P-ISSN 2664-0422	Pakistan Social Sciences Review	March-April 2026, Vol. 10, No.2
O-ISSN 2664-0430	https://doi.org/10.35484/pssr.2026(10-II)05	[55-74]

Pakistan Social Sciences Review
www.pssr.org.pk

RESEARCH PAPER

Arousal Attenuates Directed Forgetting of Affective and Neutral English Words in an Item-Method Recognition Paradigm

[1]Suleman Ahmad, [2]Ather Mujitaba and [3] Amina Afzal

1. BS Clinical Psychology, Department of Humanities and Social Science, GIFT University, Gujranwala, Punjab, Pakistan
2. Lecturer, Department of Humanities and Social Science, GIFT University, Gujranwala, Punjab, Pakistan
3. PhD Scholar, Faculty of Psychology, Southwest University, China. ORCID: https:0009-0004-1871-3794

Corresponding Author: Sulemanahmad.0148@gmail.com

ABSTRACT

This study examined whether item-method directed forgetting (DF) for English words is better explained by affective categories or item-level valence, arousal, and dominance (VAD). Directed forgetting is a laboratory model of intentional memory control, but it remains unclear whether forgetting varies more by emotional category of words or their emotional dimensions. This distinction has received limited direct testing. A quantitative experimental design was used with 67 Pakistani adults recruited through convenience sampling; after prespecified exclusions, 60 were analyzed. Participants studied positive, neutral, negative, and threatening words followed by remember or forget cues, then completed a recognition test. Category effects were tested with binomial logistic regression, and dimensional effects with weighted models. A robust DF effect emerged. It was largest for neutral words and smallest for threatening words. Arousal significantly weakened the forget effect, whereas valence and dominance did not. Future work should prioritize arousal-sensitive models and culturally grounded verbal stimuli.

KEYWORDS Directed Forgetting, Arousal, Affective Words, Recognition Memory, Item-Method Paradigm

Introduction

Directed forgetting (DF) is widely used as a laboratory model of intentional memory control because memory is typically poorer for items followed by a forget cue than for items followed by a remember cue (Taylor et al., 2018; Zwissler et al., 2015). In the item-method paradigm, each stimulus is followed by an explicit instruction to remember or forget, allowing researchers to examine how top-down goals shape later memory for individual items (Anderson & Hanslmayr, 2014; MacLeod, 1999; Taylor et al., 2018). Although the basic DF effect is robust, it is not equally strong across stimulus classes (Quinlan et al., 2010; Scotti & Maxcey, 2022). Emotional material is often less susceptible to intentional forgetting than neutral material, suggesting that memory control depends not only on cue-based goals but also on the affective characteristics of the information being processed (Hall et al., 2021).

This asymmetry is theoretically important because it challenges a purely instruction-driven view of directed forgetting. Item-method accounts have commonly emphasized selective rehearsal and related post-cue encoding differences, according to which remember cues support further processing of the just-presented item, whereas forget cues reduce such processing (Bjork, 1972; Lee, 2013; Tan et al., 2020). If forget cues reliably reduce later memory, intentional forgetting may appear to reflect top-down control alone. However, if affectively salient items remain relatively well remembered despite forget instructions, then the efficacy of memory control must also depend on properties of the stimuli themselves (Hall et al., 2021; Nowicka et al., 2011). In that case, directed forgetting is constrained not only by what people are instructed to do, but also by the mnemonic priority of the material they are trying to forget (Anderson & Hanslmayr, 2014; Bailey & Chapman, 2012).

Literature Review

The emotional DF literature points in that direction, but the mechanism remains unclear. A meta-analysis of item-method DF found that emotional material was, on average, harder to forget than neutral material, while also documenting substantial heterogeneity across studies (Hall et al., 2021). That heterogeneity suggests that the central issue is no longer whether affect matters for DF in a broad sense, but which affective properties reduce the efficacy of forget cues. Much of the literature still relies on broad contrasts such as emotional versus neutral or negative versus neutral. Although such comparisons are descriptively useful, they do not identify the property responsible for reduced forgetting. Recent work further suggests that emotional influences on directed remembering and forgetting depend on how affect is instantiated in the materials and task, especially at encoding (Chen et al., 2025). Related results from research in emotional word memory suggest that, similarly, emotion dimensions of valence and arousal can have independent effects on memory performance, and that the distinction between the categories of positive and negative emotion is insufficiently nuanced (Gao et al., 2024).

The domain of words is especially pertinent in this regard. Compared to other potential domains of study, such as complex scenes, faces, and autobiographical events, words have the advantage of allowing for a high degree of experimental control while still allowing for significant variability in affective salience. Just as importantly, words can be characterized using normative affective ratings at the item level. In affective word research, valence, arousal, and dominance (VAD) constitute a standard dimensional framework for describing emotional meaning, and large-scale English norms allow these properties to be modelled continuously rather than inferred solely from category labels (Warriner et al., 2013). This makes affectively varied word stimuli especially well-suited to testing whether reduced forgetting is better explained by broad category membership or by the underlying dimensions that define those categories.

Among the VAD dimensions, arousal is the strongest theoretical candidate. Arousal-biased competition theory proposes that arousal amplifies processing in favor of high-priority representations, strengthening the processing of salient material while disadvantaging lower-priority alternatives (Mather & Sutherland, 2011). In the context of directed forgetting, this framework implies that highly arousing items may remain relatively resistant to forget cues because their mnemonic priority is not fully neutralized by top-down instructions to disengage (Mather, 2007; Mather & Sutherland, 2011). From this perspective, reduced forgetting of affective verbal material should not be attributed to emotion in a generic sense, but rather to the extent to which arousal preserves the priority of an item despite a forget cue. Recent evidence that emotional influences on remembering and forgetting depend on emotion-related encoding dynamics is broadly consistent with this view (Chen et al., 2025).

This logic is particularly relevant when it comes to threatening words. The point is that threatening words are not simply negative, as neutral words are; threatening words tend to be negative, yet also arousing, and low in perceived controllability

(Bradley & Lang, 1999). If threatening items do prove to be particularly resilient against forget cues, it is then relevant to ask whether this is due to category membership, or to the affective profile that is typically instantiated by this category. The broader point, of course, is that if attenuation of DF is best accounted for in terms of arousal, then it is relevant to ask whether the advantage of threatening material is seen as part of a more general process.

The present study was intended to assess this question with an item-method recognition design and affective and neutral English words. The present study was also intended to follow both category-based and dimensional approaches by testing four words types that included affective and neutral words: Positive, Neutral, Negative, and Threatening, which were already normed in Pakistani youth (Mujitaba et al., 2025). Any disorder-specific lexical word lists were avoided (Mujitaba & Riaz, 2025), as the present study mainly focused on a nonclinical smaple. The combined category-based and dimensional approaches are beneficial, because category-based and dimensional analyses are non-identical but related questions (Hamann, 2012). The category-based and dimensional analyses provide different types of information. The category-based analysis is beneficial since it is more interpretable at the broad level.

Consequently, three questions were addressed in this study. First, would the standard DF effect be found across the entire range of materials? Again, in line with the larger item-method literature (e.g., Basden, 1996; MacLeod, 2012), it was predicted that recognition memory would be superior for remember-cued than for forget-cued items. Second, would an affective category be found to influence the degree of directed forgetting? The assumption was that, relative to affectively neutral categories, directed forgetting would be found to be smaller for affectively salient categories, with threatening categories showing the smallest directed forgetting effect, negative categories intermediate to threatening and neutral, and positive categories intermediate to negative and neutral, though to a lesser degree than threatening (Wessel & Merckelbach, 2006). Third, would the properties of continuous VADs influence directed forgetting? Arousal was again the focal moderator, and the central prediction was that the R-F difference would decrease with increasing arousal. Valence and dominance were again included as secondary predictors, but their roles are seen as much less certain because of their conceptual and empirical overlap with category structure. The study attempted to go beyond the simple emotional vs. neutral contrast and provide a more detailed account of the role of arousal and other affective properties in intentional forgetting.

Hypotheses

- Recognition memory would be better for remember-cued items than for forget-cued items, reflecting the standard directed forgetting effect.
- Directed forgetting will be vary across affective categories, with the smallest effect for threatening words, and intermediate effect for negative and positive words and largest effect for neutral words.
- Arousal would reduce the difference between remember-cued and forget-cued items, whereas valence and dominance would be examined as additional predictor.

Material and Methods

The study used a quantitative mixed factorial experimental design and employed the item-method directed forgetting paradigm and a computer-assisted recognition memory task (Basden, 1996). Cue Condition (Remember, Forget), Affective Category (Positive, Neutral, Negative, Threatening), and List (A, B) were the independent variables. List was the between-participants counterbalance factor.

This design was used to test directed forgetting at the level of individual lexical items while also examining whether the magnitude of the cue effect varied across affective categories and as a function of item-level affective properties. The within-subject manipulation of Cue and Category increased sensitivity to directed forgetting effects. In contrast, the list factor ensured that each Study Phase item appeared equally often in the Remember and Forget conditions across participants.

Participants

The target population for this study included English-fluent adults who lived in Gujranwala, Pakistan, and were affiliated with higher education institutions. The participants were a convenience sample of 60 Pakistani adults recruited from GIFT University and its surrounding areas. The inclusion criteria for participants were that they had to be at least 18 years old, had to be English-fluent enough to perform a computer-based lexical memory task, and had to have normal vision with or without corrective measures. Participants who had uncorrected vision problems that could impact their ability to read were excluded from the study. Participants were made aware that some of the words could be unpleasant or threat-related and that they could terminate their participation at any time without penalty. All participants provided informed consent for their participation in the study. This study was approved by the ethics review committee of university and follows the Ethical Principles of Psychologists and Code of Conduct as prescribed by the American Psychological Association (APA, 2017).

Participants were assigned in approximately equal numbers to one of two counterbalancing lists, List A (n = 30) and List B (n = 30). Assignment to list served only to counterbalance item-to-cue pairing and was not treated as a substantive experimental manipulation. The final analyzed sample had a mean age of 22.07 years (SD = 2.84, range = 18–30). Of the participants, 31 (51.7%) were men and 29 (48.3%) were women. Most participants were single (91.7%), and the majority were undergraduate students (76.7%). Additional demographic characteristics by counterbalancing list are presented in Table 1.

Table 1
Participant demographics by counterbalancing list (N=60)

	Total (N=60)		**List A (n=30)**		**List B (n=30)**	
Variable	N	%	N	%	N	%
Gender						
Male	31	51.7	16	53.3	15	50.0
Female	29	48.3	14	46.7	15	50.0
Marital status						
Single	55	91.7	28	93.3	27	90.0
Married	5	8.3	2	6.7	3	10.0
Family system						
Nuclear	31	51.7	16	53.3	15	50.0
Joint	29	48.3	14	46.7	15	50.0

Residence						
Urban	43	71.7	23	76.7	20	66.7
Rural	17	28.3	7	23.3	10	33.3
Family income (PKR/month)						
≤ 60k	25	41.7	14	46.7	11	36.7
61k–100k	23	38.3	11	36.7	12	40.0
> 100k	12	20.0	5	16.7	7	23.3
Education level						
Undergraduate	46	76.7	21	70.0	25	83.3
Graduate (MS/MPhil)	14	23.3	9	30.0	5	16.7
Graduate (PhD)	0	0.0	0	0.0	0	0.0
Prior DF task experience						
No	53	88.3	25	83.3	28	93.3
Yes	3	5.0	2	6.7	1	3.3
Not sure	4	6.7	3	10.0	1	3.3
Prior memory experiments						
No	60	100.0	30	100.0	30	100.0
Yes	0	0.0	0	0.0	0	0.0

Note. %=percentage, n= Number of Participants

Study Phase Items

The Study Phase stimulus set consisted of 80 English words drawn from a Pakistani normed affective and neutral English word pool developed by Mujitaba et al. (2025). The set comprised 20 Positive, 20 Neutral, 20 Negative, and 20 Threaening words. Category membership was based on previously available valence, arousal, and dominance ratings (Bradley & Lang, 1999; Warriner et al., 2013). Positive words were high in valence, Neutral words were near the affective midpoint and relatively low in arousal, these words were also tested in memory recall (Nadeem et al., 2026), Negative words were low in valence with moderate arousal, and Threatening words were low in valence but comparatively higher in arousal and lower in dominance. Category labels were used only for stimulus selection and later analysis and were not shown to participants during the task. Each Study Phase word was linked to its item-level valence, arousal, and dominance values for the dimensional analyses. Thus, the study tested both broad affective category differences and continuous affective properties of the words.

In the present analyses, word length was included as a lexical covariate in the dimensional model. Other lexical variables, such as frequency, familiarity, concreteness, imageability, and age of acquisition (Brysbaert et al., 2013; Brysbaert et al., 2016; Kuperman et al., 2012), were not explicitly modelled and therefore remain possible sources of residual item-level variation.

Distractors

An additional set of 80 English distractor words was presented only during the recognition phase. As with the Study Phase set, the distractor set contained 20 words from each affective category. Distractors were selected from the same normed pool to approximate the studied items in the affective range and general lexical form while ensuring that none had appeared during the study phase. These distractor items permitted estimation of false alarms and correct rejections.

Apparatus and Response Display

The experiment was administered individually on computers in a controlled laboratory setting. The task was built in OpenSesame version 4.1 (Mathôt, Schreij, & Theeuwes, 2012) with the OSWeb extension (Version 2.2.9.0; Mathôt, 2025) and was run through a local network setup in which one computer served as the local host, and connected devices accessed the task through the same Wi-Fi network using Microsoft Edge on Windows 10. Stimuli were presented on 14-inch displays with a 1920 × 1080 resolution and a 60 Hz refresh rate. Participants were seated at a distance of approximately 50–70 cm from the screen. During the study phase, words were shown centrally in 28-point Times New Roman, black, on a light blue background. The cue appeared centrally in 28-point Times New Roman, black, on a white background, with the fixation cross also displayed in the same style and position in white background. During the recognition phase, participants responded using a six-option source-confidence format: Old Forget (Sure), Old Forget (Unsure), Old Remember (Sure), Old Remember (Unsure), New (Sure), and New (Unsure). This response format allowed participants to indicate whether a probe word had appeared previously, whether they believed it had originally been paired with a Remember or Forget cue, and how confident they were in that judgment (Hennessee et al., 2019; Pastötter et al., 2016).

Procedure

The experiment consisted of three phases: a study phase, a brief retention interval, and a recognition test. Two counterbalancing lists were constructed to avoid confounding cue assignment with item memorability (Basden, 1996; MacLeod, 1999). In each list, every affective category contained 20 studied items, with 10 assigned to Remember and 10 assigned to Forget. Across the two lists, cue assignments were reversed. Thus, a word presented as a Remember item in List A appeared as a Forget item in List B, and vice versa. Across participants, each studied word therefore contributed equally to both cue conditions. This procedure reduced the possibility that any observed cue effect could be attributed solely to stable differences in the memorability of particular items.

Study Phase

Participants completed 80 study trials. On each trial, a single English word was presented centrally in 28-point Times New Roman, black text, on a light blue background for 2.0 s, followed by a white background with an instruction cue presented for 2.5 s, as shown in Figure 1. The cue was either "Remember" or "Forget" (Basden, 1996). After the cue, a blank fixation screen with a white background was shown for 0.5 s, after which the next trial began. Participants were instructed to memories words following the Remember cue, because their memory for those items would later be tested. For words followed by the Forget cue, they were instructed to stop rehearsing the word and allow it to leave their mind because it was no longer relevant for later performance (MacLeod, 1999). Across the study phase, each participant saw 40 Remember-cued words and 40 Forget-cued words. Within each affective category, 10 items were paired with Remember and 10 items with Forget. The order of trials was randomized separately for each individual, keeping track of the necessary counts of categories by cue (Minnema & Knowlton, 2008).

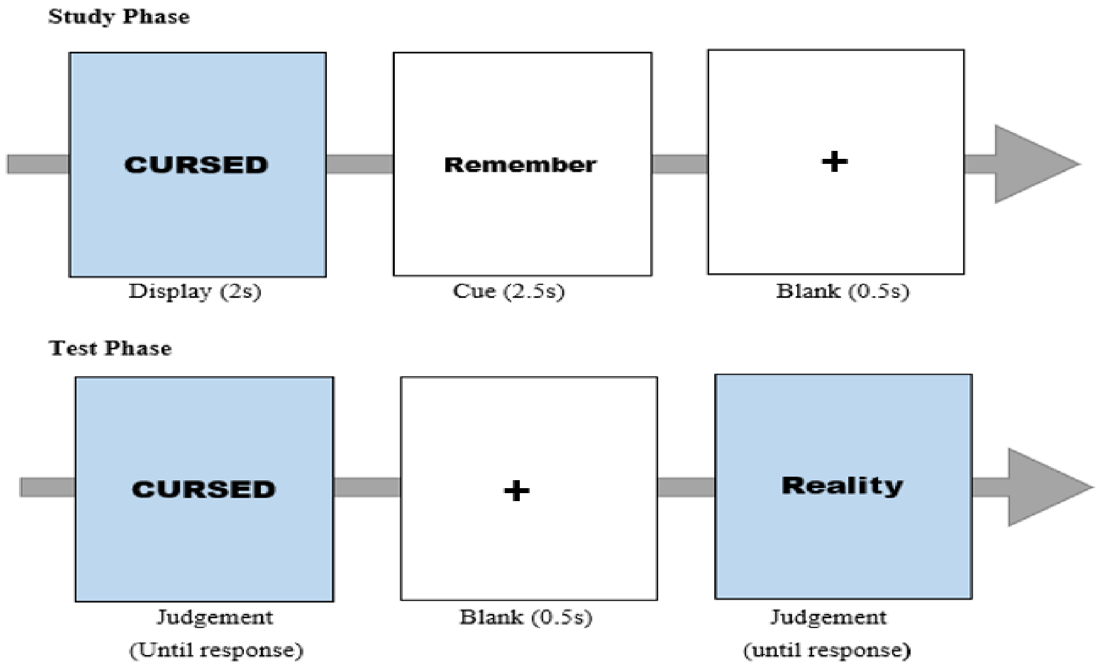

Figure 1 Trial Sequence for the Study and Test Phases of the Item-method Recognition Task

Note. Schematic illustration of the trial sequence for the Study Phase and Test Phase. In the Study Phase, each trial began with a centrally presented word for 2.0 s, followed by a cue ("Remember" or "Forget") for 2.5 s, and a fixation cross (+) for 0.5 s. In the Test Phase, each trial began with a probe word until response, followed by a fixation cross for 0.5 s, and then a source confidence judgment screen until response. Words were presented in 28-point font, black color, on a light blue background; cues and fixation crosses were presented on a white background.

Retention Interval

Next, participants were given a brief, 2-minute, neutral arithmetic task. This was done to prevent rehearsal of the previously presented words and to eliminate recency effects (MacLeod, 2012; Sahakyan & Kelley, 2002).

Test Phase Recognition

The test phase was composed of a total of 160 trials, including the previously presented words and distractors. The trials were presented in random order. For each recognition trial, a probe was presented, together with the six response options. The display was visible until a response was given. Then, a blank fixation screen was presented for 0.5 seconds, after which a new trial was presented. For each probe, participants had to select one of the six response options presented on the response screen. This format enabled us to obtain a general recognition decision as well as a specific source decision, including confidence.

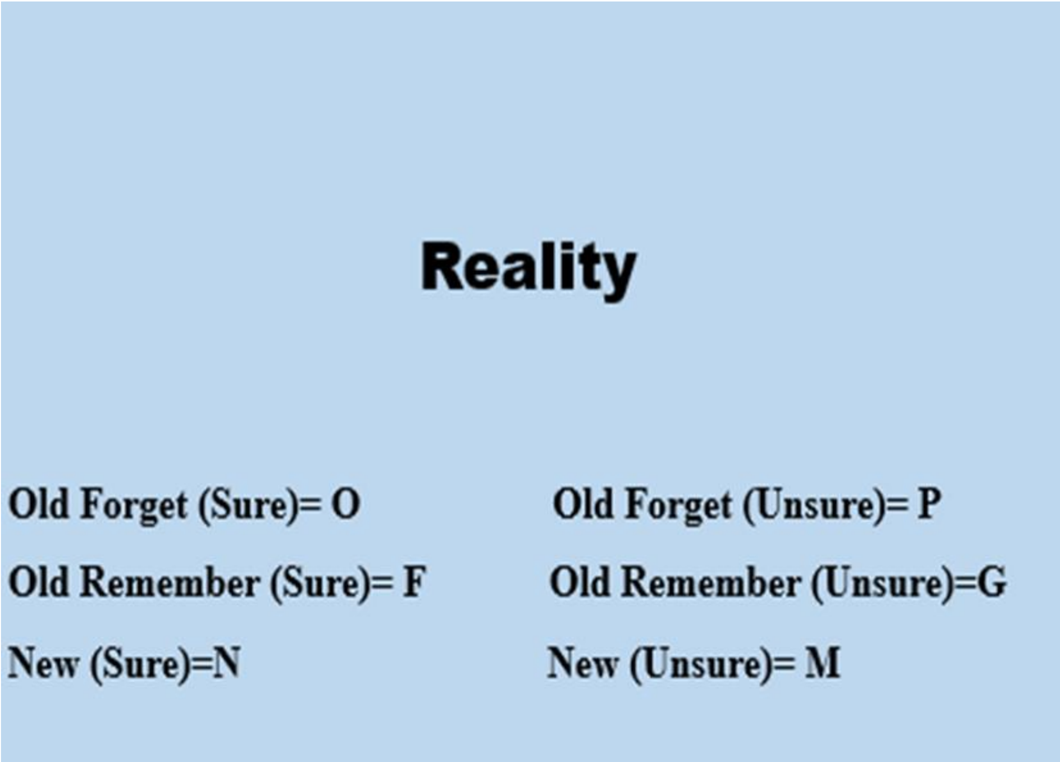

Figure 2 Example of the Options Screen Shown to the Participants During the Test Phase of Words

Note. Sample of the Source Confidence Response Screen shown to the participant during the Test Phase. Participants chose one of the six response options to express their recognition decision. The response options were Old Forget (Sure), Old Forget (Unsure), Old Remember (Sure), Old Remember (Unsure), New (Sure), or New (Unsure). The letters (O, P, F, G, N, M) are the response codes associated with the response options.

Outcome Variables and Scoring

The main dependent variable was the recognition of the studied items. For the main results, the responses were considered to be binary. For the studied words, any response that indicated that the item was previously seen—Old Remember and Old Forget, sure and unsure—was considered an old response. The studied items that were considered old were the hits, whereas the studied items that were considered new were the misses.

Cue-specific hit rates were defined by the item's original cue assignment during the study phase, not by the participant's later source judgment during recognition (Hourihan, 2021; Taylor et al., 2018). Thus, a Remember hit referred to a word originally paired with a Remember cue that was later endorsed as old. In contrast, a Forget hit referred to a word originally paired with a Forget cue that was later endorsed as old.For distractor words, any old response was coded as a false alarm, and any new response was coded as a correct rejection. These outcomes were retained for supplementary analyses of response bias and discrimination.

Design Features Supporting Internal Validity

Several design features were incorporated to enhance the study's internal validity. First, the cues for each word were counterbalanced over two word lists so that each word was presented an equal number of times in the Remember and Forget conditions (Basden, 1996; MacLeod, 1999). The randomization of trial order was also implemented for each participant to minimize the effects of primacy, recency, and other serial position effects (Murdock, 1962). Further, the same computerized instructions were provided for all participants, and the study was conducted under standard laboratory conditions. The use of the same distractor task between the study phase and the Test Phase was also implemented for all participants, minimizing the effects of immediate rehearsal and recency effects (Postman & Phillips, 1965). The recognition measure was also consistent for all participants, using the same response format and procedures.

Data Screening and Exclusion Criteria

Screening was carried out prior to analysis to detect incomplete participation and non-engagement with the task. The participants were excluded if they exhibited any of the following: more than 10% non-response during the recognition phase, an overall false alarm rate exceeding .60, or an overall hit rate below .40 across studied items (Rich, 2018). The total number of participants was 67. The data from 7 participants was excluded according to the criteria mentioned above. The final sample was composed of 60 participants.

Statistical Analysis

These analyses were carried out in two stages to mirror the structure of the available data. For the category-based analysis, aggregated participant-level data on participant recognitions were rearranged into long form such that each participant contributed one row per Cue x Category cell. The outcome variable was the number of hits out of the total number of items studied in each cell. This was analyzed using a binomial model with a logit link, with hit probability predicted from Cue (Remember, Forget), the four word categories (Neutral, Positive, Negative, Threatening), the Cue x Category interaction, and List as a nuisance variable, with standard errors clustered at the participant level (Dixon, 2008; Jaeger, 2008). The main effects of interest were Cue and the Cue x Category interaction.

For the analysis of the dimensions, the data were entered at the word by cue level since the predictor values of valence, arousal, and dominance were at the item level. Each word was entered once for the Remember condition and once for the Forget condition, and the outcome measure was the number of hits out of the total number of observations for the word in the given cue condition. Weighted binomial logistic regression analysis was performed with hit probability as the outcome variable and the predictor variables of cue, standardized valence, arousal, dominance, word length, and the interactions of these with cue (Jaeger, 2008; Warriner et al., 2013). The interaction of cue and arousal was the main test of the dimensions; the interaction of cue and valence and the interaction of cue and dominance were the secondary tests.

Results and Discussion

These analyses evaluated whether item-method directed forgetting (DF) was present, as well as examining the variability of DF across four-word categories that included affective/neutral content and continuous valence-arousal-dominance (VAD) properties. Descriptive analyses were conducted based on hit frequency, which is the number of items recognized as old. The main analyses were conducted using binomial logistic regression on hits and misses. Families of related post hoc comparisons were Holm-adjusted to control the familywise error rate.

Because List A and List B were used only to counterbalance cue assignment across items, the first step was to confirm that the two lists did not differ meaningfully in overall task performance. This analysis was necessary to ensure that any observed cue effects were not attributable to systematic differences between the counterbalancing lists. Independent-samples Welch's t tests were therefore conducted on total Remember hits, total Forget hits, and the DF gap.

Table 2
List A Versus List B Checks for Overall Performance (Hit Frequency)

	List A		List B						
Variable	**M**	**SD**	**M**	**SD**	**MD**	**95% CI**	**t(df)**	**P**	**d**
Remember hits (out of 40)	32.07	2.27	31.60	3.02	.47	[−.92, 1.85]	.68 (53.84)	.502	.17
Forget hits (out of 40)	21.63	3.72	20.97	4.12	.67	[−1.36, 2.70]	.66 (57.39)	.513	.17
DF gap (R − F)	10.43	4.01	10.63	3.52	−.20	[−2.15, 1.75]	−.21 (57.05)	.838	−.05

Note. MD = mean difference. Welch's t-tests are reported. Remember hits and forget hits are totals across categories (40 items per cue per participant). DF gap = Remember hits − Forget hits. d = standardised mean difference for independent samples.

As shown in Table 2, List A and List B did not differ significantly in total Remember hits, $t(53.84) = 0.68$, $p = .502$, $d = 0.17$, 95% CI [−0.92, 1.85], total Forget hits, $t(57.39) = 0.66$, $p = .513$, $d = 0.17$, 95% CI [−1.36, 2.70], or the DF gap, $t(57.05) = -0.21$, $p = .838$, $d = -0.05$, 95% CI [−2.15, 1.75]. These findings indicate that the two lists functioned as intended for counterbalancing and that treating List as a nuisance control factor in subsequent analyses is supported. In simple terms, participants assigned to List A and List B performed similarly, so later differences can be interpreted without concern that one list was easier or harder than the other.

The next question was whether the standard directed forgetting effect emerged across the full stimulus set and how this pattern looked descriptively across categories. To address this, an overall paired-samples t-test compared Remember and Forget hits across all studied items, and descriptive means were examined separately for each category.

Table 3
Mean Hit Frequency by Cue and Category

	Remember Hits		**Forget Hits**		**DF gap**	
Category	M	SD	M	SD	M	SD
Neutral	7.87	1.28	3.22	1.64	4.65	1.90
Positive	8.08	1.25	5.13	1.65	2.95	1.96
Negative	7.80	1.33	5.52	1.30	2.28	1.96
Threatening	8.08	1.20	7.43	1.41	0.65	1.64

Note. Hit frequency is the number of studied items that are endorsed as old. Means are based on 10 items per category and cue condition. DF gap = Remember hits − Forget hits.

Across the full stimulus set, Remember-cued items produced more hits than Forget-cued items, confirming a robust DF effect, $t(59) = 21.82$, $p < .001$, mean difference = 10.53 hits, 95% CI [9.57, 11.50], $d_z = 2.82$. Descriptively, the mean DF gap was largest for Neutral words ($M = 4.65$, $SD = 1.90$) and smallest for Threatening words ($M = 0.65$, $SD = 1.64$), with Positive ($M = 2.95$, $SD = 1.96$) and Negative ($M = 2.28$, $SD = 1.96$) words falling in between (see Table 3). In simple terms, participants remembered Remember-cued words much better than Forget-cued words overall, but this gap was much smaller for Threatening words than for the other categories.

To determine whether the DF effect was statistically reliable within each category, paired-samples t tests were conducted separately for Neutral, Positive, Negative, and Threatening words. These comparisons were Holm-adjusted across the four tests to control the familywise error rate.

Table 4
Within-Category Directed Forgetting in Hit Frequency

	Remember Hits		Forget Hits					
Category	**M**	**SD**	**M**	**SD**	**t(59)**	**dz**	**p (Holm)**	**Interpretation**
Neutral	7.87	1.28	3.22	1.64	18.93	2.44	< .001	Largest DF
Positive	8.08	1.25	5.13	1.65	11.65	1.50	< .001	Strong DF
Negative	7.80	1.33	5.52	1.30	9.03	1.17	< .001	Moderate DF
Threatening	8.08	1.20	7.43	1.41	3.06	0.40	.003	Weakest DF

Note. dz = Cohen's dz for paired samples. p-values were Holm-adjusted across the four within-category comparisons.

As shown in Table 4, significant DF effects were observed in all four categories after Holm adjustment: Neutral, $t(59) = 18.93$, $p < .001$, dz = 2.44; Positive, $t(59) = 11.65$, $p < .001$, dz = 1.50; Negative, $t(59) = 9.03$, $p < .001$, dz = 1.17; and Threatening, $t(59) = 3.06$, $p = .003$, dz = 0.40. Although DF was present in every category, the effect was markedly smaller for threatening words than for Neutral, Positive, or Negative words. In simple terms, participants were still able to forget threatening words to some degree, but those words were much less affected by the Forget cue than the others.

These results collectively suggest that the presence of DF was observed for all four categories, but the forgetting effect was much diminished for the Threatening words. This was due to the heightened recognition for the Forget-cued Threatening words compared to the other three categories.

The next issue was to determine whether the affective category moderated the extent of directed forgetting after controlling for the entire design involving Cue Category interaction. To examine this issue, a binomial logistic regression was estimated on the hits and misses for the entire design involving Cue Category interaction along with List as a predictor variable. This model was chosen to analyze the data because the dependent variable was binary in nature for the individual items (hits vs. misses), whereas the inferential interest was focused on the Cue Category interaction.

Table 5
Model 1 Fixed Effects Predicting Hit Probability from Cue × Category and List

Predictor	b	SE	Z	p	OR	95% CI for OR
Cue (Forget vs Remember) for Neutral	−2.05	.13	−16.15	< .001	.13	[.10, .16]
Positive vs. Neutral	.134	.150	.89	.372	1.14	[.85, 1.54]
Negative vs. Neutral	−.04	.14	−.28	.783	.96	[.73, 1.27]
Threatening vs. Neutral	.13	.13	1.03	.303	1.14	[.89, 1.48]
Cue × Positive	.66	.19	3.44	< .001	1.94	[1.33, 2.84]
Cue × Negative	.99	.19	5.22	< .001	2.70	[1.86, 3.92]
Cue × Threatening	1.68	.16	10.43	< .001	5.34	[3.90, 7.32]
List (B vs A)	−.07	.09	−.79	.428	.93	[.78, 1.11]

Note. b = log-odds coefficient; OR = odds ratio. Robust standard errors were clustered at the participant level. Cue is coded as "Forget" relative to "Remember".

Table 5 shows that Forget-cued items had substantially lower odds of recognition than Remember-cued items in the Neutral reference category, b = −2.05, SE = .13, z = −16.15, p < .001, OR = 0.13, 95% CI [0.10, 0.16]. There were no significant main effects of Positive, Negative, or Threatening categories relative to Neutral under the Remember reference condition, all ps > .30. Critically, however, the Cue × Positive interaction, b = 0.66, SE = .19, z = 3.44, p < .001, OR = 1.94, 95% CI [1.33, 2.84], the Cue × Negative interaction, b = 0.99, SE = .19, z = 5.22, p < .001, OR = 2.70, 95% CI [1.86, 3.92], and the Cue × Threatening interaction, b = 1.68, SE = .16, z = 10.43, p < .001, OR = 5.34, 95% CI [3.90, 7.32], were all significant and positive. List did not significantly predict recognition, b = −0.07, SE = .09, z = −0.79, p = .428, OR = 0.93, 95% CI [0.78, 1.11]. These results indicate that the Forget penalty was attenuated for Positive, Negative, and especially threatening words relative to Neutral words. In simple terms, the category-based model confirms that Threatening words were the hardest to forget intentionally.

A most theoretically important Primary binomial logistic regression tested whether continuous item-level VAD properties moderated the effect of DF. Because valence, arousal, and dominance were item-level predictors, a weighted binomial logistic regression was estimated using aggregated hit frequencies by word and cue condition. Valence, arousal, dominance, and word length were standardized prior to analysis. The central inferential target was the Cue × Arousal interaction.

Table 6
Model 2 Fixed Effects Predicting Hit Probability from Cue × VAD and Word Length

Predictor	B	SE	Z	p	OR	95% CI for OR
Cue (Forget vs Remember)	−1.22	.07	−18.27	< .001	0.30	[0.26, 0.33]
Valence (z)	−.34	.37	−.91	.362	0.71	[0.34, 1.48]
Arousal (z)	.20	.10	2.01	.045	1.22	[1.01, 1.48]
Dominance (z)	.50	.42	1.19	.232	1.65	[0.72, 3.75]
Cue × Valence (z)	.11	.49	.22	.828	1.11	[0.42, 2.91]
Cue × Arousal (z)	.61	.13	4.67	< .001	1.84	[1.42, 2.37]
Cue × Dominance (z)	−.08	.55	−.15	.881	0.92	[0.31, 2.72]
Word length (z)	−.02	.03	−.47	.641	0.98	[0.92, 1.05]

Note. Cue × Arousal was the central theoretical test. Positive coefficients indicate attenuation of the Forget penalty as the predictor increases.

As shown in Table 6, the model again showed a strong overall DF effect, with Forget-cued items having lower odds of recognition than Remember-cued items (b = −1.22, SE = .07, z = −18.27, p < .001, OR = 0.30, 95% CI [0.26, 0.33]). Arousal was positively associated with recognition, b = 0.20, SE = .10, z = 2.01, p = .045, OR = 1.22, 95% CI [1.01, 1.48]. Most importantly, the Cue × Arousal interaction was significant, b = 0.61, SE = .13, z = 4.67, p < .001, OR = 1.84, 95% CI [1.42, 2.37], indicating that the Forget-related reduction in recognition became smaller as arousal increased. In contrast, neither Cue × Valence, b = 0.11, SE = .49, z = 0.22, p = .828, OR = 1.11, 95% CI [0.42, 2.91], nor Cue × Dominance, b = −0.08, SE = .55, z = −0.15, p = .881, OR = 0.92, 95% CI [0.31, 2.72], was significant. Word length also did not predict recognition, b = −0.02, SE = .03, z = −0.47, p = .641, OR = 0.98, 95% CI [0.92, 1.05]. In simple terms, higher-arousal words were less affected by Forget cues, whereas valence and dominance did not show the same moderating effect.

To clarify the significant Cue × Arousal interaction, simple slopes of arousal were estimated separately for Remember and Forget conditions. This analysis was conducted to determine whether arousal predicted recognition similarly across both cue conditions or more strongly in one than the other.

Table 7
Simple Slopes of Arousal Within Each Cue Condition

Cue condition	B	SE	z	p	OR [95% CI]
Remember	0.197	0.098	2.01	.045	1.218 [1.005, 1.477]
Forget	0.806	0.086	9.36	< .001	2.239 [1.891, 2.650]

Note. Alternatively, values represent the multiplicative change in the odds of a hit per 1 SD increase in arousal.

As shown in Table 7, arousal was positively associated with recognition in the Remember condition, b = 0.197, SE = 0.098, z = 2.01, p = .045, OR = 1.218, 95% CI [1.005, 1.477], and even more strongly in the Forget condition, b = 0.806, SE = 0.086, z = 9.36, p < .001, OR = 2.239, 95% CI [1.891, 2.650]. Thus, although arousal improved recognition in both cue conditions, its association with recognition was substantially stronger for Forget-cued words. In simple terms, arousal especially helped words survive the Forget instruction.

Finally, exploratory three-way interactions were examined to determine whether VAD-based moderation differed across the four word categories. These analyses were treated cautiously because several estimates were imprecise and because they were not the primary inferential focus.

None of the Cue × Category × Arousal, Cue × Category × Valence, or Cue × Category × Dominance interactions reached significance. Taken together, the results support three conclusions. First, item-method directed forgetting was found to be robust. Second, DF was found to be reliably variable across categories, with threatening words having the smallest Remember-Forget difference and hence the highest resistance to forgetting. Third, the analyses of the dimensions showed that the attenuation was most associated with arousal rather than valence or dominance. For this data set, both the category-based and the dimensional analyses converged to a similar pattern: high-arousing materials showed resistance to directed forgetting.

Discussion

The present study investigated whether item-method Directed Forgetting (DF) of affective and neutral English words depends on affective category and continuous levels of valence, arousal, and dominance (VAD) properties. Three findings are central. First, the paradigm produced the expected overall DF effect: Remember-cued words were recognized more often than Forget-cued words (MacLeod, 1999). Second, DF magnitude differed across categories, with the largest Remember-Forget difference for Neutral words and the smallest for Threatening words. Third, arousal emerged as the clearest dimensional moderator of the cue effect: as arousal increased, the Forget penalty decreased, whereas comparable moderation did not emerge for valence or dominance. Taken together, these findings indicate that intentional forgetting is robust but not uniform; its behavioral impact appears reduced when the to-be-forgotten material is relatively high in arousal.

The overall DF effect is consistent with standard item-method findings and establishes the baseline needed to interpret the effects of affective moderation. At a general level, item-method accounts propose that Remember cues promote continued encoding or rehearsal, whereas Forget cues reduce subsequent processing of the just-presented item (Bjork, 1972; Wetzel & Hunt, 1977). The present data are consistent with that broad framework because cue instructions reliably altered later recognition. At the same time, the current design does not directly adjudicate among selective-rehearsal, inhibitory, or hybrid accounts of DF (Basden, 1996; MacLeod, 2012; Taylor et al., 2018). The strongest conclusion at this level is therefore modest: the cue manipulation worked in the expected direction. It provided a valid basis for testing whether its effect depended on the affective properties of the stimuli (Hall et al., 2021).

Although DF was robust overall, it was not equally strong across categories. Neutral words showed the strongest DF effect, whereas threatening words showed the weakest, with Positive and Negative words falling between these endpoints. This pattern indicates that affectively salient material was less susceptible to intentional forgetting than neutral material under the present task conditions (Hall et al., 2021; Wessel & Merckelbach, 2006). Importantly, however, DF remained statistically significant in all four categories. Thus, the present findings do not suggest that threatening verbal material is immune to memory control. Rather, they indicate that forget instructions remained effective, but less effective, when applied to relatively salient words. This distinction is important because the results do support attenuation rather than absence of control.

The category ordering is also an important aspect that needs emphasis. The results did not simply fall into the two categories of Neutral vs. Threatening. The results for the Positive and Negative words also showed an attenuation of the DF effect relative to the Neutral words, implying that the critical contrast is not simply threat, but rather salience in general. However, the particularly low DF effect for the threatening words was due mainly to the high recognition for the Forget-cued Threatening items. Although the category results are descriptively important, these kinds of results cannot pinpoint the critical property. In the present set of words, the threatening words not only had low valence but also had higher arousal and lower dominance (Mujitaba et al., 2025). Accordingly, the category findings should not be treated as evidence of a uniquely threat-specific mechanism. Instead, they identify where DF attenuation was most pronounced, leaving the dimensional analyses to clarify which affective property most plausibly accounted for that pattern.

Those dimensional analyses provided the strongest evidence for the study's main theoretical claim. Arousal significantly moderated the cue effect such that the reduction in recognition associated with a Forget cue became smaller as arousal increased. By contrast, the Cue × Valence and Cue × Dominance interactions were not statistically reliable in the present data. On that basis, arousal appears to be the clearest affective correlate of attenuated DF among the predictors tested here. This conclusion should still be stated carefully: the data show that arousal had the strongest evidence of moderation in this dataset, not that valence and dominance are irrelevant in principle. Nonetheless, the present findings suggest that reduced forgetting of affective verbal material may be better understood as a function of arousal-linked salience rather than of broad category membership alone. This explanation is consistent with the idea that the memory effects for emotional words may be influenced by separate factors of valence and arousal, rather than a general "emotionality" factor (Gao et al., 2024).

The simple slopes results provide further support for this explanation. For both cue conditions, arousal was positively related to recognition, although the relation was stronger for the Forget-cued words. This indicates that the relationship between arousal and memory may not have been due simply to arousal being related to better memory in general, but rather to the fact that arousal was related to the reduction in memory from the Forget cue. However, the exact processes by which this reduction in memory from the Forget cue occurred are not immediately clear from the present design (Basden, 1996;

MacLeod, 2012; Taylor et al., 2018). Indeed, the exact processes by which the moderation effects in a behavioral design like the present one occur are not immediately clear.

The current data generally accord with arousal-biased competition theory, which posits that arousal increases processing for high-priority representations (Mather & Sutherland, 2011). Translated to the item-method DF, this account predicts that forget instructions may counteract strategic rehearsal but may not entirely remove the memorial advantage of relatively arousing items. The current data follow this general pattern: High-arousal words showed reduced effects of Forget cues than low-arousal words. However, the evidence should be treated as supportive rather than decisive, because the present study assessed recognition outcomes rather than directly assessing attentional competition or post-cue control. More broadly, the findings align with the emotional DF literature, which shows that emotional material is often harder to forget than neutral material, and they also help explain why that literature is heterogeneous. Hall et al. (2021) not only found that, overall, emotional items are less susceptible to item-method DF effects than neutral items, but they also found that these effects differ substantially across studies and that these differences are greater when the emotional items are more arousing than the neutral items. The present results are in line with the overall pattern of these results and may help explain why arousal is a source of these differences.

Another strength of the present research is the combination of category and dimensional analyses. Category effects were useful for establishing where the attenuation of the DF effects was most marked, while the dimensional analysis was useful for establishing which affective property was most responsible for the attenuation. Category effects are more intuitively accessible, but they are mechanistically unclear. In the present case, the convergence between attenuated DF for threatening words and the Cue × Arousal interaction suggests that the limits of intentional forgetting may be understood more precisely in dimensional rather than purely categorical terms. The study also contributes to affective lexical research by treating VAD values as item-level predictors rather than merely descriptive labels, which aligns well with the broader use of large-scale lexical norms, such as those provided by Warriner et al. (2013).

Conclusion

In conclusion, the present study shows that item-method directed forgetting is robust overall, but that its magnitude depends on the affective characteristics of the material being regulated. Neutral words showed the strongest DF effect, whereas threatening words showed the weakest. More importantly, arousal provided the clearest evidence of attenuating the Forget penalty, whereas comparable evidence did not emerge for valence or dominance. The most defensible interpretation is that intentional forgetting remains effective under item-method conditions, but becomes less effective when to-be-forgotten words are relatively high in arousal. In this way, the present results extend existing models of emotionally directed forgetting by highlighting that the constraints of this process might be thought of in dimensional, rather than purely categorical, terms (Hamann, 2012; Warriner et al., 2013).

Recommendations

We propose several recommendations for future research. First, future studies should test arousal-sensitivity model of item-method directed forgetting more directly, because arousal appears to be more than valence or dominance in reducing intentional forgetting. Second, directed forgetting should be examined using native-language or culturally grounded valid way. Third, future studies should use more systematically balanced and well normed words sets and consider increasing trial per category only when this can be done without fatigue and practice effects. Fourth, replication with larger and more diverse sample is needed to determine whether present pattern generalizes beyond non-clinical Pakistan adults.

References

American Psychological Association. (2017). *Ethical principles of psychologists and code of conduct*. https://www.apa.org/ethics/code/ethics-code-2017.pdf

Anderson, M. C., & Hanslmayr, S. (2014). Neural mechanisms of motivated forgetting. *Trends in Cognitive Sciences, 18*(6), 279–292. https://doi.org/10.1016/j.tics.2014.03.002

Bailey, K., & Chapman, P. (2012). When can we choose to forget? An ERP study into item-method directed forgetting of emotional words. *Brain and Cognition, 78*(2), 133–147. https://doi.org/10.1016/j.bandc.2011.11.004

Basden, B. H., & Basden, D. R. (1996). Directed forgetting: Further comparisons of the item and list methods. *Memory, 4*(6), 633–654. https://doi.org/10.1080/741941000

Bjork, R. A. (1972). Theoretical implications of directed forgetting. In A. W. Melton & E. Martin (Eds.), *Coding processes in human memory* (pp. 217–235). Winston.

Bradley, M. M., & Lang, P. J. (1999). *Affective norms for English words (ANEW): Instruction manual and affective ratings* (Technical Report C-1). The Center for Research in Psychophysiology, University of Florida.

Brysbaert, M., Stevens, M., Mandera, P., & Keuleers, E. (2016). How many words do we know? Practical estimates of vocabulary size dependent on word definition, the degree of language input and the participant's age. *Frontiers in Psychology, 7*, Article 1116. https://doi.org/10.3389/fpsyg.2016.01116

Brysbaert, M., Warriner, A. B., & Kuperman, V. (2014). Concreteness ratings for 40 thousand generally known English word lemmas. *Behavior Research Methods, 46*(3), 904–911. https://doi.org/10.3758/s13428-013-0403-5

Chen, Z., Wang, L., Ying, S., Yuan, J., Ren, J., Yan, Y., Qin, Y., Liu, T., & Yao, D. (2025). Emotional influences on remembering and forgetting explained by frontal and parietal dynamics. *Journal of Neurophysiology, 133*(3), 784–798. https://doi.org/10.1152/jn.00484.2024

Dixon, P. (2008). Models of accuracy in repeated-measures designs. *Journal of Memory and Language, 59*(4), 447–456. https://doi.org/10.1016/j.jml.2007.11.004

Gao, C., Ren, J., Sakaki, M., & Jia, X. (2024). Memory enhancement for emotional words is attributed to both valence and arousal. *Acta Psychologica, 246*, Article 104249. https://doi.org/10.1016/j.actpsy.2024.104249

Hall, K. J., Fawcett, E. J., Hourihan, K. L., & Fawcett, J. M. (2021). Emotional memories are (usually) harder to forget: A meta-analysis of the item-method directed forgetting literature. *Psychonomic Bulletin & Review, 28*, 1313–1326. https://doi.org/10.3758/s13423-021-01914-z

Hamann, S. (2012). Mapping discrete and dimensional emotions onto the brain: Controversies and consensus. *Trends in Cognitive Sciences, 16*(9), 458–466. https://doi.org/10.1016/j.tics.2012.07.006

Hennessee, J. P., Patterson, T. K., Castel, A. D., & Knowlton, B. J. (2019). Forget me not: Encoding processes in value-directed remembering. *Journal of Memory and Language, 106*, 29–39. https://doi.org/10.1016/j.jml.2019.02.001

Hourihan, K. L. (2021). The influence of cue probability on item and source judgments in item method directed forgetting. *Memory, 29*(9), 1136–1155. https://doi.org/10.1080/09658211.2021.1967400

Jaeger, T. F. (2008). Categorical data analysis: Away from ANOVAs (transformation or not) and towards logit mixed models. *Journal of Memory and Language, 59*(4), 434–446. https://doi.org/10.1016/j.jml.2007.11.007

Kuperman, V., Stadthagen-Gonzalez, H., & Brysbaert, M. (2012). Age-of-acquisition ratings for 30,000 English words. *Behavior Research Methods, 44*(4), 978–990. https://doi.org/10.3758/s13428-012-0210-4

Lee, Y.-S. (2013). Costs and benefits in item-method directed forgetting: Differential effects of encoding and retrieval. *The Journal of General Psychology, 140*(3), 159–173. https://doi.org/10.1080/00221309.2012.750591

MacLeod, C. M. (1999). The item and list methods of directed forgetting: Test differences and the role of demand characteristics. *Psychonomic Bulletin & Review, 6*(1), 123–129. https://doi.org/10.3758/BF03210819

MacLeod, C. M. (2012). Directed forgetting. In N. M. Seel (Ed.), *Encyclopedia of the sciences of learning* (pp. 993–995). Springer. https://doi.org/10.1007/978-1-4419-1428-6_1084

Mather, M. (2007). Emotional arousal and memory binding: An object-based framework. *Perspectives on Psychological Science, 2*(1), 33–52. https://doi.org/10.1111/j.1745-6916.2007.00028.x

Mather, M., & Sutherland, M. R. (2011). Arousal-biased competition in perception and memory. *Perspectives on Psychological Science, 6*(2), 114–133. https://doi.org/10.1177/1745691611400234

Mathôt, S. (2025). *OSWeb extension for OpenSesame* (Version 2.2.9.0) [Computer software]. PyPI. https://pypi.org/project/opensesame-extension-osweb/

Mathôt, S., Schreij, D., & Theeuwes, J. (2012). OpenSesame: An open-source, graphical experiment builder for the social sciences. *Behavior Research Methods, 44*(2), 314–324. https://doi.org/10.3758/s13428-011-0168-7

Minnema, M. T., & Knowlton, B. J. (2008). Directed forgetting of emotional words. *Emotion, 8*(5), 643–652. https://doi.org/10.1037/a0013441

Mujitaba, A., & Riaz, S. (2025). Words that sting, words that soothe: A two-phase experimental and norming study on emotionally valenced word lists for social anxiety in Pakistan. *Qlantic Journal of Social Sciences, 6*(3), 103–117. https://doi.org/10.55737/qjss.vi-iii.25391

Mujitaba, A., Shaheen, A., Ahmad, S., Naeem, H., & Ihsan, A. (2025). Mapping culturally grounded English word affect in Pakistan: A valence-arousal-dominance lexicon.

Social Sciences & Humanity Research Review, 2(4), 1724–1736. https://doi.org/10.63468/sshrr.348

Murdock, B. B., Jr. (1962). The serial position effect of free recall. *Journal of Experimental Psychology, 64*(5), 482–488. https://doi.org/10.1037/h0045106

Nadeem, A., Mujitaba, A., Ahmad, S., & Ihsan, A. (2026). Surah Ar-Rahman recitation and affective-cognitive outcomes in young adults: An experimental study of anxiety, mood, and selective memory consolidation. *The Critical Review of Social Sciences Studies, 4*(1), 2013–2030. https://doi.org/10.59075/zm9x8085

Nowicka, A., Marchewka, A., Jednoróg, K., Tacikowski, P., & Brechmann, A. (2011). Forgetting of emotional information is hard: An fMRI study of directed forgetting. *Cerebral Cortex, 21*(3), 539–549. https://doi.org/10.1093/cercor/bhq117

Pastötter, B., Kliegl, O., & Bäuml, K.-H. T. (2016). List-method directed forgetting: Evidence for the reset-of-encoding hypothesis employing item-recognition testing. *Memory, 24*(1), 63–74. https://doi.org/10.1080/09658211.2014.985589

Postman, L., & Phillips, L. W. (1965). Short-term temporal changes in free recall. *Quarterly Journal of Experimental Psychology, 17*(2), 132–138. https://doi.org/10.1080/17470216508416422

Quinlan, C. K., Taylor, T. L., & Fawcett, J. M. (2010). Directed forgetting: Comparing pictures and words. *Canadian Journal of Experimental Psychology/Revue canadienne de psychologie expérimentale, 64*(1), 41–46. https://doi.org/10.1037/a0016569

Rich, J. B. (2018). Recognition memory test. In J. S. Kreutzer, J. DeLuca, & B. Caplan (Eds.), *Encyclopedia of clinical neuropsychology* (pp. 2952–2957). Springer. https://doi.org/10.1007/978-3-319-57111-9_1148

Sahakyan, L., & Kelley, C. M. (2002). A contextual change account of the directed forgetting effect. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 28*(6), 1064–1072. https://doi.org/10.1037/0278-7393.28.6.1064

Scotti, P. S., & Maxcey, A. M. (2022). Directed forgetting of pictures of everyday objects. *Journal of Vision, 22*(10), Article 8. https://doi.org/10.1167/jov.22.10.8

Taylor, T. L., Cutmore, L., & Pries, L. (2018). Item-method directed forgetting: Effects at retrieval? *Acta Psychologica, 183*, 116–123. https://doi.org/10.1016/j.actpsy.2017.12.004

Tan, P., Ensor, T. M., Hockley, W. E., Harrison, G. W., & Wilson, D. E. (2020). In support of selective rehearsal: Double-item presentation in item-method directed forgetting. *Psychonomic Bulletin & Review, 27*(3), 529–535. https://doi.org/10.3758/s13423-020-01723-w

Warriner, A. B., Kuperman, V., & Brysbaert, M. (2013). Norms of valence, arousal, and dominance for 13,915 English lemmas. *Behavior Research Methods, 45*(4), 1191–1207. https://doi.org/10.3758/s13428-012-0314-x

Wessel, I., & Merckelbach, H. (2006). Brief report: Forgetting "murder" is not harder than forgetting "circle": Listwise-directed forgetting of emotional words. *Cognition & Emotion, 20*(1), 129–137. https://doi.org/10.1080/02699930500260195

Wetzel, C. D., & Hunt, R. E. (1977). Cue delay and the role of rehearsal in directed forgetting. *Journal of Experimental Psychology: Human Learning and Memory, 3*(2), 233–245. https://doi.org/10.1037/0278-7393.3.2.233

Zwissler, B., Schindler, S., Fischer, H., Plewnia, C., & Kissler, J. M. (2015). "Forget me (not)?" Remembering forget-items versus un-cued items in directed forgetting. *Frontiers in Psychology, 6*, Article 1741. https://doi.org/10.3389/fpsyg.2015.01741