



Emotion processing in concrete and abstract words: evidence from eye fixations during reading

Bo Yao, Graham G. Scott, Gillian Bruce, Ewa Monteith-Hodge & Sara C. Sereno

To cite this article: Bo Yao, Graham G. Scott, Gillian Bruce, Ewa Monteith-Hodge & Sara C. Sereno (04 Jul 2024): Emotion processing in concrete and abstract words: evidence from eye fixations during reading, *Cognition and Emotion*, DOI: [10.1080/02699931.2024.2367062](https://doi.org/10.1080/02699931.2024.2367062)

To link to this article: <https://doi.org/10.1080/02699931.2024.2367062>



© 2024 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group



[View supplementary material](#)



Published online: 04 Jul 2024.



[Submit your article to this journal](#)



Article views: 49



[View related articles](#)



[View Crossmark data](#)

Emotion processing in concrete and abstract words: evidence from eye fixations during reading

Bo Yao ^a, Graham G. Scott ^b, Gillian Bruce ^b, Ewa Monteith-Hodge^c and Sara C. Sereno ^c

^aDepartment of Psychology, Lancaster University, Lancaster, UK; ^bSchool of Education and Social Sciences, University of the West of Scotland, Paisley, UK; ^cSchool of Psychology and Neuroscience, University of Glasgow, Glasgow, UK

ABSTRACT

We replicated and extended the findings of Yao et al. [(2018). Differential emotional processing in concrete and abstract words. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 44(7), 1064–1074] regarding the interaction of emotionality, concreteness, and imageability in word processing by measuring eye fixation times on target words during normal reading. A 3 (Emotion: negative, neutral, positive) × 2 (Concreteness: abstract, concrete) design was used with 22 items per condition, with each set of six target words matched across conditions in terms of word length and frequency. Abstract (e.g. *shocking, reserved, fabulous*) and concrete (e.g. *massacre, calendar, treasure*) target words appeared (separately) within contextually neutral, plausible sentences. Sixty-three participants each read all 132 experimental sentences while their eye movements were recorded. Analyses using Gamma generalised linear mixed models revealed significant effects of both Emotion and Concreteness on all fixation measures, indicating faster processing for emotional and concrete words. Additionally, there was a significant Emotion × Concreteness interaction which, critically, was modulated by Imageability in early fixation time measures. Emotion effects were significantly larger in higher-imageability abstract words than in lower-imageability ones, but remained unaffected by imageability in concrete words. These findings support the multimodal induction hypothesis and highlight the intricate interplay of these factors in the immediate stages of word processing during fluent reading.

ARTICLE HISTORY

Received 27 November 2023
Revised 3 June 2024
Accepted 5 June 2024

KEYWORDS


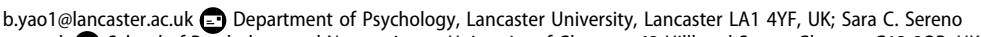

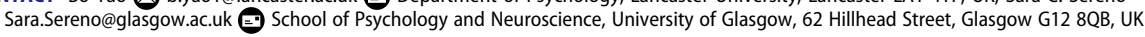

Emotion; concreteness; word recognition; eye fixations; reading

Introduction

While it is well-established that emotionally valenced (positive and negative) written words are recognised faster than neutral words, different explanations for this advantage have been proposed (Yao et al., 2018). Studies have varied word concreteness to understand the mechanisms of this facilitation (Kanske & Kotz, 2007; Palazova et al., 2013; Sheikh & Titone, 2013), but methodologies and results have been somewhat inconsistent. Recently, Yao et al. (2018) conducted a large-scale lexical decision

experiment manipulating target word Emotion and Concreteness, while controlling for other lexical variables to probe emotion word processing. The current study replicates Yao et al.'s study but, critically, measures eye fixation durations on target words presented in sentences, examining the early time course of these effects during fluent reading.

Emotionality is characterised by arousal (internal activation) and valence (value or worth) (Osgood et al., 1957). Emotion words, compared to neutral words, typically exhibit higher arousal and more

CONTACT Bo Yao  b.yao1@lancaster.ac.uk  Sara C. Sereno  Sara.Sereno@glasgow.ac.uk 
 Supplemental data for this article can be accessed online at <https://doi.org/10.1080/02699931.2024.2367062>.

© 2024 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group
This is an Open Access article distributed under the terms of the Creative Commons Attribution-NonCommercial-NoDerivatives License (<http://creativecommons.org/licenses/by-nc-nd/4.0/>), which permits non-commercial re-use, distribution, and reproduction in any medium, provided the original work is properly cited, and is not altered, transformed, or built upon in any way. The terms on which this article has been published allow the posting of the Accepted Manuscript in a repository by the author(s) or with their consent.

extreme valence (strongly negative or positive). However, emotional valence may manifest differently in concrete and abstract words. Concreteness concerns how much a word's meaning is grounded in sensory experience. Concrete words (e.g. *list*, *liquid*) evoke tangible, perceptible phenomena, while abstract words (e.g. *stay*, *subtle*) are less sensory-based. Concrete words tend to generate more vivid mental images and sensory associations than abstract words. This difference in sensory grounding is thought to influence how emotional content is processed and represented.

The interaction of emotional valence and concreteness is explained by two competing hypotheses (see Yao et al., 2018). The *representational substitution* hypothesis suggests that emotions are central in representing abstract words, giving them a residual processing advantage over concrete words, once differences in context availability and imageability are accounted for (Kousta et al., 2011; Vigliocco et al., 2014). This hypothesis predicts a more pronounced emotion advantage in processing abstract than concrete words, positing that emotional representations of abstract words “substitute” for their lack of direct sensorimotor experiences. In contrast, the *multimodal induction* hypothesis argues for a stronger emotional valence effect in concrete than abstract words, attributed to concrete words' direct sensorimotor associations facilitating their emotional processing more effectively. This model is grounded in the theory that sensorimotor experiences are required for evoking emotional responses, with higher levels of imageability acting as conduits for “inducing” emotional activation (Hess & Blairy, 2001; Niedenthal, 2007).

In their seminal work, Yao et al. (2018) conducted a large-scale lexical decision study exploring the interaction of emotionality and concreteness in word processing, examining the roles of alexithymia and word imageability within the frameworks of the representational substitution and multimodal induction hypotheses. Their study manipulated Emotion (Negative, Neutral, Positive) and Concreteness (Abstract, Concrete), while controlling for key lexical variables such as log frequency, familiarity, age of acquisition (AoA), and arousal.

Yao et al. (2018) found significant main effects of Emotion and Concreteness, indicating faster processing for emotional and concrete words over neutral and abstract words, respectively. Critically, they observed a significant Emotion \times Concreteness

interaction, consistent with the multimodal induction hypothesis, which posits a greater emotional valence advantage in concrete compared to abstract words. Moreover, the study revealed that this interaction was not influenced by participant alexithymia level but was significantly influenced by word imageability. This three-way interaction suggested that while both concrete and abstract emotion words are processed faster than neutral words, the mechanisms of this facilitation relied more on imageability, especially in abstract words. These results supported the idea that sensory associations in abstract words may enhance the activation of emotional content, highlighting the importance of sensory activation in emotional processing.

Although Yao et al. (2018) provided valuable insights on the interplay of emotionality, concreteness, and imageability in word processing, their lexical decision task – focusing on decontextualised words and prone to secondary task effects – may not reflect natural reading accurately (Diependaele et al., 2012). Furthermore, lexical decisions, a relatively late measure of word recognition, offer limited insight into the earliest interactions of Emotion, Concreteness, and Imageability, leaving a gap in understanding immediate cognitive responses to these word characteristics (Sereno & Rayner, 2003).

The current study addresses these limitations by testing the three-way interaction in naturalistic sentence reading. This approach is more representative of real-world language processing, overcoming the artificial and decontextualised nature of lexical decision (Rayner, 2009). Using eye-tracking, both earlier and later fixation measures, such as first fixation duration (FFD) and gaze duration (GD), can be captured. Indeed, recent eye-tracking investigations have illuminated the early influence of emotion on word processing (Knickerbocker et al., 2015, 2019), as well as its interaction with word frequency (Scott et al., 2012). It remains to be seen if the interactions between Emotion, Concreteness, and Imageability observed in Yao et al. (2018) affect the early stages or only the later stages of word processing in natural reading. Understanding the time course of these interactions is crucial for discerning whether sensorimotor activations through imageability are foundational to emotional activation *during* lexico-semantic processing, or emerge as a by-product or elaboration of semantic content *after* lexico-semantic processing.

Building on Yao et al. (2018), we anticipate a significant Emotion \times Concreteness interaction, with

concrete words showing a larger emotional valence advantage (faster processing of emotional vs. neutral words) than abstract words. Furthermore, we predict this interaction to be modulated by imageability, but effectively only in abstract words. These interactions are expected to emerge in early fixation measures like FFD. In line with the multimodal induction hypothesis, sensory activation (imageability) should enhance emotional processing, particularly in abstract words. As emotion effects have been demonstrated to emerge early in word recognition, we expect these interactions to manifest in early processing stages, where FFD captures rapid, pre-attentive responses to both sensory activation and emotional content. Our study seeks to elucidate how these three factors rapidly and automatically interact during the early stages of word processing in fluent reading.

Method

Participants

We used G*Power (Faul et al., 2009) for sample size estimation to replicate Yao et al.'s (2018) Emotion-Concreteness-Imageability interaction, with an observed Cohen's $d = 0.43$ from a lexical decision task. Due to the lack of established sample size estimation methods for generalised linear mixed models (GLMMs) and the original effect derived from lexical decisions rather than eye movements, we opted for

a simplified t -test approximation (two-tailed, $\alpha = .05$, power = .80), which indicated a minimum of 45 participants for replicating this three-way interaction.

Sixty-three members of the University of Glasgow community (42 female, 21 male; $M_{\text{age}} = 23$) took part in the experiment. Additional participants (none from the main experiment) were utilised in associated norming studies (see *Design and materials*). All were native English speakers, had normal or corrected-to-normal vision, and had never been diagnosed with any reading disorder. Participants were compensated at £6/hr or were given course credits. The study conformed to British Psychological Society ethical guidelines and procedures were approved by the College of Science and Engineering Ethics Committee at the University of Glasgow. Participants gave written informed consent prior to testing.

Design and materials

The experiment employed a 3 (Emotion: Negative, Neutral, Positive) \times 2 (Concreteness: Abstract, Concrete) design. We selected 132 words (22 in six conditions) to appear as target words in individual sentences. Most target words were nouns, with some being verbs or adjectives. Across the six conditions, words were matched on an item-by-item basis for word length and frequency of occurrence. Word frequencies were acquired from the British

Table 1. Word stimuli.

Negative	Abstract Neutral	Positive	Negative	Concrete Neutral	Positive
risk	stay	wish	fire	list	cash
fear	duty	trust	bomb	path	gold
danger	method	belief	battle	winter	castle
suffer	affect	create	murder	mirror	dinner
curse	mood	luck	storm	jury	silk
fierce	subtle	honest	weapon	liquid	flower
chaos	array	bonus	crash	canal	ocean
evil	mild	bold	riot	horn	toys
cruel	overt	brave	flood	cliff	movie
hostile	passive	sincere	robbery	stadium	rainbow
scared	casual	playful	poison	basket	balloon
agony	irony	glory	snake	arrow	bunny
shocking	reserved	fabulous	massacre	calendar	treasure
atrocious	abstract	exquisite	terrorist	mechanic	butterfly
irritating	conceptual	sensational	earthquake	technician	millionaire
punish	oblige	admire	bullet	sponge	jewels
deceive	convert	refresh	shotgun	luggage	bouquet
manic	tame	cosy	shark	mast	pearl
obscenity	deviation	gratitude	tarantula	container	limousine
insanity	deduction	devotion	assassin	wrestling	comedian
malice	satire	valour	dagger	crater	kitten
paranoid	thrifty	graceful	scorpion	herring	ladybird

National Corpus, a database of 90 million written word tokens (Davies, 2004). Emotion (valence) and concreteness values for targets were obtained from the Glasgow Norms (Scott et al., 2019). Target words are listed in Table 1 and their lexical specifications are presented in Table 2. The specifications of additional lexical variables that were used in our analyses are included in Table 2, namely arousal (calming or exciting), familiarity (unfamiliar or familiar), age of acquisition (AoA; estimated age at which a word was learned), and imageability (ease or difficulty to imagine or picture), all of which were Likert-based ratings from the Glasgow Norms.

Although each set of three Abstract and three Concrete words were length- and frequency-matched, separate sets of sentence frames were constructed. That is, for each set of three Abstract targets (Negative, Neutral, Positive), three alternative sentence frames that could accommodate each of the three targets (separately) were created. The same was

done for each set of three Concrete targets. As each target could appear in any of its three possible sentence frames, three versions of experimental materials were devised, presented to three participant groups. In this way, all participants read all target words in all sentence frames, but with alternative mappings between target words and sentence frames. Thus, each set of three target words appeared in identical contexts across participants. Moreover, the use of multiple sentence frames for multiple targets reduced the possibility of creating idiosyncratic contexts, increasing the generalisability of results. An example set of materials is presented in Table 3. The full set of materials is presented in Supplementary Materials A.

Our aim was for the target words to be both contextually neutral (not predictable) and yet plausible (not anomalous) in their sentences. In reading studies, when target words are either predictable or anomalous, their fixation times are correspondingly

Table 2. Means (with SDs) of target word specifications across conditions.

	Abstract			Concrete		
	Negative	Neutral	Positive	Negative	Neutral	Positive
<i>N</i>	22	22	22	22	22	22
Valence	2.35 (0.66)	5.21 (0.48)	7.46 (0.49)	2.55 (0.81)	5.08 (0.51)	6.91 (0.50)
Valence	3.65 (0.66)	1.44 (0.26)	3.46 (0.49)	3.45 (0.81)	1.41 (0.31)	2.91 (0.50)
Concreteness	2.84 (0.34)	2.70 (0.43)	2.73 (0.39)	6.01 (0.65)	6.07 (0.41)	6.24 (0.42)
Frequency	22.00 (33.31)	25.92 (32.81)	24.91 (34.43)	21.33 (31.61)	22.62 (30.53)	20.52 (26.62)
Log frequency	0.97 (0.57)	1.12 (0.54)	1.04 (0.61)	0.98 (0.59)	1.09 (0.47)	0.90 (0.81)
Length	6.32 (1.73)	6.18 (1.79)	6.36 (1.89)	6.32 (1.73)	6.18 (1.79)	6.36 (1.89)
Arousal	5.67 (0.66)	4.15 (0.68)	5.91 (0.71)	5.76 (0.71)	3.91 (0.70)	5.70 (0.68)
Familiarity	5.29 (0.56)	4.95 (0.85)	5.47 (0.78)	5.19 (0.54)	5.35 (0.84)	5.79 (0.53)
AoA	4.43 (1.13)	5.16 (1.08)	4.15 (1.18)	3.83 (0.96)	3.61 (1.02)	2.88 (1.00)
Imageability	3.67 (0.71)	2.71 (0.42)	3.46 (0.61)	6.14 (0.57)	6.13 (0.38)	6.41 (0.39)

Note. Units of measures are as follows: Valence on a scale from 1 (very negative) to 5 (neutral) to 9 (very positive); |Valence| = abs(valence-5) + 1, ranging from 1 (neutral) to 5 (highly valenced); Concreteness on a scale from 1 (very abstract) to 7 (very concrete); Frequency in occurrences per million; Length in number of letters; Arousal on a scale from 1 (very unarousing) to 9 (very arousing); Familiarity on a scale from 1 (very unfamiliar) to 7 (very familiar); AoA (age of acquisition) on a scale from 1 to 7 (ages 0–2, 3–4, 5–6, 7–8, 9–10, 11–12, and 13+ years, respectively); and Imageability on a scale from 1 (very unimageable) to 7 (very imageable).

Table 3. Example stimulus set.

	Condition	Targets	Sentence frames
Abstract	Negative	shocking	The crime novel detailed how ____ the affair had been.
	Neutral	reserved	Bethany's remarks were ____ and helped promote their cause.
	Positive	fabulous	The model's new look was ____ as she strode down the catwalk.
Concrete	Negative	massacre	It turned out that information on the ____ had been fabricated.
	Neutral	calendar	Scholars continue to study the Chinese ____ and its history.
	Positive	treasure	The origin of the ancient ____ is still a mystery.

Note. One set of the 22 sets of length- and frequency-matched Abstract and Concrete Negative, Neutral, and Positive target words is presented. Separate sets of sentence frames were constructed for Abstract and Concrete targets. The “____” designates where one of the target words would appear. Any of the Abstract sentence frames could accommodate any of the Abstract targets; any of the Concrete sentence frames could accommodate any of the Concrete targets. Three versions of the materials were designed, presented to three different participant groups. In this way, all participants read all targets and all sentence frames, only the mapping between target and sentence frame differed across the three participant groups.

faster or slower, respectively (Rayner et al., 2004; Sereno et al., 2018). Contextual predictability is typically approximated via a Cloze task and plausibility via a rating task. We ran both norming studies to validate our materials.

Predictability

Twenty-five native English participants (17 female, 8 male; $M_{\text{age}} = 23$) took part in the Cloze task. For each of the 132 experimental sentences, only the first part of each sentence up to but not including the target word was presented, and participants were asked to generate the next word. Items were scored as “1” for correct responses and “0” for all other guesses. With three possible targets for each of the 132 sentences, there were 396 possible correct guesses. The overwhelming majority of target words were never guessed. Of the few that were, eight words were guessed once (*battle, crash, flood, earthquake, ocean, butterfly, luck, playful*; Cloze = .01), one was guessed twice (*bomb*; Cloze = .03), one guessed three times (*duty*; Cloze = .04), and one guessed four times (*murder*; Cloze = .05). Sereno et al. (2018) had defined the following levels of predictability with respect to Cloze values: low (0.00–0.05); medium (0.20–0.75); and high (0.85–1.00). Thus, none of the targets in the current study were deemed predictable from their prior contexts.

Plausibility

For the plausibility task, the three versions of the materials were rated by 57 participants (38 female, 19 male; $M_{\text{age}} = 25$). An additional 30 filler sentences, designed to be lower in plausibility, were included. Participants were instructed to judge how likely, reasonable, or believable the events described in each sentence were, by rating each sentence on a scale of 1 (very low plausibility) to 4 (medium plausibility) to 7 (very high plausibility). For experimental sentences, the average plausibility across all six conditions ranged from 4.9 to 5.1; for filler sentences, the average plausibility was 2.7. Thus, the experimental sentences were deemed plausible.

Apparatus

Eye movements were monitored via an SR Research Desktop-Mount EyeLink 2K eyetracker (spatial resolution 0.01°), with participants’ heads stabilised via a chin/forehead rest. Viewing was binocular and eye position was sampled from the right eye at 1000 Hz.

EyeTrack software (<https://blogs.umass.edu/eyelab/software/>) was used to control text presentation. Text (black letters on a white background, 14-point nonproportional Bitstream Vera Sans Mono font) was presented on a Dell P1130 19” flat screen cathode ray tube (CRT; 1024 × 768 resolution; 150 Hz). All experimental sentences appeared on a single line (maximum 70 characters). At a viewing distance of 72 cm, approximately four characters subtended 1° of visual angle.

Procedure

Participants were informed they would be reading sentences on the computer screen while their eyes were monitored. They were told that although Yes-No comprehension questions followed two-thirds of the sentences, these were there to keep them engaged and they should read normally as if they were reading a magazine article. After the initial calibration and validation procedure (9-point, full horizontal and vertical range), participants read six practice sentences (with four questions). They then read the 132 experimental sentences (with 88 associated questions). Participants self-paced their breaks, and calibration and validation procedures were repeated after each break and as necessary throughout the session.

Each trial began with a centrally-displayed dot. When the participant’s eye position was verified, the dot disappeared and a black box appeared to the left, at the location of the first letter of the sentence. When their eye position was again detected, the box disappeared and a sentence was presented. When the participant finished reading the sentence, they looked to the bottom right of the screen and pressed the left or right trigger on the game controller to clear the screen. On one-third of the trials, the next trial began. On the remaining two-thirds of the trials, a question first appeared with the words “No” and “Yes” below, and participants pressed the left or right trigger to indicate their response. On average, participants answered 96% of the questions correctly.

Results

To prepare the eye-tracking data for statistical analyses, a suite of data management programmes (<https://blogs.umass.edu/eyelab/software/>) were used. The target region comprised the target word and the space before it. Lower and upper cutoff

values for individual fixations were 100 and 750 ms, respectively. Data were additionally eliminated if there was a blink or track loss on the target, or if a first-pass fixation on the target was either the first or last fixation on that line. Overall, 1% of the data were excluded for these reasons. The percentages of the remaining data for first-pass single fixation, immediate refixation, and first-pass skipping of the target were 72, 17, and 11%, respectively.

The resulting data were analysed over a number of standard measures: first fixation duration (FFD; the initial first-pass fixation duration, regardless of whether the target was refixated); single fixation duration (SFD; first-pass fixation time when a target was only fixated once); and gaze duration (GD; the sum of all first-pass fixations before the eyes move to another word).

Model choice and software

Although Yao et al. (2018) employed Linear Mixed Models (LMMs), we opted for Gamma Generalised Linear Mixed Models (GLMMs) with identity links to better model positively skewed fixation duration data. This choice was supported by our empirical comparisons, which consistently showed that Gamma GLMMs had better model fit, as indicated by lower Bayesian Information Criterion values compared to LMMs (see [Supplementary Materials B](#)). Thus, switching to Gamma GLMMs in our analysis was justified by both theoretical considerations and empirical evidence of their suitability for our data.

All GLMMs were fitted in R (<https://www.r-project.org/>) using the `glmer()` function from the `lme4` package (Bates et al., 2015). The predictors' p -values were calculated using Satterthwaite's approximations via the `lmerTest` package (Kuznetsova et al., 2017). Variance Inflation Factors (VIFs) for model predictors were calculated using the `vif()` function from the `car` package (Fox & Weisberg, 2019). Figures were generated using `ggplot2` (Wickham, 2016).

Replicating Emotion \times Concreteness

We used three GLMMs with identity links to examine the effects of Emotion (absolute valence) and Concreteness, as per Yao et al. (2018). Our models included log frequency, familiarity, AoA, and arousal as fixed effects for word-level variation, and maximised the random-effect structure with by-subject random intercepts and slopes. The coefficients for FFD, SFD,

Table 4. GLMM coefficients for fixed effects of Emotion, Concreteness, and their interactions by fixation duration measures.

	FFD	SFD	GD
Intercept	*** 222.5	*** 225.7	*** 262.7
Emo	* -4.6	* -5.4	*** -10.4
CNC	*** -8.9	*** -9.3	** -10.2
Emo \times CNC	** -8.8	-8.5	*** -13.6
Covar: Log frequency	*** -5.4	*** -7.7	*** -25.5
Covar: Familiarity	*** -8.0	*** -7.5	0.0
Covar: AoA	-1.1	-1.9	* 5.0
Covar: Arousal	** 6.4	** 7.3	*** 30.5

Note: Emo = Emotion (absolute valence); CNC = Concreteness; Covar = Covariate; AoA = age of acquisition; FFD = first fixation duration; SFD = single fixation duration; GD = gaze duration. Significant effects are highlighted in **bold**, with significance levels indicated by asterisks: *** $p < .001$, ** $p < .01$, * $p < .05$; marginal effects are highlighted in *italics*.

and GD are reported in [Table 4](#), with the full results reported in [Supplementary Materials B](#). VIFs for critical predictors were below 1.46 across all GLMMs, indicating low multicollinearity.

As predicted, both Emotion and Concreteness significantly affected all fixation measures. Emotional words (positive and negative) were processed faster than neutral words, and concrete words faster than abstract ones. The interaction of Emotion and Concreteness was significant in FFD and GD, and marginally so in SFD. Exploring the interactions, we found that the effects of Emotion were significantly larger in Concrete ($M + SD$) words ($b_{\text{FFD}} = -9$ ms, $95\%CI_{\text{FFD}} [-14, -4]$; $b_{\text{SFD}} = -10$ ms, $95\%CI_{\text{SFD}} [-17, -2]$; $b_{\text{GD}} = -17$ ms, $95\%CI_{\text{GD}} [-25, -10]$) than in Abstract ($M - SD$) words ($b_{\text{FFD}} = 0$ ms, $95\%CI_{\text{FFD}} [-6, 5]$; $b_{\text{SFD}} = -1$ ms, $95\%CI_{\text{SFD}} [-8, 5]$; $b_{\text{GD}} = -4$ ms, $95\%CI_{\text{GD}} [-11, 4]$), consistent with Yao et al.'s (2018) findings.

Replicating Emotion \times Concreteness \times Imageability

We used three Gamma GLMMs with identity links to examine the effects of Emotion (absolute valence), Concreteness, and Imageability, and their interactions on fixation durations (except for Concreteness and Emotion \times Concreteness, as in Yao et al., 2018, to address multicollinearity). Our fixed-effect structure included log frequency, familiarity, AoA, and arousal for word-level variation. By-subject variation was addressed by maximising the random-effect structure with random intercepts and slopes for all fixed effects. We report the fixed-effect coefficients across FFD, SFD, and GD in [Table 5](#), and report the full results in [Supplementary Materials B](#). VIFs for all predictors

Table 5. GLMM coefficients for fixed effects of Emotion, Concreteness, Imageability, and their interactions by fixation duration measures.

	FFD	SFD	GD
Intercept	*** 226.2	*** 230.9	*** 257.7
Emo	*** -9.6	*** -10.9	** -10.0
IMAG	*** -11.0	*** -11.4	*** -13.8
Emo × IMAG	** -8.7	* -7.6	*** -19.2
CNC × IMAG	*** -12.8	*** -18.9	*** 24.3
Emo × CNC × IMAG	*** 16.8	*** 16.6	7.8
Covar: Log frequency	*** -6.0	*** -8.5	*** -25.0
Covar: Familiarity	*** -9.1	*** -8.8	0.1
Covar: AoA	* 4.2	* 5.2	*** 29.0
Covar: Arousal	0.5	0.0	* 5.6

Note: Emo = Emotion (absolute valence); CNC = Concreteness; IMAG = Imageability; Covar = Covariate; AoA = age of acquisition; FFD = first fixation duration; SFD = single fixation duration; GD = gaze duration. Significant effects are highlighted in **bold**, with significance levels indicated by asterisks: *** $p < .001$, ** $p < .01$, * $p < .05$; marginal effects are highlighted in *italics*.

were below 1.45 across all GLMMs, indicating low multicollinearity.

Emotion and Imageability significantly influenced all fixation time measures, indicating faster processing of emotional compared to neutral words, and of high to low imageability words. Both Emotion × Imageability and Concreteness × Imageability were also significant across all measures. Emotion had a stronger effect on high imageability ($M + SD$) words ($b_{\text{FFD}} = -14$ ms, 95% $CI_{\text{FFD}} [-19, -9]$; $b_{\text{SFD}} = -15$ ms, 95% $CI_{\text{SFD}} [-21, -9]$; $b_{\text{GD}} = -20$ ms, 95% $CI_{\text{GD}} [-27, -12]$) compared to low imageability ($M - SD$) words ($b_{\text{FFD}} = -5$ ms, 95% $CI_{\text{FFD}} [-10, 0]$; $b_{\text{SFD}} = -7$ ms, 95% $CI_{\text{SFD}} [-13, -1]$; $b_{\text{GD}} = 0$ ms, 95% $CI_{\text{GD}} [-7, 6]$). In FFD and SFD, Imageability significantly affected Concrete ($M + SD$) words ($b_{\text{FFD}} = -17$ ms, 95% $CI_{\text{FFD}} [-23, -12]$; $b_{\text{SFD}} = -21$ ms, 95% $CI_{\text{SFD}} [-27, -15]$), but not Abstract ($M - SD$) words ($b_{\text{FFD}} = -5$ ms, 95% $CI_{\text{FFD}} [-10, 1]$; $b_{\text{SFD}} = -2$ ms, 95% $CI_{\text{SFD}} [-8, 4]$). In GD, the pattern reversed;

Imageability significantly affected Abstract words ($b_{\text{GD}} = -26$ ms, 95% $CI_{\text{GD}} [-32, -20]$), but not Concrete words ($b_{\text{GD}} = -2$ ms, 95% $CI_{\text{GD}} [-8, 5]$).

The critical three-way interaction between Emotion, Concreteness, and Imageability was significant in FFD and SFD, and was marginally significant in GD. The range of Imageability ratings differed markedly between Abstract (1.9–4.8) and Concrete words (4.7–6.9), clustering at opposite ends of the distribution. To accurately assess Imageability effects within these respective ranges, we followed Yao et al. (2018) in conducting separate *post hoc* Emotion × Imageability GLMM analyses for Abstract and Concrete words. Both Gamma GLMMs included Emotion, Imageability, their interaction, and log frequency, familiarity, AoA, and arousal as fixed effects, and incorporated by-subject random intercepts and slopes for all fixed effects in the random-effect structure. The VIFs for critical fixed effects were below 1.58 in the Abstract GLMM and below 1.64 in the Concrete GLMM, indicating minimal collinearity. Fixed effects' coefficients are reported in Table 6, with the full results reported in Supplementary Materials B.

Consistent with Yao et al. (2018), the Emotion × Imageability interaction was significant in Abstract words (except marginal for GD), but not in Concrete words (Table 6), as illustrated in Figure 1. Exploring this interaction revealed that in Abstract words (Figure 1A), the Emotion effect was significant at higher imageability levels ($M + SD$) in FFD and SFD ($b_{\text{FFD}} = -15$ ms, 95% $CI_{\text{FFD}} [-24, -6]$; $b_{\text{SFD}} = -21$ ms, 95% $CI_{\text{SFD}} [-32, -11]$), but not in GD ($b_{\text{GD}} = -15$ ms, 95% $CI_{\text{GD}} [-24, -6]$), or at lower imageability levels ($M - SD$; $b_{\text{FFD}} = 6$ ms, 95% $CI_{\text{FFD}} [-2, 13]$; $b_{\text{SFD}} = 6$ ms, 95% $CI_{\text{SFD}} [-3, 14]$; $b_{\text{GD}} = 10$ ms, 95% $CI_{\text{GD}} [-3, 23]$). For Concrete words (Figure 1B), the Emotion effect showed no significant variation with Imageability,

Table 6. GLMM coefficients for Emotion × Imageability by fixation duration measures in Abstract and Concrete words.

	Abstract			Concrete		
	FFD	SFD	GD	FFD	SFD	GD
Intercept	*** 232.0	*** 236.4	*** 277.2	*** 216.2	*** 219.3	*** 250.9
Emo	-4.8	* -7.7	1.3	** -8.7	** -9.6	** -14.6
IMAG	3.5	* 7.9	-9.6	* -7.3	* -8.1	* -11.3
Emo × IMAG	*** -21.1	*** -27.3	-16.8	-6.4	-6.2	-5.6
Covar: Log frequency	** -8.4	** -8.0	*** -19.1	-3.0	* -5.8	*** -30.2
Covar: Familiarity	* -6.5	* -8.2	0.2	*** -12.2	*** -12.0	-1.4
Covar: AoA	*** 7.5	*** 9.3	*** 35.6	-2.5	-1.7	** 11.9
Covar: Arousal	-3.5	-4.6	6.1	1.4	1.4	6.1

Note: Emo = Emotion (absolute valence); IMAG = Imageability; Covar = Covariate; AoA = age of acquisition; FFD = first fixation duration; SFD = single fixation duration; GD = gaze duration. Significant effects are highlighted in **bold**, with significance levels indicated by asterisks: *** $p < .001$, ** $p < .01$, * $p < .05$; marginal effects are highlighted in *italics*.

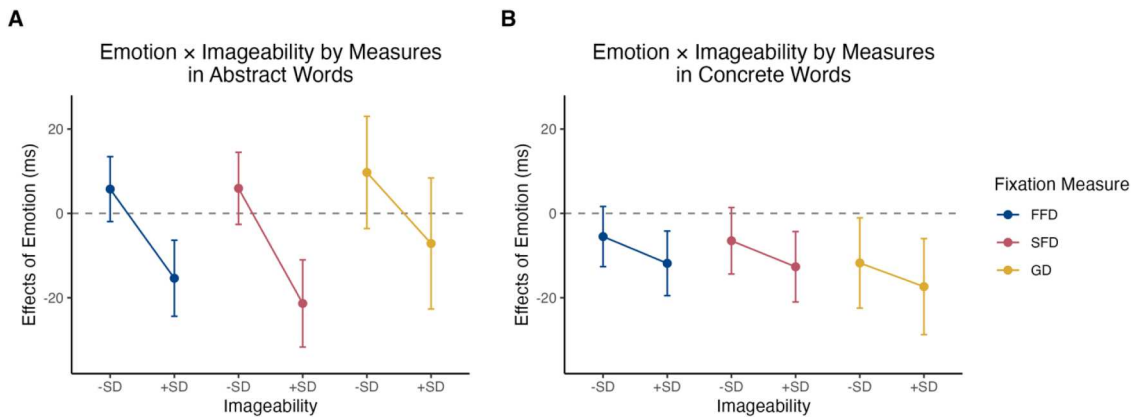


Figure 1. Effects (coefficients) of Emotion by Imageability across fixation measures in Abstract and Concrete words. Panel A = Abstract Words; Panel B = Concrete Words. The points in the graph represent the coefficients (slopes) of the Emotion effects (difference between emotional and neutral words) in milliseconds. Error bars around the points depict the 95% confidence intervals for these coefficients. The horizontal dashed lines represent a zero effect. If the errors bars cross this line, it indicates that the Emotion effects at that specific level of Imageability are not significantly different from zero.

although a trend of higher Emotion effects was observed at higher versus lower imageability levels.

Discussion

The present study aimed to replicate and extend Yao et al.'s (2018) research on the interplay of emotionality, concreteness, and imageability in word processing in natural reading. We used a 3 (Emotion: Negative, Neutral, Positive) \times 2 (Concreteness: Abstract, Concrete) design, and measured fixation durations on 132 target words in contextually neutral, plausible sentences. Gamma GLMM analyses revealed significant main effects of Emotion and Concreteness as well as significant interactions of Emotion \times Concreteness and Emotion \times Concreteness \times Imageability.

In line with previous literature (Knickerbocker et al., 2015, 2019; Scott et al., 2012; Yao et al., 2018), the observed main effects of Emotion and Concreteness indicated faster processing of emotional and concrete words compared to neutral and abstract ones, respectively. The Emotion \times Concreteness interaction was significant in FFD and GD, and marginal in SFD, indicating a larger emotional valence advantage in concrete over abstract words. This finding aligns with the multimodal induction hypothesis, which posits that the emotional activation of concrete words is enhanced by their direct links to sensory and motor experiences that evoke these emotions. This underlines the critical role of sensory experiences in the processing and representation of word emotions.

The multimodal induction hypothesis is further supported by the three-way interaction between Emotion, Concreteness, and Imageability, which was significant in both FFD and SFD, and was marginal in GD. While the emotion facilitation effect was not modulated by imageability in concrete words, it depended on relatively higher imageability in abstract words. These findings are in line with Yao et al.'s (2018), and additionally point to an early influence of imageability on emotion word processing in natural reading. This early interaction underscores the integral role of imageability in processing emotion words, suggesting that the activation of emotional content in abstract words necessitates a critical threshold of sensorimotor activations. Since concrete words inherently possess relatively high levels of imageability, their emotional processing is less affected by variations in imageability. In contrast, for abstract words, reaching this critical threshold of sensorimotor activations is crucial for enabling emotional activation, thereby explaining the interaction between emotion and imageability exclusive to abstract words.

The eye-tracking methodology provides a distinct advantage for assessing word recognition processes during fluent reading. Unlike traditional RT or electrophysiological paradigms that typically present words in isolation, eye-tracking captures the dynamic processes of reading, where word meanings are rapidly activated and integrated online into a broader context. Our fixation time results reveal that the interplay between Emotion, Concreteness, and

Imageability begins in the earliest stages of word processing. Thus, the engagement of sensorimotor simulations is not merely a by-product or elaboration of semantic activation but is foundational to the activation of emotional content in word processing.

Overall, our findings suggest that imageability may have a more pronounced impact on emotional processing in reading than previously understood. While concreteness is tied to sensory experience, imageability extends to how words, including abstract ones, resonate emotionally and cognitively with readers. Neuroimaging studies support this distinction, showing that imageability engages broader brain networks associated with visual imagery and emotional processing, beyond the sensory-motor areas typically activated by concrete words (Binder et al., 2005). It is important to recognise the independent contributions of concreteness and imageability in language processing models, especially in how they interact with emotional valence during word recognition in reading.

Disclosure statement

No potential conflict of interest was reported by the author(s).

Funding

This research was supported by a Carnegie Collaborative Research Grant (Trust Reference 50084) from the Carnegie Dunfermline Trust for the Universities of Scotland awarded to Sara C. Sereno, Graham G. Scott, Gillian Bruce, and Patrick J. O'Donnell (who has since passed away).

Data availability

The data and analysis code and output are shared on the Open Science Framework server: <https://osf.io/zcb8f/>.

ORCID

Bo Yao  <http://orcid.org/0000-0003-1852-2774>
 Graham G. Scott  <http://orcid.org/0000-0002-7296-6264>
 Gillian Bruce  <http://orcid.org/0000-0003-2593-1488>
 Sara C. Sereno  <http://orcid.org/0000-0001-7957-9542>

References

- Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, 67(1), 1–48. <https://doi.org/10.18637/jss.v067.i01>
- Binder, J. R., Westbury, C. F., McKiernan, K. A., Possing, E. T., & Nedler, D. A. (2005). Distinct brain systems for processing concrete and abstract concepts. *Journal of Cognitive Neuroscience*, 17(6), 905–917. <https://doi.org/10.1162/0898929054021102>
- Davies, M. (2004). *BYU-BNC: The British National Corpus*. Oxford University Press. <http://corpus.byu.edu/bnc/>
- Diependaele, K., Brysbaert, M., & Neri, P. (2012). How noisy is lexical decision? *Frontiers in Psychology*, 3, 1–9. <https://doi.org/10.3389/fpsyg.2012.00348>
- Faul, F., Erdfelder, E., Buchner, A., & Lang, A.-G. (2009). Statistical power analyses using G*Power 3.1: Tests for correlation and regression analyses. *Behavior Research Methods*, 41(4), 1149–1160. <https://doi.org/10.3758/BRM.41.4.1149>
- Fox, J., & Weisberg, S. (2019). *An R companion to applied regression* (3rd ed.). Sage Publications.
- Hess, U., & Blairy, S. (2001). Facial mimicry and emotional contagion to dynamic emotional facial expressions and their influence on decoding accuracy. *International Journal of Psychophysiology*, 40(2), 129–141. [https://doi.org/10.1016/S0167-8760\(00\)00161-6](https://doi.org/10.1016/S0167-8760(00)00161-6)
- Kanske, P., & Kotz, S. A. (2007). Concreteness in emotional words: ERP evidence from a hemifield study. *Brain Research*, 1148, 138–148. <https://doi.org/10.1016/j.brainres.2007.02.044>
- Knickerbocker, H., Johnson, R. L., & Altarriba, J. (2015). Emotion effects during reading: Influence of an emotion target word on eye movements and processing. *Cognition and Emotion*, 29(5), 784–806. <https://doi.org/10.1080/02699931.2014.938023>
- Knickerbocker, F., Johnson, R. L., Starr, E. L., Hall, A. M., Preti, D. M., Slate, S. R., & Altarriba, J. (2019). The time course of processing emotion-laden words during sentence reading: Evidence from eye movements. *Acta Psychologica*, 192, 1–10. <https://doi.org/10.1016/j.actpsy.2018.10.008>
- Kousta, S.-T., Vigliocco, G., Vinson, D. P., Andrews, M., & Del Campo, E. (2011). The representation of abstract words: Why emotion matters. *Journal of Experimental Psychology: General*, 140(1), 14–34. <https://doi.org/10.1037/a0021446>
- Kuznetsova, A., Brockhoff, P. B., & Christensen, R. H. B. (2017). lmerTest package: Tests in linear mixed effects models. *Journal of Statistical Software*, 82(13), 1–26. <https://doi.org/10.18637/jss.v082.i13>
- Niedenthal, P. M. (2007). Embodying emotion. *Science*, 316(5827), 1002–1005. <https://doi.org/10.1126/science.1136930>
- Osgood, C. E., Suci, G. J., & Tannenbaum, P. H. (1957). *The measurement of meaning*. University of Illinois Press.
- Palazova, M., Sommer, W., & Schacht, A. (2013). Interplay of emotional valence and concreteness in word processing: An event-related potential study with verbs. *Brain and Language*, 125(3), 264–271. <https://doi.org/10.1016/j.bandl.2013.02.008>
- Rayner, K. (2009). The 35th Sir Frederick Bartlett Lecture: Eye movements and attention in reading, scene perception, and visual search. *Quarterly Journal of Experimental Psychology*, 62(8), 1457–1506. <https://doi.org/10.1080/17470210902816461>
- Rayner, K., Warren, T., Juhasz, B. J., & Liversedge, S. P. (2004). The effect of plausibility on eye movements in reading. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 30(6), 1290–1301. <https://doi.org/10.1037/0278-7393.30.6.1290>
- Scott, G. G., Keitel, A., Becirspahic, M., Yao, B., & Sereno, S. C. (2019). The Glasgow Norms: Ratings of 5,500 words on nine

- scales. *Behavior Research Methods*, 51(3), 1258–1270. <https://doi.org/10.3758/s13428-018-1099-3>
- Scott, G. G., O'Donnell, P. J., & Sereno, S. C. (2012). Emotion words affect eye fixations during reading. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 38(3), 783–792. <https://doi.org/10.1037/a0027209>
- Sereno, S. C., Hand, C. J., Shahid, A., Yao, B., & O'Donnell, P. J. (2018). Testing the limits of contextual constraint: Interactions with word frequency and parafoveal preview during fluent reading. *Quarterly Journal of Experimental Psychology*, 71(1), 302–313. <https://doi.org/10.1080/17470218.2017.1327981>
- Sereno, S. C., & Rayner, K. (2003). Measuring word recognition in reading: Eye movements and event-related potentials. *Trends in Cognitive Sciences*, 7(11), 489–493. <https://doi.org/10.1016/j.tics.2003.09.010>
- Sheikh, N. A., & Titone, D. A. (2013). Sensorimotor and linguistic information attenuate emotional word processing benefits: An eye-movement study. *Emotion*, 13(6), 1107–1121. <https://doi.org/10.1037/a0032417>
- Vigliocco, G., Kousta, S.-T., Della Rosa, P. A., Vinson, D. P., Tettamanti, M., Devlin, J. T., & Cappa, S. F. (2014). The neural representation of abstract words: The role of emotion. *Cerebral Cortex*, 24(7), 1767–1777. <https://doi.org/10.1093/cercor/bht025>
- Wickham, H. (2016). *ggplot2: Elegant graphics for data analysis*. Springer-Verlag. <https://ggplot2.tidyverse.org>
- Yao, B., Keitel, A., Bruce, G., Scott, G. G., O'Donnell, P. J., & Sereno, S. C. (2018). Differential emotional processing in concrete and abstract words. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 44(7), 1064–1074. <https://doi.org/10.1037/xlm0000464>