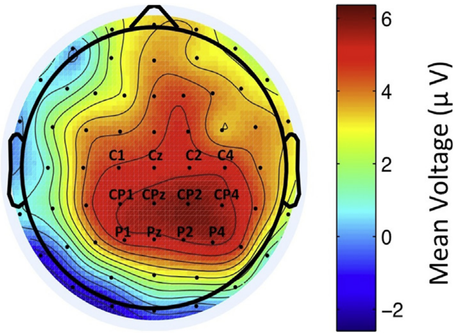

Fig. 2. Scalp topography displaying grand average mean voltage (μV) from 250 to 700 ms post-feedback onset, calculated for all trials across all participants.

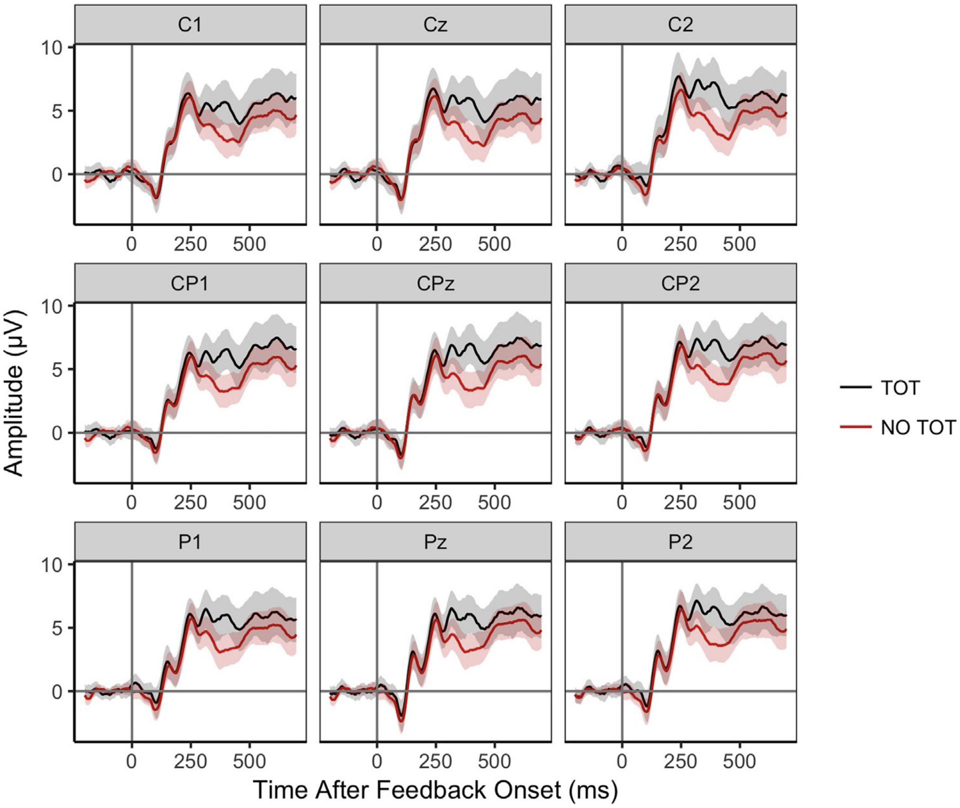

Fig. 3. ERPs time-locked to feedback onset during the initial TOT/feedback phase. ERPs to feedback showed enhanced late positivity in centro-parietal electrodes in TOT trials (red) compared to no-TOT trials (black). Shading = 95% confidence interval at each time point. N = 26. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

Overall ERP amplitudes did not differ significantly across the 9 electrodes ($F(1.69, 42.154) = 2.34$, $p = .116$), and electrode location did not interact significantly with TOT state ($F(3.31, 82.69) = 1.19$, $p = .565$).

3.2.4. Single-trial analysis

ERP amplitudes in response to feedback predict subsequent recall. To further examine the relationship between processing of TOT feedback and subsequent recall, we isolated the single-trial ERP data for each participant using custom MATLAB and R scripts. For each trial, the ERP amplitude was averaged between 250 and 700 ms post-stimulus across the 9 electrodes of interest. We then regressed[3] the correct/incorrect post-test answers on feedback ERP amplitude, whether the participant was in a TOT, and their

---

[3] We estimated a multilevel logistic regression with the R package brms (Bürkner, 2017; R Core Team, 2017) built on Stan (Gelman et al., 2015). We allowed the effects of TOT and amplitude to vary across participants, and used software-provided uninformative priors on all parameters.

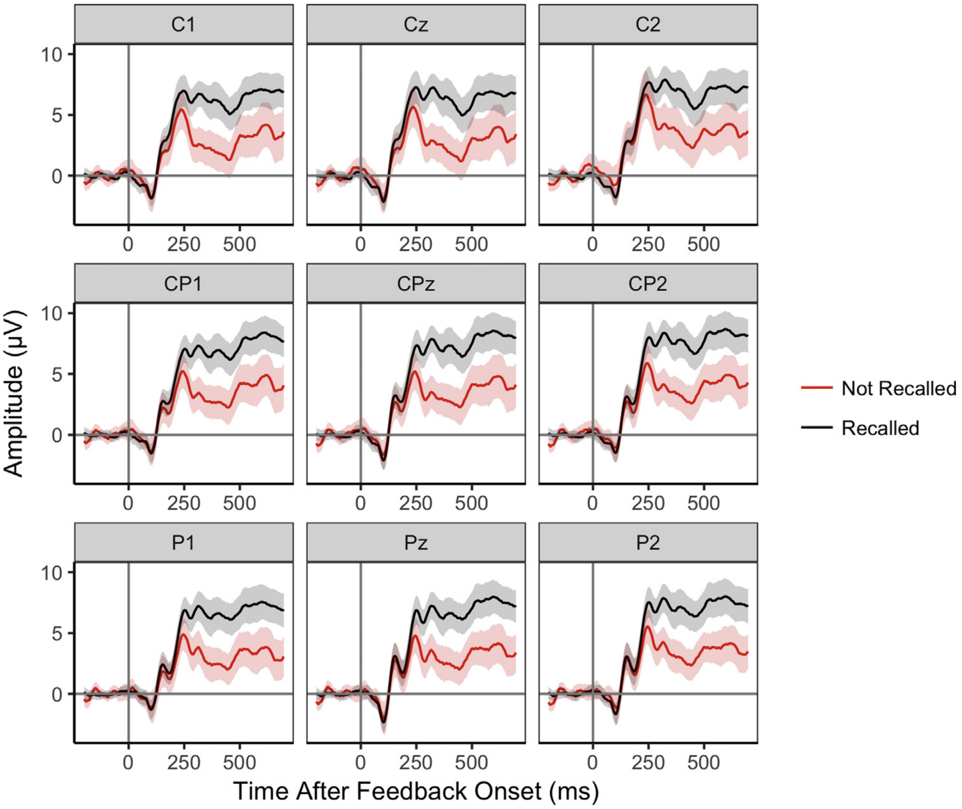

Fig. 4. ERPs to feedback showed enhanced late positivity for trials in which the item was subsequently recalled in the test phase. N = 26.

interaction (Fig. 6). Consistent with the previous results, TOT state positively predicted probability of subsequent recall such that recall was better for TOT than no-TOT trials (β = .94, 95% CI = [0.63, 1.24])[4]. Single-trial ERP amplitude also positively predicted probability of subsequent recall (β = .04, 95% CI = [0.03, 0.06]). There was no interaction between TOT state and ERP amplitude (β = .01, 95% CI = [−0.02, 0.04]).

To take into account effects of individual electrodes, we conducted a similar regression in which we did not average across electrodes, but instead allowed effects of TOT and amplitude to vary across both participants and our 9 electrodes of interest. This model produced very similar predictions to the first, such that TOT state (β = 1.03, 95% CI = [0.59, 1.51]) and ERP amplitude (β = .04, 95% CI = [0.01, 0.07]) both positively predicted recall probability. There was no interaction between TOT state and ERP amplitude (β = .01, 95% CI = [−0.02, 0.05]).

3.2.5. Mediation analysis

The above results suggested a causal model according to which feedback presented during a TOT state elicits greater positive centro-parietal ERP responses, and this ERP magnitude in turn leads to a higher probability of correctly recalling the item at a later test. To formally address this causal model, we performed a multilevel mediation analysis, with TOT as the independent variable, single-trial ERP (averaged across electrodes) as the mediator, and binary recalled/not recalled as the dependent variable[5] (Vuorre & Bolger, 2017; Vuorre, 2017). This mediation analysis (see Fig. 7) indicated that ERP amplitude partially mediated (indirect effect = 0.10, 95% CI = [0.04, 0.11]) the relationship between TOT states and subsequent recall (proportion mediated effect = 0.093, 95% CI = [0.036, 0.16]). A direct effect of TOT states on subsequent recall remained after taking ERP amplitudes into account (c′ = 1.02, 95% CI = [0.74, 1.29]). Although the mediation effect was partial, this finding lends support for a model in which the effect of TOT states on subsequent recall is mediated by ERP amplitudes.

---

[4] Note that the β for TOT state is the effect, in log-odds, of a categorical predictor coded as a 1 for TOT trials and 0 for no-TOT trials. β for ERP amplitude is the effect, in log-odds, of a continuous predictor with microvolts (µV) as units.

[5] This analysis was conducted using the same measure as our previous ERP analyses; single-trial averaged amplitude 250-700ms post-stimulus at centro-parietal electrodes. We modeled the binary outcomes with a logistic link function—therefore the relevant effects are in log-odds. We again used software-provided uninformative priors for all parameters.

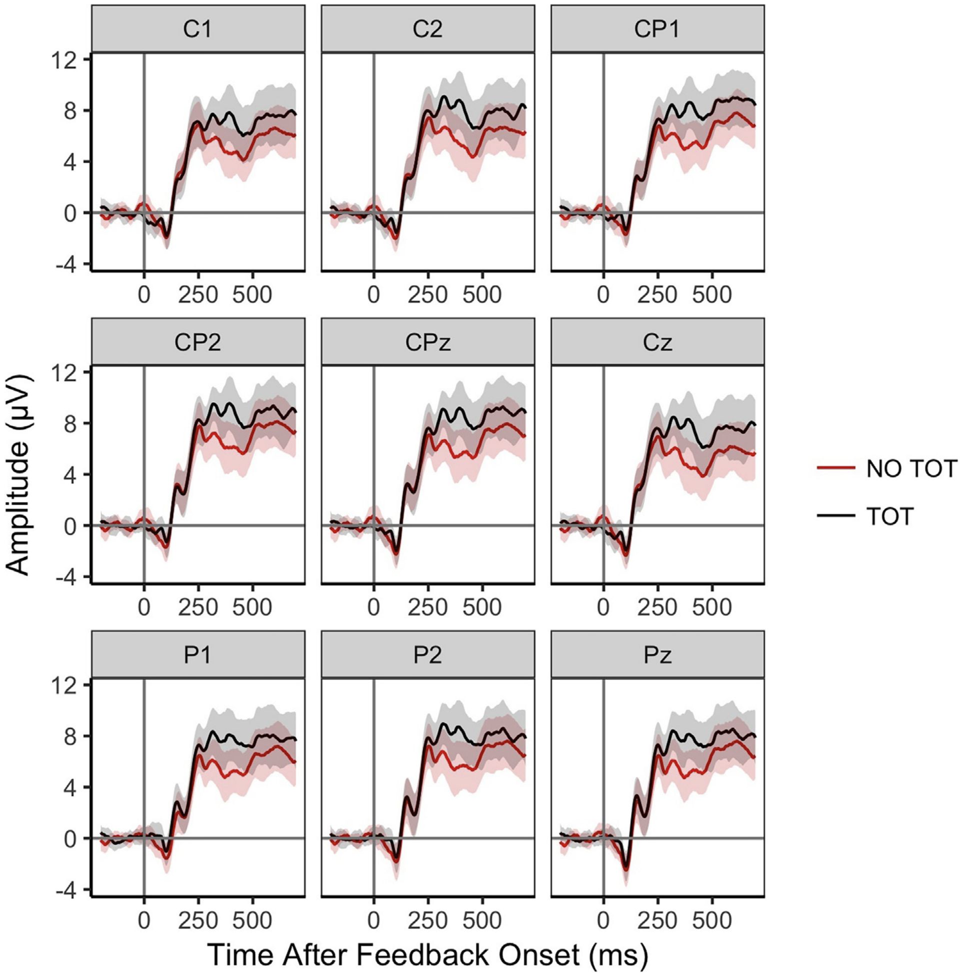

Fig. 5. ERPs to TOT feedback only on items that were subsequently recalled. N = 26.

4. Discussion

This experiment examined whether TOT states were associated with enhanced learning, and whether this relationship was mediated by neurocognitive processing of feedback, as measured through ERPs. In addition, we tested whether subsequent recall was improved for items in which TOTs had been evoked during the TOT/feedback phase. If neural processing of feedback mediated the relationship between TOT states and recall, we expected that feedback-evoked ERPs in TOT trials (compared to no-TOT trials) would be increased in amplitude, and that this increased amplitude would positively predict subsequent recall probability. The data indeed showed that TOT states preceding the feedback phase predicted subsequent recall, and that this relationship was partially mediated by ERP amplitudes over centro-parietal sites. Although our experimental design precludes direct claims of causality, our results support a model in which TOT states facilitate learning through heightened processing of feedback.

With our single-trial analyses we asked whether ERP amplitudes predicted subsequent recall. We found that items evoking higher average amplitudes over central-posterior electrodes in the TOT/feedback phase were more likely to be recalled in the test phase, and that this ERP measure mediated the relationship between TOT states and subsequent recall. All subjects appeared to similarly show this mediation effect (see S1). Consistent with the mediation model, higher ERP amplitudes were observed for feedback to TOT trials, compared to no-TOT trials, even between items in which the target was subsequently recalled. These results provide convergent support for a framework in which TOT states themselves contribute to improved learning due to enhanced neurocognitive processing of feedback.

The difference in ERP waveforms evoked by feedback in TOT versus no-TOT trials was highly similar to that observed for ERPs to subsequently remembered versus forgotten items in the Dm literature. While Dm effects do vary in both location and latency depending on the task, the enhanced late processing at centro-parietal electrodes that we observed in response to feedback during TOT states resembles the neural signatures evoked during encoding of subsequently remembered items in several studies employing semantic tasks (Friedman & Johnson, 2000; Friedman et al., 1996; Paller & Wagner, 2002; Paller et al., 1987; Van Petten & Senkfor, 1996). While we wish to be cautious about the extent to which we equate our TOT findings with Dm effects, the Dm literature can act

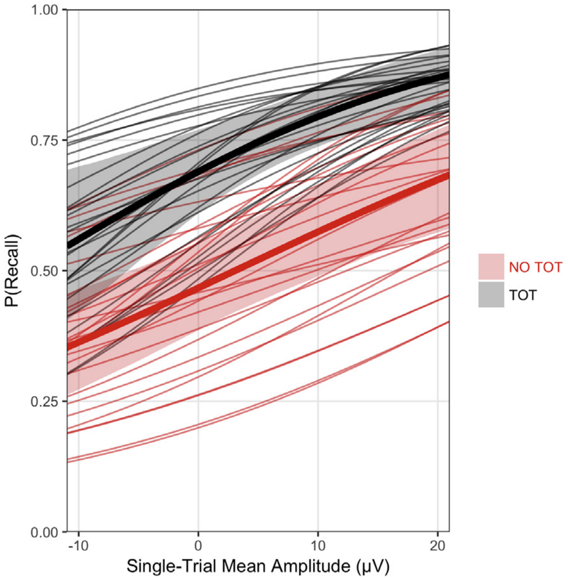

Fig. 6. Multilevel logistic regression model for post-test recall probability using single-trial EEG data. Amplitude averaged across the time window (250–700 ms post-feedback onset) was measured in microvolts (μV). Thick lines = population-level regression line; thin lines = subject-specific regression lines; shading = 95% CI of the population-level regression line. N = 26.

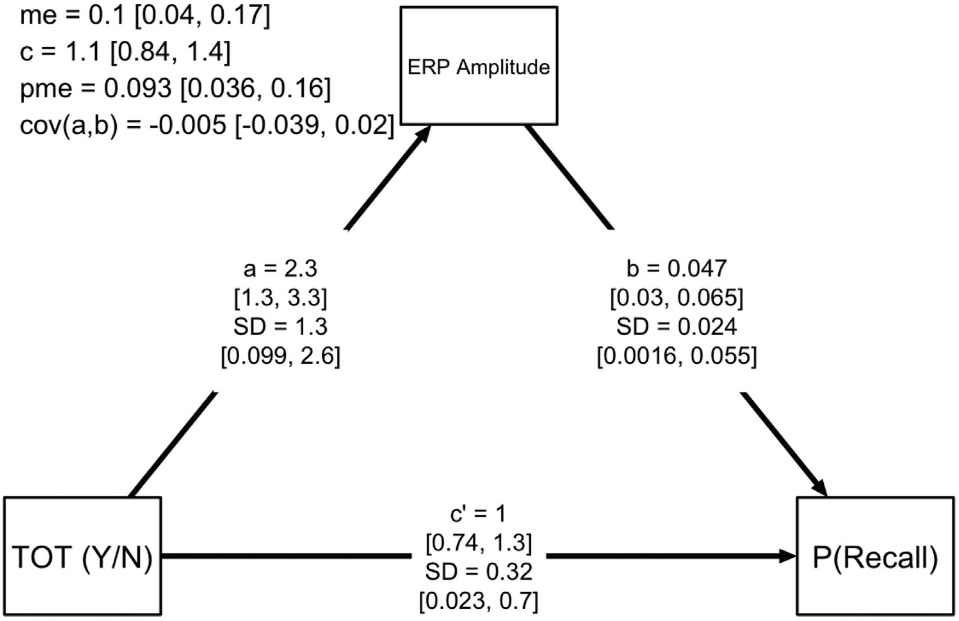

Fig. 7. Path diagram of the multilevel mediation model. c = total effect (direct + indirect effect of TOT on P(Recall), me = mediated effect, c′ = direct effect, pme = proportion of effect that is mediated, cov(a,b) = covariance of subject-level a and b parameters. Parameters are reported with 95% credible intervals in square brackets.

as a useful backdrop for understanding how TOT states predict greater encoding of feedback. Notably, our finding that ERP amplitudes predicted subsequent memory (independent of TOT states) for our task is itself a true Dm effect—subsequently remembered items display enhanced late processing in response to feedback.

We may also be able to contextualize the relationship between TOTs and subsequent recall within previous work on ‘hypercorrection’. In paradigms involving general information questions, high-confidence errors result in better learning of feedback, compared to low-confidence errors (Butler, Fazio, & Marsh, 2011; Butterfield & Metcalfe, 2001, 2006; Metcalfe et al., 2015; Metcalfe, 2017). Hypercorrection effects have been suggested to stem from heightened attention during the feedback following the error, as well as to the semantic similarity of the errors to the corrective feedback. The former postulated mechanism for the enhanced memory was supported both by evidence that the P3 component was increased to high-confidence error feedback (Butterfield & Mangels, 2003; Metcalfe et al, 2015) and by participants' failure to detect auditory signals in a secondary task during such feedback (Butterfield & Metcalfe, 2006). These latter results suggest that, in addition to the informational difference (Finn & Metcalfe, 2007; Metcalfe & Finn, 2011; Sitzman, Rhodes, Tauber, & Liceralde, 2015), there may be a feedback loop during hypercorrection that recruits attention and strengthens encoding processes (Butterfield & Metcalfe, 2006; Metcalfe, 1993). Further, items evoking high-confidence errors may reside within an individual's region of proximal learning and thus signal that it will be beneficial to invest effort into learning them (Metcalfe & Finn, 2011).

In the case of the TOT state specifically, it may be that heightened curiosity for the target information engages attention, boosts encoding, and leads to improved recall. In this model, the TOT state acts as a 'gain control' to amplify attention to and encoding of feedback, increasing the likelihood of subsequent memory. The enhanced encoding due to the TOT state thus spurs learning of feedback. Such a model is also consistent with predictions that TOT-evoking items sit within an individual's region of proximal learning, and thus recruit heightened attention to feedback. Indeed, similar work has suggested that there is an increase in processing at parietal electrodes when participants were attentive to and focused on learning, relative to when they were mind wandering (Xu, Friedman, & Metcalfe, 2018), further supporting the idea that participants might have been more attentive to feedback during TOT states.

It should be noted, though, that the observed differences in recall and in the ERPs to feedback might not be due exclusively to enhanced attentional processing. These differences might also be attributed in part to the likelihood that participants possessed more prior information for answers in which they experienced TOT states. Both metacognitive models such as Schwartz and Metcalfe's (2011) heuristic model and linguistic models such as Burke et al.'s (1991) models based upon Node Structure Theory (and see, Burke & Shafto, 2004), indicate that TOT experiences arise when there is considerable accumulated information from a variety of sources (syntactic, partial semantic, phonetic, or cue-based information). Subsequent memory is also strongest when there is a high degree of prior information available for the items being learned (Paller et al., 1987; Van Petten & Senkfor, 1996). It follows that enhanced memory for feedback to TOT states would be expected based only on prior information. Interpretation of the extent to which our results are due to curiosity and increased attention as compared to only the availability of semantic information remains a challenge: both components probably contributed to our effects.

While it is very likely, then, that differences in semantic access contributed to our results, the finding (1) of a mediating role of ERP amplitudes and (2) that ERP amplitudes were enhanced more for TOT as compared to non-TOT related feedback even when only items that were later recalled were examined provides support for the idea that enhanced processing, rather than only differences in the amount of information, plays a role in the learning associated with TOT states. To more directly address whether TOT states influence ERPs and subsequent learning above and beyond the effects of semantic access, future experimental designs are needed in which TOT states are evoked while the degree of semantic access is held constant.

Given the experimental design in the present study, we were only able to directly compare learning and ERPs to feedback for items where participants did not overtly provide an answer ('TOT' versus 'Don't Know' responses). Future studies are needed to compare long-term recall and neurocognitive responses to feedback following TOT states versus successful retrieval. In addition, the current design did not investigate 'blocked' TOTs, in which an intruding word occurred (Brown & McNeill, 1966; Burke et al., 1991; Jones & Langford, 1987; Kornell & Metcalfe, 2006). Research is needed on whether the ERP signatures and learning for blocked TOTs are the same or different from those shown here.

While the results of our study indicate that TOT states may facilitate learning when correct feedback is provided, we do note that it has also been suggested that the effects of TOTs may not be uniformly positive: they may result in error learning through implicit learning of an unsuccessful retrieval attempt (D'Angelo & Humphreys, 2015), and recurrence on a second test (Warriner & Humphreys, 2008). It is certainly possible that TOTs help learning through an initial provocation of curiosity, yet may have some harmful effects if no resolution is attained despite an extended time period. Nevertheless, the finding that being in a TOT state when correct feedback is provided is associated with better subsequent memory, as was found here, and that the neural correlates point to enhanced processing of the feedback, may have valuable applications for learning. More broadly, research on how initial failed retrievals and initial errors impact subsequent memory may contribute to the development of new strategies for improving learning in both educational and therapeutic settings.

Acknowledgements

We would like to thank Zach Bucknoff, Irma Alineri, Gilian Carling, Michael Helzer, Kelsey Ables, and the other members of the Memory & Metacognition Lab for their help and feedback. The authors are responsible for the data and the interpretation of the data reported here. This research was funded by a Natural Science and Engineering Research Council of Canada PGS-D3 award to J.X., an Institute for Education Sciences award (R305A150467) and a James S. McDonnell Foundation grant (220020166) to J.M, and a National Institute of Aging grant (#AG005213) to DF.

Appendix A. Supplementary material

Supplementary data associated with this article can be found, in the online version, at https://doi.org/10.1016/j.concog.2018.05.010.

References

Berlyne, D. E. (1954). A theory of human curiosity. *British Journal of Psychology. General Section, 45*(3), 180–191. http://dx.doi.org/10.1111/j.2044-8295.1954.tb01243.x.

Berlyne, D. E. (1960). *Conflict, arousal, and curiosity.* McGraw-Hill.

Brown, A. S. (1991). A review of the tip-of-the-tongue experience. *Psychological Bulletin, 109*(2), 204–223. http://dx.doi.org/10.1037/0033-2909.109.2.204.

Brown, R., & McNeill, D. (1966). The "tip of the tongue" phenomenon. *Journal of Verbal Learning and Verbal Behavior, 5*(4), 325–337. http://dx.doi.org/10.1016/S0022-5371(66)80040-3.

Burke, D. M., MacKay, D. G., Worthley, J. S., & Wade, E. (1991). On the tip of the tongue: What causes word finding failures in young and older adults? *Journal of Memory and Language, 30*(5), 542–579. http://dx.doi.org/10.1016/0749-596X(91)90026-G.

Burke, D. M., & Shafto, M. A. (2004). Aging and language production. *Current Directions in Psychological Science, 13*(1), 21–24. http://dx.doi.org/10.1111/j.0963-7214.2004.01301006.x.

Bürkner, P.-C. (2017). brms: Bayesian regression models using Stan (Version 1.6.1). Retrieved from https://cran.r-project.org/web/packages/brms/index.html.

Butler, A. C., Fazio, L. K., & Marsh, E. J. (2011). The hypercorrection effect persists over a week, but high-confidence errors return. *Psychonomic Bulletin & Review, 18*(6), 1238–1244. http://dx.doi.org/10.3758/s13423-011-0173-y.

Butterfield, B., & Mangels, J. A. (2003). Neural correlates of error detection and correction in a semantic retrieval task. *Cognitive Brain Research, 17*(3), 793–817. http://dx.doi.org/10.1016/S0926-6410(03)00203-9.

Butterfield, B., & Metcalfe, J. (2001). Errors committed with high confidence are hypercorrected. *Journal of Experimental Psychology. Learning, Memory, and Cognition, 27*(6), 1491–1494.

Butterfield, B., & Metcalfe, J. (2006). The correction of errors committed with high confidence. *Metacognition and Learning, 1*(1), 69–84. http://dx.doi.org/10.1007/s11409-006-6894-z.

Cohen, G., & Faulkner, D. (1986). Memory for proper names: Age differences in retrieval. *British Journal of Developmental Psychology, 4*(2), 187–197. http://dx.doi.org/10.1111/j.2044-835X.1986.tb01010.x.

D'Angelo, M. C., & Humphreys, K. R. (2015). Tip-of-the-tongue states reoccur because of implicit learning, but resolving them helps. *Cognition, 142*, 166–190. http://dx.doi.org/10.1016/j.cognition.2015.05.019.

Delorme, A., & Makeig, S. (2004). EEGLAB: An open source toolbox for analysis of single-trial EEG dynamics including independent component analysis. *Journal of Neuroscience Methods, 134*(1), 9–21. http://dx.doi.org/10.1016/j.jneumeth.2003.10.009.

Díaz, F., Lindín, M., Galdo-Álvarez, S., & Buján, A. (2014). Neurofunctional correlates of the tip-of-the-tongue state. In B. Schwartz & A. Brown (Eds.), Tip-of-the-tongue states and related phenomena (pp. 198–231). Cambridge: Cambridge University Press. http://doi.org/10.1017/CBO9781139547383.010.

Díaz, F., Lindín, M., Galdo-Alvarez, S., Facal, D., & Juncos-Rabadán, O. (2007). An event-related potentials study of face identification and naming: The tip-of-the-tongue state. *Psychophysiology, 44*(1), 50–68. http://dx.doi.org/10.1111/j.1469-8986.2006.00483.x.

Fabiani, M., & Donchin, E. (1995). Encoding processes and memory organization: A model of the von Restorff effect. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 21*(1), 224–240. http://dx.doi.org/10.1037/0278-7393.21.1.224.

Fabiani, M., Karis, D., & Donchin, E. (1986). P300 and recall in an incidental memory paradigm. *Psychophysiology, 23*(3), 298–308. http://dx.doi.org/10.1111/j.1469-8986.1986.tb00636.x.

Fabiani, M., Karis, D., & Donchin, E. (1990). Effects of mnemonic strategy manipulation in a Von Restorff paradigm. *Electroencephalography and Clinical Neurophysiology, 75*(1–2), 22–35. http://dx.doi.org/10.1016/0013-4694(90)90149-E.

Finn, B., & Metcalfe, J. (2007). The role of memory for past test in the underconfidence with practice effect. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 33*(1), 238–244. http://dx.doi.org/10.1037/0278-7393.33.1.238.

Friedman, D. (1990). ERPs during continuous recognition memory for words. *Biological Psychology, 30*(1), 61–87. http://dx.doi.org/10.1016/0301-0511(90)90091-A.

Friedman, D. (2003). Cognition and aging: A highly selective overview of event-related potential (ERP) data. *Journal of Clinical and Experimental Neuropsychology, 25*(5), 702–720. http://dx.doi.org/10.1076/jcen.25.5.702.14578.

Friedman, D., & Johnson, R. (2000). Event-related potential (ERP) studies of memory encoding and retrieval: A selective review. *Microscopy Research and Technique, 51*(1), 6–28. http://dx.doi.org/10.1002/1097-0029(20001001)51:1<6::AID-JEMT2>3.0.CO;2-R.

Friedman, D., Ritter, W., & Snodgrass, J. G. (1996). ERPs during study as a function of subsequent direct and indirect memory testing in young and old adults. *Cognitive Brain Research, 4*(1), 1–13. http://dx.doi.org/10.1016/0926-6410(95)00041-0.

Friedman, D., & Sutton, S. (1987). Event-related potentials during continuous recognition memory. *Electroencephalography and Clinical Neurophysiology. Supplement, 40*, 316–321.

Friedman, D., & Trott, C. (2000). An event-related potential study of encoding in young and older adults. *Neuropsychologia, 38*(5), 542–557. http://dx.doi.org/10.1016/S0028-3932(99)00122-0.

Gardiner, F. M., Craik, F. I. M., & Bleasdale, F. A. (1973). Retrieval difficulty and subsequent recall. *Memory & Cognition, 1*(3), 213–216. http://dx.doi.org/10.3758/BF03198098.

Gelman, A., Lee, D., & Guo, J. (2015). Stan: A probabilistic programming language for Bayesian inference and optimization. *Journal of Educational and Behavioral Statistics, 40*(5), 530–543. http://dx.doi.org/10.3102/1076998615606113.

James, W. (1890). The principles of psychology. Read Books Ltd.

Jennings, R., & Wood, C. C. (1976). The ε-adjustment procedure for repeated-measures analyses of variance. *Psychophysiology, 13*(3), 277–278. http://dx.doi.org/10.1111/j.1469-8986.1976.tb00116.x.

Jones, G. V., & Langford, S. (1987). Phonological blocking in the tip of the tongue state. *Cognition, 26*(2), 115–122. http://dx.doi.org/10.1016/0010-0277(87)90027-8.

Kang, M. J., Hsu, M., Krajbich, I. M., Loewenstein, G., McClure, S. M., Wang, J. T., & Camerer, C. F. (2009). The wick in the candle of learning: Epistemic curiosity activates reward circuitry and enhances memory. *Psychological Science, 20*(8), 963–973. http://dx.doi.org/10.1111/j.1467-9280.2009.02402.x.

Kikyo, H., Ohki, K., & Sekihara, K. (2001). Temporal characterization of memory retrieval processes: An fMRI study of the "tip of the tongue" phenomenon. *European Journal of Neuroscience, 14*(5), 887–892. http://dx.doi.org/10.1046/j.0953-816x.2001.01711.x.

Koriat, A., & Lieblich, I. (1974). What does a person in a "TOT" state know that a person in a "don't know" state doesn't know. *Memory & Cognition, 2*(4), 647–655. http://dx.doi.org/10.3758/BF03198134.

Kornell, N., & Metcalfe, J. (2006). "Blockers" do not block recall during tip-of-the-tongue states. *Metacognition and Learning, 1*(3), 248–261. http://dx.doi.org/10.1007/s11409-007-9003-z.

Litman, J., Hutchins, T., & Russon, R. (2005). Epistemic curiosity, feeling-of-knowing, and exploratory behaviour. *Cognition and Emotion, 19*(4), 559–582. http://dx.doi.org/10.1080/02699930441000427.

Litman, J. A., & Spielberger, C. D. (2003). Measuring epistemic curiosity and its diversive and specific components. *Journal of Personality Assessment, 80*(1), 75–86. http://dx.doi.org/10.1207/S15327752JPA8001_16.

Loewenstein, G. (1994). The psychology of curiosity: A review and reinterpretation. *Psychological Bulletin, 116*(1), 75. http://dx.doi.org/10.1037/0033-2909.116.1.75.

Lopez-Calderon, J., & Luck, S. J. (2014). ERPLAB: An open-source toolbox for the analysis of event-related potentials. *Frontiers in Human Neuroscience, 8*. http://dx.doi.org/10.3389/fnhum.2014.00213.

Luck, S. J. (2014). *An introduction to the event-related potential technique.* MIT Press.

MacKay, D. G., & Burke, D. M. (1990). Chapter five cognition and aging: A theory of new learning and the use of old connections. In F. G. A. Stone & R. West (Eds.), Advances in psychology (Vol. 71, pp. 213–263). North-Holland. https://doi.org/10.1016/S0166-4115(08)60159-4.

Makeig, S., & Onton, J. (2011). ERP features and EEG dynamics. *The Oxford Handbook of Event-Related Potential Components*. http://dx.doi.org/10.1093/oxfordhb/9780195374148.013.0035.

Maril, A., Simons, J. S., Weaver, J. J., & Schacter, D. L. (2005). Graded recall success: An event-related fMRI comparison of tip of the tongue and feeling of knowing.

NeuroImage, 24(4), 1130–1138. http://dx.doi.org/10.1016/j.neuroimage.2004.10.024.

Maril, A., Wagner, A. D., & Schacter, D. L. (2001). On the tip of the tongue: An event-related fmri study of semantic retrieval failure and cognitive conflict. *Neuron, 31*(4), 653–660. http://dx.doi.org/10.1016/S0896-6273(01)00396-8.

Metcalfe, J. (2002). Is study time allocated selectively to a region of proximal learning? *Journal of Experimental Psychology: General, 131*(3), 349. http://dx.doi.org/10.1037/0096-3445.131.3.349.

Metcalfe, J. (2017). Learning from errors. *Annual Review of Psychology, 68*(1), 465–489. http://dx.doi.org/10.1146/annurev-psych-010416-044022.

Metcalfe, J., Casal-Roscum, L., Radin, A., & Friedman, D. (2015). On teaching old dogs new tricks. *Psychological Science, 26*(12), 1833–1842. http://dx.doi.org/10.1177/0956797615597912.

Metcalfe, J., Schwartz, B.L., & Bloom, P. A. (2017). The tip-of-the-tongue state and curiosity. Cognitive Research: Principles and Implications. http://doi.org/10.1186/s41235-017-0065-4.

Metcalfe, J., & Finn, B. (2011). People's hypercorrection of high confidence errors: Did they know it all along? *Journal of Experimental Psychology Learning, Memory, and Cognition, 37*(2), 437–448. http://dx.doi.org/10.1037/a0021962.

Metcalfe, J., & Kornell, N. (2005). A region of proximal learning model of study time allocation. *Journal of Memory and Language, 52*(4), 463–477. http://dx.doi.org/10.1016/j.jml.2004.12.001.

Metcalfe, J. (1993). Monitoring and gain control in an episodic memory model: relation to P300 event-related potentials. In A. Collins, S. Gathercole, M. A. Conway, & P. E. Morris (Eds.), Theories of memory (pp. 327–354). Lawrence Erlbaum.

Miozzo, M., & Caramazza, A. (1997). Retrieval of lexical-syntactic features in tip-of-the-tongue states. *Journal of Experimental Psychology. Learning, Memory, and Cognition, 23*(6), 1410–1423. http://dx.doi.org/10.1037//0278-7393.23.6.1410.

Münte, T. F., Heinze, H. J., Scholz, M., & Künkel, H. (1988). Effects of a cholinergic nootropic (WEB 1881 FU) on event-related potentials recorded in incidental and intentional memory tasks. *Neuropsychobiology, 19*(3), 158–168. http://dx.doi.org/10.1159/000118453.

Nelson, T. O., & Narens, L. (1980). Norms of 300 general-information questions: Accuracy of recall, latency of recall, and feeling-of-knowing ratings. *Journal of Verbal Learning and Verbal Behavior, 19*(3), 338–368. http://dx.doi.org/10.1016/S0022-5371(80)90266-2.

Paller, K. A. (1990). Recall and stem-completion priming have different electrophysiological correlates and are modified differentially by directed forgetting. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 16*(6), 1021–1032. http://dx.doi.org/10.1037/0278-7393.16.6.1021.

Paller, K. A., & Kutas, M. (1992). Brain potentials during memory retrieval provide neurophysiological support for the distinction between conscious recollection and priming. *Journal of Cognitive Neuroscience, 4*(4), 375–392. http://dx.doi.org/10.1162/jocn.1992.4.4.375.

Paller, K. A., Kutas, M., & Mayes, A. R. (1987). Neural correlates of encoding in an incidental learning paradigm. *Electroencephalography and Clinical Neurophysiology, 67*(4), 360–371.

Paller, K. A., McCarthy, G., & Wood, C. C. (1988). ERPs predictive of subsequent recall and recognition performance. *Biological Psychology, 26*(1–3), 269–276. http://dx.doi.org/10.1016/0301-0511(88)90023-3.

Paller, K. A., & Wagner, A. D. (2002). Observing the transformation of experience into memory. *Trends in Cognitive Sciences, 6*(2), 93–102. http://dx.doi.org/10.1016/S1364-6613(00)01845-3.

Peirce, J. W. (2007). PsychoPy—psychophysics software in Python. *Journal of Neuroscience Methods, 162*(1–2), 8–13. http://dx.doi.org/10.1016/j.jneumeth.2006.11.017.

Ripley, B., & Hornik, K. (2010). agrep: Approximate string matching (Fuzzy Matching) (Version 3.4.0). Retrieved from https://stat.ethz.ch/R-manual/R-devel/library/base/html/agrep.html.

Ryan, M. P., Petty, C. R., & Wenzlaff, R. M. (1982). Motivated remembering efforts during tip-of-the-tongue states. *Acta Psychologica, 51*(2), 137–147. http://dx.doi.org/10.1016/0001-6918(82)90058-0.

Sanquist, T. F., Rohrbaugh, J. W., Syndulko, K., & Lindsley, D. B. (1980). Electrocortical signs of levels of processing: Perceptual analysis and recognition memory. *Psychophysiology, 17*(6), 568–576. http://dx.doi.org/10.1111/j.1469-8986.1980.tb02299.x.

Schwartz, B. L. (2001). The relation of tip-of-the-tongue states and retrieval time. *Memory & Cognition, 29*(1), 117–126. http://dx.doi.org/10.3758/BF03195746.

Schwartz, B. L. (2006). Tip-of-the-tongue states as metacognition. *Metacognition and Learning, 1*(2), 149–158. http://dx.doi.org/10.1007/s11409-006-9583-z.

Schwartz, B. L., & Díaz, F. (2014). Quantifying human metacognition for the neurosciences. In The Cognitive Neuroscience of Metacognition (pp. 9–23). Springer, Berlin, Heidelberg. https://doi.org/10.1007/978-3-642-45190-4_2.

Schwartz, B. L., & Metcalfe, J. (2011). Tip-of-the-tongue (TOT) states: Retrieval, behavior, and experience. *Memory & Cognition, 39*(5), 737–749. http://dx.doi.org/10.3758/s13421-010-0066-8.

Sitzman, D. M., Rhodes, M. G., Tauber, S. K., & Liceralde, V. R. T. (2015). The role of prior knowledge in error correction for younger and older adults. *Aging, Neuropsychology, and Cognition, 22*(4), 502–516. http://dx.doi.org/10.1080/13825585.2014.993302.

Smith, M. E. (1993). Neurophysiological manifestations of recollective experience during recognition memory judgments. *Journal of Cognitive Neuroscience, 5*(1), 1–13. http://dx.doi.org/10.1162/jocn.1993.5.1.1.

Van Petten, C., & Senkfor, A. J. (1996). Memory for words and novel visual patterns: Repetition, recognition, and encoding effects in the event-related brain potential. *Psychophysiology, 33*(5), 491–506. http://dx.doi.org/10.1111/j.1469-8986.1996.tb02425.x.

Vuorre, M., & Bolger, N. (2017). Within-subject mediation analysis for experimental data in cognitive psychology and neuroscience. *Behavior Research Methods, 1–19*. http://dx.doi.org/10.3758/s13428-017-0980-9.

Vuorre, M. (2017). bmlm: Bayesian multilevel mediation. Retrieved from https://cran.r-project.org/package=bmlm.

Wagner, A. D., Koutstaal, W., & Schacter, D. L. (1999). When encoding yields remembering: Insights from event-related neuroimaging. *Philosophical Transactions of the Royal Society B: Biological Sciences, 354*(1387), 1307–1324.

Wagner, A. D., Schacter, D. L., Rotte, M., Koutstaal, W., Maril, A., Dale, A. M., ... Buckner, R. L. (1998). Building memories: Remembering and forgetting of verbal experiences as predicted by brain activity. *Science, 281*(5380), 1188–1191.

Warriner, A. B., & Humphreys, K. R. (2008). Learning to fail: Reoccurring tip-of-the-tongue states. *The Quarterly Journal of Experimental Psychology, 61*(4), 535–542. http://dx.doi.org/10.1080/17470210701728867.

Weyerts, H., Tendolkar, I., Smid, H. G., & Heinze, H. J. (1997). ERPs to encoding and recognition in two different inter-item association tasks. *Neuroreport, 8*(7), 1583–1588.

Xu, J., Friedman, D., & Metcalfe, J. (2018). Attenuation of deep semantic processing during mind wandering: An event-related potential study. *NeuroReport, 29*(5), 380. http://dx.doi.org/10.1097/WNR.0000000000000978.

Xu, J., & Metcalfe, J. (2016). Studying in the region of proximal learning reduces mind wandering. *Memory & Cognition, 44*(5), 681–695. http://dx.doi.org/10.3758/s13421-016-0589-8.