Socialness Effects in Lexical-Semantic Processing

Diveica, Veronica; Muraki, Emiko J.; Binney, Richard J.; Pexman, Penny M.

Journal of Experimental Psychology: Learning, Memory, and Cognition

DOI:
10.1037/xlm0001328

E-pub ahead of print: 21/03/2024

Peer reviewed version

Socialness Effects in Lexical-Semantic Processing

Veronica Diveica[1,2]*, Emiko J. Muraki[3], Richard J. Binney[1], Penny M. Pexman[3,4]

[1] Cognitive Neuroscience Institute, Department of Psychology, Bangor University, Gwynedd, Wales, UK

[2] Montreal Neurological Institute, Department of Neurology and Neurosurgery, McGill University, Montreal, Quebec, Canada

[3] Department of Psychology and Hotchkiss Brain Institute, University of Calgary, Calgary, Alberta, Canada

[4] Department of Psychology, Western University, London, Ontario, Canada

Abstract

Contemporary theories of semantic representation posit that social experience is an important source of information for deriving meaning. However, there is a lack of behavioural evidence in support of this proposal. The aim of the present work was to test whether words' degree of social relevance, or *socialness*, influences lexical-semantic processing. In Study 1, across a series of item-level regression analyses, we found (1) that socialness can facilitate responses in lexical, semantic and memory tasks, and (2) limited evidence for an interaction of socialness with concreteness. In Studies 2-3, we tested the pre-registered hypothesis that social words, compared to non-social words, will be associated with faster and more accurate responses during a syntactic classification task. We found that socialness has a facilitatory effect on noun decisions (Study 3), but not verb decisions (Study 2). Overall, our results suggest that the socialness of a word affects lexical-semantic processing but also that this is task-dependent. These findings constitute novel evidence in support of proposals that social information is an important dimension of semantic representation.

Accepted Author Manuscript to appear in *Journal of Experimental Psychology: Learning, Memory and Cognition*

***Address correspondence to:** Veronica Diveica, PhD via email: veronica.diveica@mcgill.ca/ veronicadiveica@gmail.com

A central claim of embodied theories of semantic representation is that sensory and motor features are the primary building blocks of concepts (Meteyard et al., 2012). According to this view, perceptual and motor brain systems are necessary for the processing of concepts, even in the absence of direct sensory-motor stimulation by an exemplar (Barsalou et al., 2003; Glenberg, 2015; Martin, 2016). A key challenge for these theories is to account for the representation of abstract concepts, like *confused* and *democracy*, whose referents cannot be directly experienced through the senses (Barsalou, 2016; Dove et al., 2022; Mahon & Caramazza, 2008; Meteyard et al., 2012). Therefore, *multiple representation* theories have proposed that the cognitive systems underpinning emotion, introspection, and language, additionally contribute to the representation of concepts, particularly those that are more abstract (Borghi et al., 2019; Borghi & Binkofski, 2014; Dove et al., 2022; Reilly et al., 2016; Vigliocco et al., 2009). As such, a key tenet of these theories is that all concepts are multidimensional and the contributing dimensions vary systematically depending on concept type (e.g., abstract vs concrete).

Some of these accounts have suggested that social experience provides a key source of semantic information. For example, according to the Situated Action Cycle theory, language understanding, as well as conceptual knowledge, and cognition more broadly, emerges from an interaction between the perceptual modalities and the body situated in physical and social environments (Barsalou, 2020). The Words as Social Tools theory proposes that social experience interacts with linguistic experience to facilitate the acquisition of abstract concepts (Borghi et al., 2019; Borghi & Binkofski, 2014). However, these accounts have yet to provide a clear characterization of the nature of this social information. Consequently, there is limited empirical evidence to support these claims (for a review and in-depth discussion, see Pexman et al., 2023). Given the lack of consensus on the defining features of social words (for a discussion,

see Pexman et al., 2023), Diveica et al. (2023) recently collected and validated socialness norms for a large sample of English words using a *broad* and *inclusive* definition that can encompass a wide range of social experiences. Specifically, they asked participants to rate the degree to which a word's referent has social relevance by referring to a social characteristic of a person or group of people (e.g., trustworthy), a social behaviour (e.g., fight), a social role (e.g., teacher), a social space (e.g., pub), a social institution or system (e.g., nation), a social value or ideology (e.g., feminism), or any other socially relevant construct. This broad and open-ended definition allowed the raters to decide what information is socially relevant, at the expense of clarity on what specific aspects of social experience have driven their judgements. Moreover, this definition can be readily applied to different parts of speech, like verbs referring to social behaviours as well as nouns referring to people, and socially relevant places and entities. The availability of these norms provides the opportunity to directly study the contribution made by socialness, as a broad construct, to lexical-semantic processing.

If social information contributes to conceptual representations, then socialness should be related to behavioural indices of lexical-semantic processing in a similar vein as other established dimensions of word meaning, like concreteness and valence (e.g., Pexman & Yap, 2018; Yap & Seow, 2014). Yet, the relationship between socialness and behaviour has not been systematically explored. Preliminary analyses of the novel socialness ratings have shown that words' degree of social relevance is related to reaction times (RTs) and accuracy scores in a lexical task (Diveica et al., 2023). The facilitatory nature of this association suggests that socialness might make semantic representations richer, by, for example, providing them with additional features or thematic associations, and thereby enabling more efficient word processing (for a review on semantic richness effects, see Pexman, 2012). Notably, socialness accounted for

unique variance in behaviour that could not be explained by basic lexical properties (e.g., letter length, age of acquisition) or other established semantic dimensions that tap into sensorimotor (as measured by Brysbaert et al., 2014), affective (as indexed by valence ratings; Warriner et al., 2013), and linguistic experience (quantified as semantic diversity; Hoffman et al., 2013). These results suggest that the socialness measure captures a distinct aspect of meaning, and that socially relevant information might enrich conceptual representations. However, key questions remain about the contribution of socialness to lexical-semantic processing, including (1) whether its relationship with behaviour is modulated by word concreteness, and (2) whether it generalizes to other behavioural datasets and types of tasks.

Theories that pinpoint social experience as an important source of information for conceptual representation (e.g., Barsalou, 2020; Borghi et al., 2019) also predict that socialness contributes to abstract word meaning more than concrete meaning. Indeed, several feature generation and rating studies have shown that, compared to concrete words, relatively more abstract words tend to be associated with more social features (Barsalou & Wiemer-Hastings, 2005; Harpaintner et al., 2018; Wiemer-Hastings & Xu, 2005) and higher ratings on socially relevant dimensions (Diveica et al., 2023; Troche et al., 2014, 2017; Villani et al., 2019). Likewise, Zdrazilova et al. (2018) showed that, when asked to communicate the meaning of a word without using the word itself, participants used more references to social agents in their descriptions of abstract compared to concrete nouns. Moreover, there is evidence that the influence of some other experience-based semantic dimensions, like emotional information, on lexical-semantic performance is dependent on word concreteness (Newcombe et al., 2012). However, this has not been explored in the case of social semantic content. Therefore, the main

aim of the present work was to test whether the relationship between socialness and behaviour depends on word concreteness.

Some contemporary theories of conceptual representation propose that concepts are dynamic, and that the conceptual information activated at any one time is critically dependent on the ongoing task and the concurrent context (Pexman, 2020; Yee & Thompson-Schill, 2016). Indeed, there is evidence that even robust semantic richness effects are task-dependent (Goh et al., 2016; Pexman et al., 2008; Tousignant & Pexman, 2012). Thus, our secondary aim was to test the relationship between socialness and behaviour across multiple types of lexical-semantic tasks. We expected that the socialness effect will vary as a function of whether social information was task relevant.

To this end, we performed a series of studies using two methodological approaches with complementary strengths. In Study 1, we conducted item-level regression analyses by capitalizing on openly available behavioural megastudy datasets. This allowed us to assess whether socialness correlates with behavioural indices of lexical-semantic processing (e.g., RTs and accuracy scores) across several types of tasks, while accounting for potentially confounding lexical and semantic variables. Strengths of this approach include the fact that it can be applied to large word samples in which word properties vary naturally. However, this approach provides only correlational evidence. To overcome this limitation, in Studies 2-3 we adopted an experimental approach, which also allows for a higher level of control over stimuli characteristics. Specifically, we tested whether manipulating the words' degree of social relevance affects participants' responses on a syntactic classification task (SCT). The SCT has been used in previous semantic richness studies (e.g., Muraki et al., 2020; Sidhu et al., 2014; Yap & Pexman, 2016). Consistent with the semantic richness literature, we hypothesized that higher

levels of socialness will be associated with more efficient lexical-semantic processing, as indexed by faster RTs and higher accuracy scores. Further, in line with the Situated Action Cycle theory and other contemporary models of conceptual processing, we hypothesized that the relationship between word socialness and performance in lexical-semantic tasks will be larger for more abstract words.

Study 1: Socialness as a Predictor of Word Processing across Tasks

Transparency and Openness

For each study, we report how we determined our sample size, all data exclusions (if any), all manipulations, and all measures in the study. All data, analysis code, and research materials are publicly available on the Open Science Framework (OSF) and can be accessed at: https://osf.io/pmyut/ . The design and analysis script for Studies 2-3 were pre-registered on OSF (Study 2 - https://osf.io/jc2s5 ; Study 3 - https://osf.io/2btyd). Data were analysed using open-source software as detailed below.

Method

In this study, we conducted a series of item-level regression analyses with socialness and its interaction with concreteness as the predictors of interest and behavioural indices of lexical-semantic processing as outcome variables. All analyses were conducted using R [version 4.1.1] (R Core Team, 2022) and the following packages: ‘lme4’ (Bates, Mächler, et al., 2015), lmerTest (Kuznetsova et al., 2017) and ‘afex’ (Singmann et al., 2017).

Dependent variables. The outcome variables were obtained from four behavioural mega-studies and included RTs and error rates from the English Lexicon Project visual lexical decision task (LDT) (Balota et al., 2007), the Auditory English Lexicon Project auditory LDT (Goh et al.,

2020), the Calgary Semantic Decision Project semantic decision task (SDT) (Pexman et al., 2017) and hits, false alarms and rate of hits minus false alarms in a recognition memory task (RMT) as collected by (Cortese et al., 2010, 2015). The details of each behavioural mega-study dataset analyzed, including word sample size, are summarized in Table 1. The full methods for each mega-study are provided in their respective papers, thus only brief descriptions are provided below.

- 'Word or non-word?' The LDT outcome variables quantify the speed and accuracy with which participants could distinguish between words and non-word letter strings that were presented visually (LDT visual) and auditorily in either American, British or Singapore accents (LDT auditory). In the case of the auditory LDT, we additionally investigated RT minus stimulus duration (henceforth RT-Duration) because this outcome variable controls for the high variation in the duration of the auditorily-presented word stimuli. In LDT, words that have richer semantic representations are expected to be associated with more efficient processing due to stronger feedback from semantic to orthographic representations (Hino et al., 2002; Hino & Lupker, 1996; Pexman et al., 2002).
- 'Abstract or concrete word?' The SDT outcome variables quantify the speed and accuracy with which participants could classify visually presented words as being concrete or abstract. We analysed the responses to concrete and abstract words separately. The reason for this was that previous research has reported considerable differences in the semantic richness variables that explain concrete and abstract decisions. For example, the processing of concrete words is mainly facilitated by concreteness, which hinders the processing of abstract words; in

contrast, abstract word processing benefits from higher emotional valence and greater semantic diversity (Newcombe et al., 2012; Pexman et al., 2017; Pexman & Yap, 2018). Therefore, we wanted to ensure that our analysis could reveal if the socialness effects are task-specific. As in the original mega study (Pexman et al., 2017), words with concreteness scores below 2.5 were considered abstract while those above 3.5 were considered concrete. In SDT, words associated with richer semantic representations are expected to be associated with more efficient processing due to faster semantic settling (Pexman et al., 2003) and semantic variables tend to explain more variance than in LDT (e.g., Taikh et al., 2015; Yap et al., 2012).

- 'Old or new word?' In the RMT, participants were asked to study a list of ("old") words and were later tested on their ability to recognize these words in a new list that also contained an equal number of previously unstudied ("new") words. The hit rate quantifies the proportion of time a study word was correctly identified as "old". The false alarm rate refers to the proportion of time an unstudied (i.e., ''new'') word was incorrectly judged as ''old''. Hits minus false alarms measure is thought to reflect the ability to discriminate old from new words, and, hence, indexes memory efficiency. There is evidence that words with richer semantic meanings tend to be more accurately remembered (Hargreaves et al., 2012; Paivio et al., 1994; Sidhu & Pexman, 2016).

Each analysis used the maximum number of words available in each behavioural dataset (e.g., a word not being present in the RMT dataset did not preclude its inclusion in the SDT analyses).

We used RTs standardized as z-scores to control for individual differences in overall processing speed (Faust et al., 1999).

Table 1. Details of each behavioural mega-study dataset analysed.

	English Lexicon Project	**Auditory English Lexicon Project**	**Calgary Semantic Decision Project**		**Mega Recognition Memory Studies**
Task type	LDT	LDT	SDT – Concrete decisions	SDT– Abstract decisions	RMT
Task question	'Word or non-word?'	'Word or non-word?'	'Abstract or concrete?'	'Abstract or concrete?'	'Old or new?'
Dependent variables	RTs Error rates	RTs- Duration Error rates	Error rates	Error rates	Hits False alarms Hits – False alarms
Sample size	7991	4609	2485	2340	2560

Note. LDT = lexical decision task; SDT = semantic decision task; RMT = recognition memory task

Independent variables. For the predictors of interest, we used the socialness norms collected by Diveica et al. (2023), which index the degree to which words' referents have social relevance, and the concreteness norms from Brysbaert et al. (2014), which quantify the extent to which the words' referents can be experienced through one of the five senses. Both measures were available for 8388 words, inclusive of nouns, verbs and adjectives.

Control variables. To account for potentially confounding variables, we additionally included several standard lexical and semantic control predictors. The lexical variables included:

letter length, frequency (log subtitle frequency; Brysbaert & New, 2009), orthographic Levenshtein distance (OLD; Yarkoni et al., 2008), and rating-based age of acquisition (Kuperman et al., 2012). For the analyses on auditory LDT responses, the number of phonemes and phonological Levenshtein distance were used as control predictors instead of letter length and OLD, respectively; we additionally controlled for the uniqueness point (the point at which enough phonetic information has been heard to leave only one word-form as a possibility). The semantic variables included valence extremity (the degree to which the word evokes positive/negative feelings; this was measured as the absolute difference between the valence rating and the neutral point of the original valence scale by (Warriner et al., 2013), in addition to concreteness (Brysbaert et al., 2014) and socialness (Diveica et al., 2023). All predictors were standardized. To aid interpretation, the means and standard deviations of the predictors of interest are provided in Table S1 for each dataset analysed.

Results

The effects associated with the predictors of interest are illustrated in Figure 1 and a summary of the statistical estimates can be found in Table S2. In the analyses predicting LDT performance, socialness was a significant predictor, contributing to faster RTs and/or fewer errors regardless of whether the stimuli were presented in the visual or auditory modality. The interaction between socialness and concreteness was only significant in the analyses predicting RT-Duration during the auditory LDT using a British or Singapore accent, such that the effect of socialness on word processing was greater for more abstract words.

In the analyses predicting SDT performance, socialness was a significant predictor in the analysis of concrete word responses (ranging in concreteness from 3.55 to 5 on a 5-point Likert

scale), contributing to faster RTs and fewer errors, but was not significant in the analysis of abstract word responses (ranging in concreteness from 1.07 to 2.48). In the analysis of concrete decisions RTs, there was an interaction between the socialness and the concreteness (continuous) predictors, such that the effect of socialness was greater for relatively more abstract words. In both the analysis of error rates for concrete decisions and the analysis of abstract decisions (RTs and error rates), the interaction between socialness and concreteness was non-significant.

In the analyses predicting RMT performance, socialness was a significant predictor of hit rates only – words with more social relevance tended to be associated with more hits. This effect was accompanied by a significant interaction between socialness and concreteness such that the positive association between socialness and hit rates was greater for more abstract words. Socialness was not a significant predictor of false alarm rates or memory efficiency as indexed by hits minus false alarms. However, the interaction between socialness and concreteness was significant in the case of hits minus false alarms, such that the association was stronger for more abstract words.

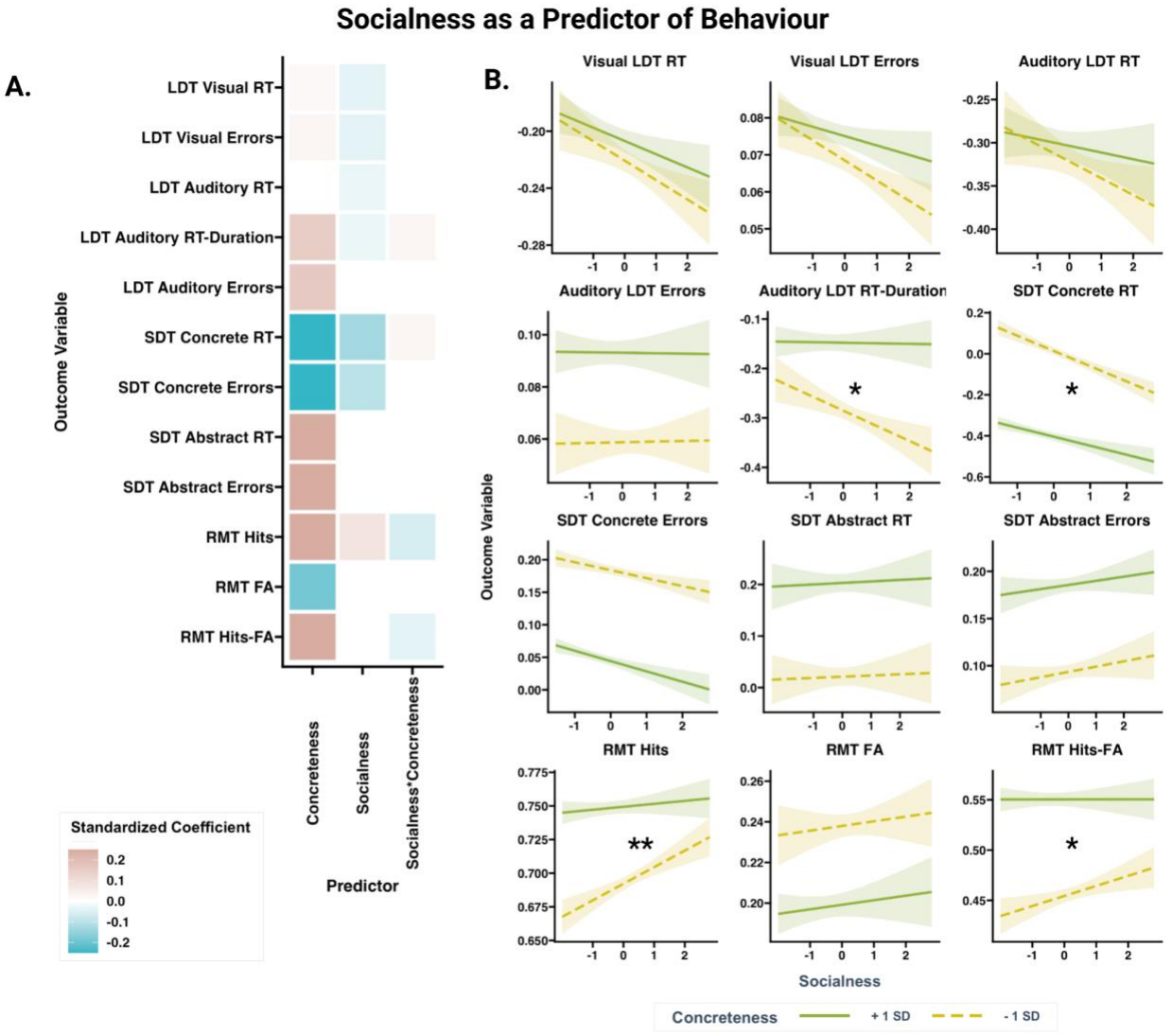

Figure 1. The results of the regression analyses with socialness as a predictor of behavioural performance in lexical-semantic tasks. Panel A illustrates the size of the main effect of socialness, and of concreteness, and of the interaction term between socialness and concreteness, for each of the 12 outcome variables. The pink colour indicates a positive association and blue indicates a negative association. Only significant effects ($p < .05$) are displayed. Panel B illustrates the effect of socialness (x axes) on each outcome variable (y axes) for concrete words (+ 1 SD from the mean concreteness of the respective word sample; depicted in green) and for abstract words (- 1 SD; depicted in yellow). In the case of the Auditory LDT, only the effects

observed in response to stimuli pronounced in a UK accent are illustrated. See Table S2 for the detailed results of the analyses on the American and Singapore accents. * $p < .05$, ** $p < .01$. LDT = lexical decision task, SDT = semantic decision task, RMT = recognition memory task.

Interim discussion

In line with our first prediction, words with higher socialness scores were associated with more efficient lexical-semantic processing. This was evidenced by the finding that increased socialness was associated with shorter RTs and fewer errors when distinguishing words from non-words (in both the visual and auditory modalities) and when categorizing words as concrete, as well as higher hit rates when deciding whether words were old or new. As such, we extended the facilitatory effect of socialness on lexical decisions on visually presented words reported by (Diveica et al., 2023) to the auditory modality, and to two additional types of task – semantic decisions and recognition memory. This pattern of results is in line with the semantic richness literature (Pexman, 2012), and suggests that socialness might enrich a word's conceptual representation, resulting in more efficient processing.

Yet, the association between socialness and lexical-semantic processing was task-specific. We found that socialness was related to concrete decisions but not abstract decisions. This finding was unexpected because of previous claims that social experience is more important for abstract than concrete words (Borghi et al., 2019). A potential explanation for this finding is that, although abstract words tend to have more social features (Barsalou & Wiemer-Hastings, 2005; Troche et al., 2014), this social information is not relevant when judging abstractness, but its absence is diagnostic of concreteness. This possibility is discussed in more depth in the General Discussion.

We found only partial support for our second prediction that the relationship between socialness and lexical-semantic performance is stronger for more abstract words, and thus, for theories that posit a greater contribution of social experience for abstract word meaning (Barsalou, 2020; Borghi et al., 2019). This pattern was found in the case of RTs when distinguishing words from non-words and when categorizing words as concrete, as well as for hit rates and hits-FA when deciding whether words were old or new. However, the association between socialness and behaviour was not modulated by word concreteness in most analyses. Therefore, the next two studies investigated the possibility that the modulatory effect of word concreteness on the relationship between socialness and task performance is detectable in more controlled experimental designs.

Study 2: The Effect of Word Socialness on Verb Judgements

Next, we investigated differences in RTs and accuracy during a syntactic classification task (SCT) for two types of words: social verbs and non-social verbs. If, as some multiple representation theories would propose, social experience enriches conceptual representations, we expected verbs that have high socialness scores to demonstrate facilitatory effects similar to those observed in previous research for words higher in sensorimotor and affective semantic content (Siakaluk, Pexman, Aguilera, et al., 2008; Yap & Seow, 2014). The hypotheses and method were pre-registered prior to data collection at: https://osf.io/jc2s5.

Method

Participants

This study received ethical approval from The University of Calgary Conjoint Faculties Research Ethics Board. The participants were recruited via the online platform Prolific

(www.prolific.co). Responders were restricted to those who had a 100% approval rate on Prolific and self-reported being fluent in English and having no language disorders. Participants completed the study in 20 minutes on average and were compensated with GBP £4. We collected data from a sample of 73 participants, with ages ranging from 18 to 65 (M = 33.32, SD = 10.94). Of the participants, 33 were female, 38 male, 1 non-binary and 1 not reported. On average, participants had completed 15.95 years (SD = 2.51) of formal education.

We determined the target sample by conducting a simulation-based power analysis to make sure that our study had more than .8 power with a .05 alpha error probability to detect an effect of socialness on RTs using linear mixed effect modelling (LMM). As recommended by DeBruine and Barr (2021), we based our simulations on parameters estimated from pilot data (n = 10). We ran 1000 simulations and calculated power by computing the percentage of iterations in which we observed a significant effect of socialness on RTs. The pilot data and power analysis script are available at https://osf.io/pmyut/ . Our power analysis suggested that 70 participants would provide 0.817 power to detect a main effect of socialness on RTs with an effect size of beta = 0.072.

Stimuli

The experimental stimuli were 134 verbs, 67 of which were social verbs, while the other 67 were non-social verbs. In addition, 134 nouns (67 social nouns and 67 non-social nouns) were used as distractor stimuli. The word stimuli were selected starting with a sample of words for which socialness ratings (Diveica et al., 2023) were available. Social words were required to have mean socialness ratings greater than 5 (out of 7), while non-social words have mean ratings less than 3. All words were known by more than 90% of respondents in the prevalence study conducted by Brysbaert et al. (2018) and are used as verbs/nouns more than 80% of the time

according to the dominant part-of speech norms by Brysbaert et al. (2012). The R package 'LexOPS' (Taylor et al., 2020) was used to select words split according to a 2 by 2 factorial design with socialness as the first factor (social, non-social) and dominant part of speech as the second factor (verb, noun), and controlled for letter length, word frequency (Brysbaert & New, 2009), age of acquisition (Kuperman et al., 2012), valence (Warriner et al., 2013) and frequency of the dominant part of speech measured as percentage of occurrences (Brysbaert et al., 2012). In addition, we made sure that the selected words cover a range of concreteness values according to the concreteness norms collected by Brysbaert et al. (2014), which index concreteness using mean ratings on a 5-point Likert scale from abstract (1) to concrete (5). The concreteness of the experimental stimuli ranged from 1.61 to 4.33 and was largely equivalent across the socialness conditions ($t(132) = -1.44$, $p = 0.15$). Summary statistics on each of the above-mentioned variables for the words in each condition are presented in Table 2.

Table 2. Mean and standard deviations for the word stimuli used in Studies 2-3 by experimental condition (before stimulus exclusions). Each table cell contains the following information: Mean (*SD*).

	Verbs		**Nouns**	
	Social	**Non-social**	**Social**	**Non-social**
Letter Length	7.51 (*1.73*)	7.42 (*1.58*)	7.82 (*1.49*)	7.54 (*1.65*)
Frequency	2.16 (*0.6*)	1.99 (*0.54*)	2.15 (*0.61*)	2.06 (*0.6*)
Age of Acquisition	9.55 (*1.82*)	9.62 (*1.68*)	9.71 (*1.79*)	9.73 (*1.81*)
PoS Frequency	0.99 (*0.03*)	1 (*0.02*)	0.99 (*0.02*)	0.99 (*0.02*)
Valence	5.35 (*1.08*)	5.36 (*1.01*)	5.44 (*1.03*)	5.33 (*0.88*)
Concreteness	2.59 (*0.55*)	2.72 (*0.53*)	2.76 (*0.61*)	2.82 (*0.57*)
Socialness	5.51 (*0.4*)	2.57 (*0.27*)	5.56 (*0.4*)	2.49 (*0.4*)

Procedure

Participants were asked to complete a go/no-go syntactic classification task. They were instructed to look at each individual word presented at the center of the screen and determine if it is a verb. If the word is a verb, they were asked to respond by pressing “k” on the keyboard and if the word is not a verb or they do not know the meaning of the word, they were instructed to make no response. Each trial began with a blank screen for 500 ms, followed by a fixation cross for 500 ms which was replaced by a word that remained on the screen for 3000 ms or until a response was made. Participants were asked to respond as quickly and accurately as possible. Before the main experimental task, participants completed a practice block in which they

responded to eight words and received feedback as to whether their response was correct or incorrect. The task (available at: https://osf.io/pmyut/) was implemented in PsychoPy [version 3.2.4] (Peirce et al., 2019) and was presented using Pavlovia (https://pavlovia.org/).

Data cleaning

We sequentially implemented exclusion criteria at the participant, item and individual trial levels. Data from participants with an overall accuracy score less than 0.597 were excluded. This threshold is equivalent to above chance accuracy as determined by a binomial test using a p-value of 0.001. Data associated with words that received correct responses from less than half of the remaining participants were excluded. Finally, trials with RTs less than 250 ms or greater than +3 SD from each participant's mean RT were excluded.

We collected a total of 19564 observations. No participants had below-chance accuracy. Four words were excluded because less than 50% of participants provided correct responses. Of the experimental trials, 251 (2.57%) were identified as RT outliers and were excluded from the analyses. Thus, the analyses reported are based on 9531 experimental observations, out of which 8621 are correct trials.

Analyses

We used the packages 'lme4' (Bates, Mächler, et al., 2015) and 'afex' (Singmann et al., 2017) to perform our statistical analysis in R [version 4.1.1] (R Core Team, 2022). We performed separate confirmatory analyses for each dependent variable. A LMM was used to test the possible effect of socialness on RTs. This analysis was restricted to the verb trials that received correct responses. Word socialness was effect-coded (social: +0.5; non-social: -0.5) and concreteness values were standardised as z-scores. We first estimated the model with the maximal random effects structure justified by the design using the following formula: RT ~

Socialness*Concreteness + (1 + Socialness* Concreteness | participant) + (1 | Word). Given that we did not have hypotheses about the random effects, we tested whether the maximal model was overfitted by following the procedure outlined by Bates, Kliegl et al. (2015) to identify the optimally parsimonious model for our data. We used likelihood ratio tests for model comparisons. The detailed procedure for model selection, along with the code used for the entire process, can be found at https://osf.io/pmyut/. To test the possible effect of socialness on response accuracy, a logistic mixed-effects model was used on all verb trials (correct & incorrect). Like in the case of the analysis performed on RTs, we started with the maximal model and followed an iterative method (Bates, Kliegl, et al., 2015) to reduce model complexity. The R package 'lmerTest' (Kuznetsova et al., 2017) was used to compute p-values for the estimated mixed-effects models via the Satterthwaite's degrees of freedom method. An alpha level of $p < .05$ was used to make inferences about the statistical significance of the results.

To quantify the evidence in favour of the null hypothesis, we conducted a version of the above analysis using Bayesian mixed-effects regression (for a detailed description and tutorial of this approach, see Vasishth et al., 2018). This analysis was not pre-registered. Specifically, we conducted a region of practical equivalence (ROPE) analysis (for a detailed explanation, see Kruschke, 2018), which calculates what percentage of the 95% most credible values, termed the highest density interval (HDI), for each estimated parameter falls within a range of values that are practically equivalent to 0 (i.e., the ROPE), thus representing the null. We adopted Kruschke's (2018) recommendation of a range of ±0.1 SDs in the dependent variable as the ROPE, as this is equivalent to what Cohen (1988) categorized as a negligible effect size. As suggested by Kruschke's (2018), we do not make a discrete decision regarding the null or alternate hypothesis; instead, we report the percentage of the HDI that falls within the ROPE.

This value can intuitively be understood as the percentage of the most likely effect sizes that can be considered negligible. In addition, we report the probability of direction (PD), which varies between 50% and 100%, and quantifies the certainty with which an effect goes in a particular direction (i.e., is positive or negative). Like the frequentist p-value, this index is sensitive only to the amount of evidence for the alternative hypothesis (Makowski et al., 2019). We used the R packages 'rstanarm' (Goodrich et al., 2020) and 'bayestestR' (Makowski et al., 2019) to conduct the Bayesian analyses.

Results

The raw RT data did not meet the assumption of homogeneity of variance. Therefore, below we report the analyses conducted on log transformed RTs, which was not pre-registered. The results of the pre-registered analysis on raw RTs are consistent with the results described below and can be accessed at: osf.io/pmyut/. The model with optimal random effects structure for our logRT data included one random intercept per word, and one random intercept per participant: logRT ~ Socialness*Concreteness + (1|Participant) + (1|Word). The main effects of concreteness and socialness, as well as their interaction, were non-significant (see Table 3 and Figure 2). The Bayesian analyses estimated that 82.91% of the HDI for socialness (90.4% PD), 100% for concreteness (72.99% PD), and 96.87% for the interaction (75.1% PD), fell within the ROPE (-0.015 – 0.015).

Table 3. Study 2 LMM estimates for the effect of socialness and concreteness on log transformed RTs in the verb judgement task (only correct responses).

Predictors	***β***	***CI***	***t***	***p***	***df***
(Intercept)	-0.03	-0.05 – -0.01	-2.62	**0.011**	83.34
Socialness	-0.01	-0.02 – 0.00	-1.33	0.185	123.78
Concreteness	0.00	-0.00 – 0.01	0.63	0.533	124.78
Socialness * Concreteness	-0.00	-0.02 – 0.01	-0.68	0.496	124.76
Random Effects					
σ^2	0.01				
$\tau_{00\ Word}$	0.00				
$\tau_{00\ Participant}$	0.01				
ICC	0.42				
Marginal R^2 / Conditional R^2	0.001 / 0.418				

Note. σ^2: Standard deviation of the residuals; τ_{00}: Standard deviation of the random intercepts; ICC: intra-class correlation coefficient; Marginal R^2: the proportion of variance explained by the fixed effects relative to the overall variance; Conditional R^2: the proportion of variance explained by both fixed and random effects relative to the overall variance.

The model with optimal random effects structure for the accuracy data included one random intercept per word and one random intercept per participant: Accuracy ~ Socialness*Concreteness + (1|Participant) + (1|Word). In these analyses, the main effects of concreteness and socialness, as well as their interaction were non-significant (see Table 4 and

Figure 2). The Bayesian analysis estimated that 57.96% of the HDI for socialness (73.86% PD), 90.39% for concreteness (75.37% PD), and 66.15% for the interaction (60.92% PD), fell within the ROPE (-0.181 – 0.181).

Table 4. Study 2 LMM estimates for the effect of socialness and concreteness on accuracy in the verb judgement task.

Predictors	***Odds Ratio***	***CI***	***t***	***p***
(Intercept)	32.42	20.78 – 50.59	15.32	**<0.001**
Socialness	1.13	0.78 – 1.64	0.65	0.516
Concreteness	0.94	0.78 – 1.13	-0.70	0.483
Socialness * Concreteness	1.06	0.73 – 1.53	0.31	0.756
Random Effects				
σ^2	3.29			
$\tau_{00\ Word}$	0.92			
$\tau_{00\ Participant}$	2.80			
ICC	0.53			
Marginal R^2 / Conditional R^2	0.001 / 0.531			

Note. σ^2: Standard deviation of the residuals; τ_{00}: Standard deviation of the random intercepts; ICC: intra-class correlation coefficient; Marginal R^2: the proportion of variance explained by the fixed effects relative to the overall variance; Conditional R^2: the proportion of variance explained by both fixed and random effects relative to the overall variance.

Interim discussion

Our predictions that (1) the degree of socialness influences syntactic classification responses and (2) this socialness effect is larger for more abstract words were not supported by the data. This could indicate that social information is not relevant for lexical-semantic processing. However, this is unlikely given the associations between socialness and task performance identified in Study 1, and those reported by Diveica et al. (2023). Alternatively, social information might not be diagnostic of a word's verb-ness and so might not be useful when making verb judgements. Therefore, in the next Study we tested the possibility that social information might be relevant when making noun judgements.

Study 3: The Effect of Word Socialness on Noun Judgements

There is evidence that the effects of different semantic dimensions on behavioural measures are dependent on the task (Pexman, 2020) and word type (Pexman et al., 2017; Pexman & Yap, 2018; also see Strik-Lievers et al., 2021). Therefore, the aim of this study was to test whether the findings from Study 1 generalize to noun judgements. We chose noun judgements because nouns tend to have higher socialness scores than verbs when controlling for concreteness (see Figure S1, TableS3). Moreover, for most people, nouns are generally characterized as referring to people, places and things, and therefore, social information might be perceived as particularly relevant for noun decisions. Like in Study 1, we hypothesized that, compared to non-social nouns, social nouns would lead to faster and more accurate responses in a syntactic classification task, and that this effect would be larger for more abstract nouns. We pre-registered the hypotheses and method at: https://osf.io/2btyd.

Method

Participants

This study received ethical approval from The University of Calgary Conjoint Faculties Research Ethics Board. The participants were recruited via the online platform Prolific (www.prolific.co). Responders were restricted to those who did not participate in the first experiment, had a 100% approval rate on Prolific, and self-reported being fluent in English and having no language disorders. Participants completed the study in 20 minutes on average and were compensated with GBP £4. We collected data from a sample of 73 participants, with ages ranging from 18 to 70 (M = 31.22, SD = 11.23). Of the participants, 21 were women, 47 men and 1 not reported. On average, participants had completed 15.85 years of formal education (SD = 2.65).

We conducted a simulation-based sensitivity analysis to estimate the power (alpha = .05) provided by a sample of 70 participants to detect a range of effect sizes. We used parameter estimates based on Study 2 and 1000 iterations to estimate power by calculating the percentage of iterations in which we observed a significant effect of socialness on RTs using LMM. The script used is available at https://osf.io/pmyut/ . Our sensitivity analysis suggested that a study with 70 participants would afford only 34.5% power to detect the effect size observed in the first experiment (25 ms difference between responses to social and non-social words). However, it would provide 84% power to detect a 50 ms difference between conditions, which is similar to the effect sizes reported in a previous study that investigated semantic richness effects on syntactic classification responses (Muraki et al., 2022).

Stimuli

We used the same set of stimuli as in Study 2. However, in this study, the 134 verbs were used as distractor stimuli, whereas the experimental stimuli were the 134 nouns, 67 of which were social nouns while the other 67 were non-social nouns. The concreteness of the experimental stimuli ranged from 1.63 to 4.39 and was largely equivalent across the socialness conditions ($t(132) = 0.56$, $p = 0.57$).

Procedure, data cleaning and analyses

The experimental procedure was the same as in Study 2, with the exception that participants were asked to respond only if the individually presented words were nouns. The approach to data cleaning and analyses was identical to that adopted in Study 2.

We collected a total of 19564 observations, including verb and noun trials. We excluded data from 4 participants with below-chance accuracy. All noun stimuli received correct responses from at least 50% of the remaining participants. However, 643 of the experimental trials (6.57%) were identified as RT outliers and were excluded from the analyses. Therefore, the analyses reported below are based on 9139 experimental observations, out of which 7801 are correct trials.

Results

We conducted the analyses on log transformed RTs because the raw RT data did not meet the assumption of homogeneity of variance. The model with optimal random effects structure for our logRT data included one random intercept per word, one random intercept and three random slopes per participant, but no slope-intercept correlations: LogRT ~ Socialness* Concreteness + (1 + Socialness*Concreteness||Participant) + (1|Word). The main effects of concreteness and socialness were significant: responses were faster for social nouns compared to non-social nouns,

and for nouns with higher concreteness values (see Table 5, Figure 2). There was no significant interaction between socialness and concreteness. The Bayesian analyses estimated that 33.25% of the HDI for socialness (98.58% PD), 37.01% for concreteness (100% PD), and 96.14% for the interaction (62.8% PD), fell within the ROPE (-0.017 – 0.017).

Table 5. Study 3 LMM estimates for the effect of socialness and concreteness on log transformed RTs in the noun judgement task (only correct responses).

Predictors	***β***	***CI***	***t***	***p***	***df***
(Intercept)	0.08	0.05 – 0.10	6.07	**<0.001**	82.58
Socialness	-0.02	-0.04 – -0.00	-2.30	**0.023**	137.64
Concreteness	-0.02	-0.03 – -0.01	-3.87	**<0.001**	151.75
Socialness * Concreteness	-0.00	-0.02 – 0.02	-0.31	0.756	132.67
Random Effects					
σ^2	0.02				
$\tau_{00\ Word}$	0.00				
$\tau_{00\ Participant}$	0.01				
$\tau_{11\ Socialness\ by\ Concreteness}$	0.00				
$\tau_{11\ Concreteness}$	0.00				
$\tau_{11\ Socialness}$	0.00				
ICC	0.11				
Marginal R^2 / Conditional R^2	0.022 / 0.134				

Note: σ^2: Standard deviation of the residuals; τ_{00}: Standard deviation of the random intercepts; τ_{11}: Standard deviation of the random slopes; ICC: intra-class correlation coefficient; Marginal

R^2: the proportion of variance explained by the fixed effects relative to the overall variance; Conditional R^2: the proportion of variance explained by both fixed and random effects relative to the overall variance.

The model with optimal random effects structure for the accuracy data included one random intercept per word, and one random intercept and three random slopes per participant: Accuracy ~ Socialness*Concreteness + (1 + Socialness*Concreteness|Participant) + (1|Word). In these analyses, the main effects of concreteness and socialness were significant: responses were more accurate for social nouns compared to non-social nouns, and for nouns with higher concreteness values (see Table 6, Figure 2). There was no significant interaction between socialness and concreteness. The Bayesian analyses estimated that 8.89% of the HDI for socialness (98.67% PD), 0% for concreteness (99.99% PD), and 69.35% for the interaction (52.91% PD), fell within the ROPE (-0.181 – 0.181). Figure 2 illustrates the effects identified in Studies 2-3.

Table 6. Study 3 LMM estimates for the effect of socialness and concreteness on accuracy in the noun judgement task.

Predictors	***Odds Ratios***	***CI***	***t***	***p***
(Intercept)	12.24	8.82 – 17.00	14.96	**<0.001**
Socialness	1.54	1.07 – 2.23	2.30	**0.022**
Concreteness	1.50	1.25 – 1.82	4.26	**<0.001**
Socialness * Concreteness	0.98	0.68 – 1.39	-0.14	0.892
Random Effects				
σ^2	3.29			
τ_{00} Word	0.74			
τ_{00} Participant	1.35			
τ_{11} Socialness	0.26			
τ_{11} Concreteness	0.09			
τ_{11} Socialness by Concreteness	0.12			
ρ_{01} Socialness	0.28			
ρ_{01} Concreteness	-0.33			
ρ_{01} Socialness by Concreteness	-0.33			
ICC	0.41			
Marginal R^2 / Conditional R^2	0.035 / 0.430			

Note: σ^2: Standard deviation of the residuals/ error variance; τ_{00}: Standard deviation of the random intercepts; τ_{11}: Standard deviation of the random slopes; ρ_{01}: Random correlation between intercepts and slopes; ICC: intra-class correlation coefficient. Marginal R^2 represents

the proportion of variance explained by the fixed effects relative to the overall variance, whereas Conditional R^2 represents the proportion of variance explained by both fixed and random effects relative to the overall variance.

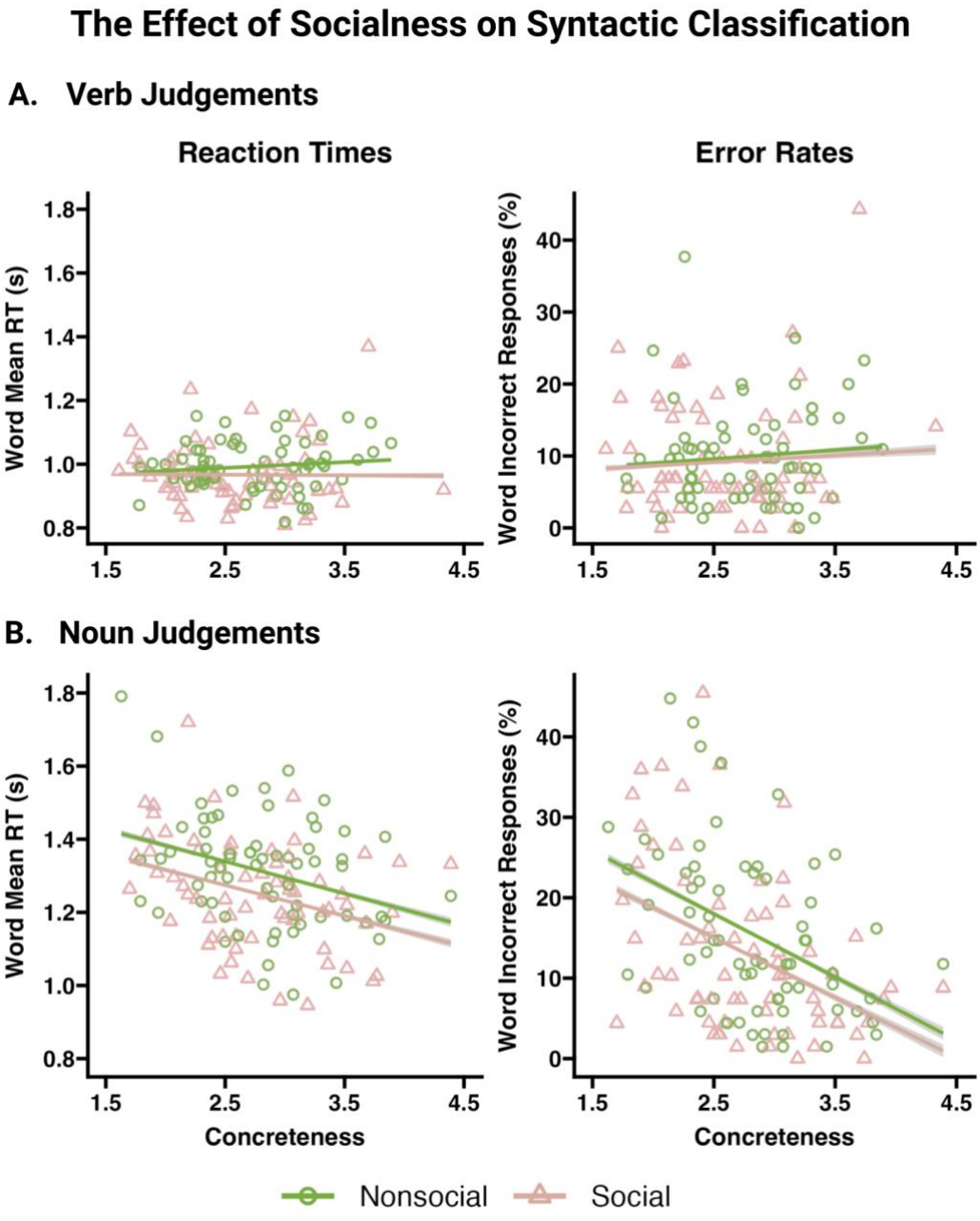

Figure 2. The relationship between concreteness and behavioural responses in the SCT focusing on verb judgements (Panel A, Study 2) and noun judgements (Panel B, Study 3) as a function of

socialness. The datapoints represent mean RTs across all participants for each word (column 1), and percentage of participants that provided incorrect responses for each word (column 2). The pink triangles correspond to the social words and the green circles correspond to the non-social words.

Interim discussion

In line with our first prediction, social nouns were processed faster and more accurately than non-social nouns when making noun judgements. As such, we found that the facilitatory socialness effects found in the correlational analyses conducted in Study 1 replicate in a more tightly controlled experimental design. The results pattern observed is consistent with a semantic richness effect wherein words that are semantically rich (i.e., they have more semantic features, associates etc.) are typically processed more quickly and accurately (Pexman, 2012). Therefore, these results suggest that social experience contributes to the semantic richness of words.

The finding that socialness affects noun judgements but not verb judgements (Study 2) even though the noun and verb stimuli were matched on socialness scores is indicative of susceptibility of the socialness effect to task demands. This is discussed in more depth in the following General Discussion section.

Our second prediction that the influence of socialness on task performance should be stronger for more abstract words was not empirically supported. One explanation for the failure to detect an interaction between socialness and concreteness in our analysis is low statistical power. Nonetheless, our sensitivity analysis demonstrated that we possessed adequate statistical power to detect an effect size akin to those previously reported in studies focusing on semantic

richness (Muraki et al., 2022). This indicates that if there is an interaction, its effect size is likely to be minimal.

General Discussion

There is limited behavioural evidence in support of multiple representation theories that proffer an important role for social experience in conceptual representation. The purpose of the current work was to test the prediction arising from these models that the knowledge derived from social experience influences behavioural responses in lexical-semantic tasks. Specifically, we operationalized social experience using a novel and inclusive socialness measure that assesses the social relevance of a word's referent (Diveica et al., 2023), and tested whether there is an association between socialness and lexical-semantic performance that (1) is modulated by word concreteness, and (2) generalizes across datasets and tasks. Across three studies, we found evidence that increased social relevance can have a faciliatory effect on lexical-semantic processing. In Study 1, we showed that words with higher socialness scores tended to be associated with faster and/or more accurate responses when distinguishing words from non-words, and concrete decisions when distinguishing concrete from abstract words, as well as with more hits when judging whether a word had been previously seen. In Studies 2-3, we found that social words received faster and more accurate responses than non-social words when participants made noun judgements, but not when they made verb judgements. Although the socialness effect was larger for words that are more abstract in some of the datasets we investigated, this interaction was only significant in some tasks (see detailed discussion below). Together, our results demonstrate that socialness influences lexical-semantic processing.

Our main finding was that increased socialness can facilitate responses in lexical-semantic tasks, even after controlling for other key semantic dimensions like concreteness and valence. Therefore, we demonstrate that the facilitatory socialness effect observed by (Diveica et al., 2023) in a visual lexical decision task and a word recognition task generalizes to other behavioural datasets, and across different input modalities (visual and auditory LDT) and task types (SDT, RMT, SCT). Our results are consistent with the semantic richness literature (for a review, see Pexman, 2012), and the idea that "more is better" (Balota et al., 1991, p. 214) – social experience might enrich conceptual representations, which benefits the lexical-semantic processing of social (compared to non-social) words. Indeed, there is evidence that words associated with more semantic information enjoy processing benefits across lexical (e.g., Pexman et al., 2008; Siakaluk, Pexman, Aguilera, et al., 2008; Sidhu et al., 2014; Yap et al., 2015), semantic (Bennett et al., 2011; Goh et al., 2016; Siakaluk, Pexman, Sears, et al., 2008), syntactic categorization (e.g., Muraki et al., 2022; Sidhu et al., 2014; Yap & Pexman, 2016), and memory tasks (e.g., Hargreaves et al., 2012; Sidhu & Pexman, 2016; Taylor et al., 2019). This is believed to be due to more semantic activation and/or faster semantic settling that improves performance in semantic tasks (Pexman, 2012). Greater semantic activation can result in more feedback from semantic to orthographic nodes, facilitating lexical decisions (Hino et al., 2002; Hino & Lupker, 1996; Pexman et al., 2002). Likewise, semantically rich words might lead to more elaborative encoding, which is thought to improve later memory performance (Lockhart & Craik, 1990). Thus, the observed association between socialness and behaviour suggests that socialness contributes to word meaning, providing empirical evidence in support of multiple representation theories that posit a key role for social experience in the acquisition and grounding of concepts (e.g., Barsalou, 2020; Borghi et al., 2019).

Nonetheless, we found that the socialness effect is not ubiquitous and that it is sometimes greater for words that are more abstract. The effect of socialness was greater for more abstract words in the case of RTs when distinguishing auditorily-presented words from non-words, and hits when judging whether words were old or new, as well as for RTs in response to relatively less concrete words when categorizing words as concrete. These former findings are consistent with our hypothesis, but the latter finding that socialness interacts with concreteness for concrete, but not abstract decisions was unexpected. One possible explanation is that socialness helps provide evidence for concrete word meanings when there is less sensorimotor information available. However, interactions were only present in a few task contexts in Study 1. Moreover, we failed to find interactions in Studies 2-3, perhaps due to lower statistical power. Yet, our sensitivity analysis suggested that in Studies 2-3 we had sufficient statistical power to detect an effect size of similar magnitude as previously reported in the semantic richness literature (e.g., Muraki et al., 2022), suggesting that, if the interaction does exist, its effect size is negligible.

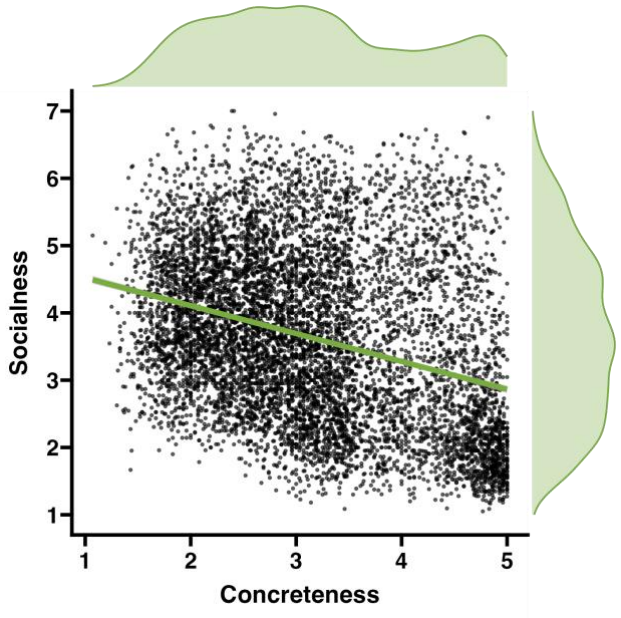

Figure 3. Scatterplot illustrating socialness as a function of concreteness in a sample of 8388 words. The linear relationship between the two variables is highlighted by the green line. The

distributions of concreteness and socialness scores are displayed as density plots on the top and right of the scatterplot. This figure shows that words that are more concrete tend to be less social.

Our secondary finding was that socialness influences lexical-semantic performance in a task-dependent way. Specifically, we found that socialness was related to concrete but not abstract decisions, and that it influenced noun but not verb decisions. Performance in all the tasks we examined is a function of both bottom-up stimulus processing (e.g., semantically richer words elicit greater semantic activation) and top-down task constraints, but the latter might have a greater influence on SDT and SCT responses. Indeed, there is evidence that task influences can attenuate bottom-up effects to the point that behavioural semantic richness effects are no longer observable (Muraki et al., 2023; Tousignant & Pexman, 2012). Thus, the task decisions chosen in the SDT and SCT shape behavioural responses - participants tend to focus on dimensions of meaning that are relevant to the specific decision (Newcombe et al., 2012; Tousignant & Pexman, 2012). The pattern of results observed in these two tasks can therefore be attributed to two sources (i) differences in the contribution of social information to the conceptual representation of the word stimuli and (ii) differences in the usefulness of the socialness dimension when making judgements. Bottom-up influences are unlikely to explain the finding that socialness is related to concrete but not abstract decisions because the concrete word sample had lower mean socialness scores than the abstract word sample. Moreover, they are unlikely to account for the inconsistent results of Studies 2 and 3 because the noun and verb stimuli were matched on socialness. Nonetheless, future research should consider and directly test the possibility that socialness effects, as well as semantic richness effects more generally, are modulated by part of speech. In contrast, task-related top-down influences could have affected

both the semantic and syntactic task results. Abstract words cover the entire socialness continuum (see Figure 3; Diveica et al., 2023; Pexman et al., 2023), perhaps rendering an emphasis on social information unhelpful when making abstractness decisions. For example, both *trust* and the *time* should be categorized as abstract, even though the former is considered social, while the latter is non-social. Thus, the presence of social information might not be useful when making abstractness decisions. In contrast, words that are more concrete tend to have lower socialness scores (see Figure 3), so the lack of social information might be informative when making concreteness decisions. In the SCT, social information might be more diagnostic of noun than verb decisions because verbs have, on average, lower socialness scores compared to nouns. Therefore, we interpret the present results as indicative of the task-dependency of the socialness effect, consistent with a dynamic view of conceptual representation according to which meaning is constructed flexibly to meet the demands of the concurrent context or ongoing task (Pexman, 2020; Yee & Thompson-Schill, 2016).

An important caveat is that our results are dependent on the way socialness was operationalised. We used an inclusive definition of socialness and, thus, it is unclear what specific social information is driving the observed effects. The socialness measure captures a variety of socially relevant constructs, such as social roles, social behaviours, social places, and social institutions. Furthermore, the failure to detect a socialness effect on abstract decisions might be explained by the fact that social information contributes to only some sub-types of abstract words (Harpaintner et al., 2018; Villani et al., 2019). Our analysis treated abstract words as one undifferentiated category, making it challenging to detect more subtle effects specific to only some types of abstract words. This explanation is in line with the growing literature supporting the importance of social features for abstract word meanings (Barsalou & Wiemer-

Hastings, 2005; Harpaintner et al., 2018; Wiemer-Hastings & Xu, 2005; Zdrazilova et al., 2018). Future research should consider more narrowly defined socialness measures (for examples, see Pexman et al., 2023) to better pinpoint what aspects of social experience are most relevant to lexical-semantic performance.

Although our results provide evidence for the benefits of socialness during lexical-semantic processing, they do not provide information about the source of those benefits. There are several possible explanations which cannot be evaluated using the current data. One possibility is that social information is a dimension of semantic richness similar to other experience-based dimensions, like sensorimotor and emotional information, that characterize a unified semantic space. The second possibility is that social experience is particularly important for the acquisition of words, and especially for those that are abstract in nature (Borghi et al., 2019). The importance of socialness for meaning acquisition might have consequences for lexical-semantic processing, such as the effects observed here. However, future research should directly investigate the effect of social information on language acquisition specifically, and its proposed interaction with concreteness. The third possibility is that socialness captures a unique aspect of meaning representation that can also be observed at the neural level. Indeed, there is growing neuroimaging evidence that social information processing relies on distinct neural pathways (for a review, see Pexman et al., 2023; also see Binney & Ramsey, 2020; Conca et al., 2021).

In conclusion, the present work provides a direct demonstration that socialness, even when defined broadly, explains unique variance in lexical-semantic processing, and that task demands dictate when these socialness effects are observed. These findings suggest that socialness is an important dimension of meaning that should be incorporated in future models of

conceptual representation. Multiple representation accounts are well positioned to incorporate these findings.

Funding

This work was supported by the Economic and Social Research Council (ESRC) Wales Doctoral Training Partnership in the form of a PhD studentship (grant no. ES/P00069X/1), a joint award from UK Research and Innovation (UKRI) and Mitacs under the UK-Canada Globalink Doctoral Exchange Scheme (grant no. NE/T014180/1) (both awarded to V.D. and R.J.B.; PhD student: V.D.), and a Natural Sciences and Engineering Research Council (SSHRC) of Canada Discovery Grant (grant no. RGPIN/03860-2018) (awarded to P.M.P.)

Author's Contributions

Veronica Diveica - Conceptualization, Methodology, Software, Formal Analysis, Investigation, Data Curation, Writing - Original Draft, Writing - Review and Editing, Visualisation, Funding Acquisition;

Emiko J. Muraki - Conceptualization, Methodology, Software, Resources, Writing - Original Draft, Writing - Review and Editing;

Richard J. Binney - Conceptualization, Methodology, Writing - Original Draft, Writing - Review and Editing, Supervision, Funding Acquisition;

Penny M. Pexman - Conceptualization, Methodology, Writing - Original Draft, Writing - Review and Editing, Supervision, Project Administration, Funding Acquisition.

References

Balota, D. A., Ferraro, F. R., & Connor, L. T. (1991). On the early influence of meaning in word recognition: A review of the literature. In *The Psychology of Word Meanings* (Schwanenflugel P.J., pp. 187–222). Erlbaum.

Balota, D. A., Yap, M. J., Cortese, M. J., Hutchison, K. A., Kessler, B., Loftis, B., Neely, J. H., Nelson, D. L., Simpson, G. B., & Treiman, R. (2007). The english lexicon project. *Behavior Research Methods*, *39*(3), 445–459. https://doi.org/10.3758/BF03193014

Barsalou, L. W. (2016). On Staying Grounded and Avoiding Quixotic Dead Ends. *Psychonomic Bulletin & Review 2016 23:4*, *23*(4), 1122–1142. https://doi.org/10.3758/S13423-016-1028-3

Barsalou, L. W. (2020). Challenges and Opportunities for Grounding Cognition. *Journal of Cognition*, *3*(1), 1–24. https://doi.org/10.5334/JOC.116

Barsalou, L. W., Simmons, W. K., Barbey, A. K., & Wilson, C. D. (2003). Grounding conceptual knowledge in modality-specific systems. *Trends in Cognitive Sciences*, *7*(2), 84–91. https://doi.org/10.1016/S1364-6613(02)00029-3

Barsalou, L. W., & Wiemer-Hastings, K. (2005). Situating Abstract Concepts. In D. Pecher & R. A. Zwaan (Eds.), *Grounding Cognition: The Role of Perception and Action in Memory, Language, and Thinking* (pp. 129–164). Cambridge University Press.

Bates, D., Kliegl, R., Vasishth, S., & Baayen, R. H. (2015). *Parsimonious Mixed Models*. https://doi.org/10.48550/arxiv.1506.04967

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting Linear Mixed-Effects Models Using lme4. *Journal of Statistical Software*, *67*, 1–48. https://doi.org/10.18637/jss.v067.i01

Bennett, S. D. R., Burnett, A. N., Siakaluk, P. D., & Pexman, P. M. (2011). Imageability and body–object interaction ratings for 599 multisyllabic nouns. *Behavior Research Methods*, *43*(4), 1100–1109. https://doi.org/10.3758/s13428-011-0117-5

Binney, R. J., & Ramsey, R. (2020). Social Semantics: The role of conceptual knowledge and cognitive control in a neurobiological model of the social brain. *Neuroscience & Biobehavioral Reviews*, *112*, 28–38. https://doi.org/10.1016/j.neubiorev.2020.01.030

Borghi, A. M., Barca, L., Binkofski, F., Castelfranchi, C., Pezzulo, G., & Tummolini, L. (2019). Words as social tools: Language, sociality and inner grounding in abstract concepts. *Physics of Life Reviews*, *29*, 120–153. https://doi.org/10.1016/J.PLREV.2018.12.001

Borghi, A. M., & Binkofski, F. (2014). *Words as Social Tools: An Embodied View on Abstract Concepts*. Springer. https://doi.org/10.1007/978-1-4614-9539-0

Brysbaert, M., Mandera, P., McCormick, S. F., & Keuleers, E. (2018). Word prevalence norms for 62,000 English lemmas. *Behavior Research Methods*, *51*(2), 467–479. https://doi.org/10.3758/S13428-018-1077-9

Brysbaert, M., & New, B. (2009). Moving beyond Kučera and Francis: A critical evaluation of current word frequency norms and the introduction of a new and improved word frequency measure for American English. *Behavior Research Methods*, *41*(4), 977–990. https://doi.org/10.3758/BRM.41.4.977

Brysbaert, M., New, B., & Keuleers, E. (2012). Adding part-of-speech information to the SUBTLEX-US word frequencies. *Behavior Research Methods*, *44*(4), 991–997. https://doi.org/10.3758/S13428-012-0190-4

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

Cohen, J. (1988). *Statistical Power Analysis for the Behavioral Sciences* (2nd ed.). Erlbaum.

Conca, F., Borsa, V. M., Cappa, S. F., & Catricalà, E. (2021). The multidimensionality of abstract concepts: A systematic review. *Neuroscience & Biobehavioral Reviews*, *127*, 474–491. https://doi.org/10.1016/J.NEUBIOREV.2021.05.004

Cortese, M. J., Khanna, M. M., & Hacker, S. (2010). Recognition memory for 2,578 monosyllabic words. *Memory*, *18*(6), 595–609. https://doi.org/10.1080/09658211.2010.493892

Cortese, M. J., McCarty, D. P., & Schock, J. (2015). A mega recognition memory study of 2897 disyllabic words. *Quarterly Journal of Experimental Psychology*, *68*(8), 1489–1501. https://doi.org/10.1080/17470218.2014.945096

DeBruine, L. M., & Barr, D. J. (2021). Understanding Mixed-Effects Models Through Data Simulation. *Advances in Methods and Practices in Psychological Science*, *4*(1), 1–15. https://doi.org/10.1177/2515245920965119

Diveica, V., Pexman, P. M., & Binney, R. J. (2023). Quantifying social semantics: An inclusive definition of socialness and ratings for 8388 English words. *Behavior Research Methods*, *55*(2), 461–473. https://doi.org/10.3758/s13428-022-01810-x

Dove, G., Barca, L., Tummolini, L., & Borghi, A. M. (2022). Words have a weight: Language as a source of inner grounding and flexibility in abstract concepts. *Psychological Research*, *86*(8), 2451–2467. https://doi.org/10.1007/s00426-020-01438-6

Faust, M. E., Ferraro, F. R., Balota, D. A., & Spieler, D. H. (1999). Individual differences in information-processing rate and amount: Implications for group differences in response latency. *Psychological Bulletin*, *125*(6), 777–799. https://doi.org/10.1037/0033-2909.125.6.777

Glenberg, A. M. (2015). Few Believe the World Is Flat: How Embodiment Is Changing the Scientific Understanding of Cognition. *Canadian Journal of Experimental Psychology*, *69*(2), 165–171. https://doi.org/10.1037/CEP0000056

Goh, W. D., Yap, M. J., & Chee, Q. W. (2020). The Auditory English Lexicon Project: A multi-talker, multi-region psycholinguistic database of 10,170 spoken words and nonwords. *Behavior Research Methods*. https://doi.org/10.3758/s13428-020-01352-0

Goh, W. D., Yap, M. J., Lau, M. C., Ng, M. M. R., & Tan, L.-C. (2016). Semantic Richness Effects in Spoken Word Recognition: A Lexical Decision and Semantic Categorization Megastudy. *Frontiers in Psychology*, *7*. https://www.frontiersin.org/articles/10.3389/fpsyg.2016.00976

Goodrich, B., Gabry, J., Ali, I., & Brilleman, S. (2020). *rstanarm: Bayesian applied regression modeling via Stan* [Computer software]. https://mc-stan.org/rstanarm

Hargreaves, I., Pexman, P., Johnson, J., & Zdrazilova, L. (2012). Richer concepts are better remembered: Number of features effects in free recall. *Frontiers in Human Neuroscience*, *6*. https://www.frontiersin.org/articles/10.3389/fnhum.2012.00073

Harpaintner, M., Trumpp, N. M., & Kiefer, M. (2018). The Semantic Content of Abstract Concepts: A Property Listing Study of 296 Abstract Words. *Frontiers in Psychology*, *9*, 1748. https://doi.org/10.3389/fpsyg.2018.01748

Hino, Y., & Lupker, S. J. (1996). Effects of polysemy in lexical decision and naming: An alternative to lexical access accounts. *Journal of Experimental Psychology: Human Perception and Performance*, *22*(6), 1331. https://doi.org/10.1037/0096-1523.22.6.1331

Hino, Y., Lupker, S. J., & Pexman, P. M. (2002). Ambiguity and Synonymy Effects in Lexical Decision, Naming, and Semantic Categorization Tasks: Interactions between Orthography, Phonology, and Semantics. *Journal of Experimental Psychology: Learning Memory and Cognition*, *28*(4), 686–713. https://doi.org/10.1037/0278-7393.28.4.686

Hoffman, P., Lambon Ralph, M. A., & Rogers, T. T. (2013). Semantic diversity: A measure of semantic ambiguity based on variability in the contextual usage of words. *Behavior Research Methods*, *45*(3), 718–730. https://doi.org/10.3758/s13428-012-0278-x

Kruschke, J. K. (2018). Rejecting or Accepting Parameter Values in Bayesian Estimation. *Advances in Methods and Practices in Psychological Science*, *1*(2), 270–280. https://doi.org/10.1177/2515245918771304

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

Kuznetsova, A., Brockhoff, P. B., & Christensen, R. H. B. (2017). lmerTest Package: Tests in Linear Mixed Effects Models. *Journal of Statistical Software*, *82*, 1–26. https://doi.org/10.18637/jss.v082.i13

Lockhart, R. S., & Craik, F. I. M. (1990). Levels of processing: A retrospective commentary on a framework for memory research. *Canadian Journal of Psychology / Revue Canadienne de Psychologie*, *44*, 87–112. https://doi.org/10.1037/h0084237

Mahon, B. Z., & Caramazza, A. (2008). A critical look at the embodied cognition hypothesis and a new proposal for grounding conceptual content. *Journal of Physiology-Paris*, *102*(1), 59–70. https://doi.org/10.1016/j.jphysparis.2008.03.004

Makowski, D., Ben-Shachar, M. S., & Lüdecke, D. (2019). BayestestR: Describing Effects and their Uncertainty, Existence and Significance within the Bayesian Framework. *The Journal of Open Source Software*, *4*(40), 541. https://doi.org/10.21105/joss.01541

Martin, A. (2016). GRAPES—Grounding representations in action, perception, and emotion systems: How object properties and categories are represented in the human brain. *Psychonomic Bulletin & Review*, *23*(4), 979–990. https://doi.org/10.3758/s13423-015-0842-3

Meteyard, L., Cuadrado, S. R., Bahrami, B., & Vigliocco, G. (2012). Coming of age: A review of embodiment and the neuroscience of semantics. *Cortex*, *48*(7), 788–804. https://doi.org/10.1016/j.cortex.2010.11.002

Muraki, E. J., Doyle, A., Protzner, A. B., & Pexman, P. M. (2023). Context matters: How do task demands modulate the recruitment of sensorimotor information during language processing? *Frontiers in Human Neuroscience*, *16*, 976954. https://doi.org/10.3389/fnhum.2022.976954

Muraki, E. J., Sidhu, D. M., & Pexman, P. M. (2020). Mapping semantic space: Property norms and semantic richness. *Cognitive Processing*, *21*(4), 637–649. https://doi.org/10.1007/s10339-019-00933-y

Muraki, E. J., Sidhu, D. M., & Pexman, P. M. (2022). Heterogenous abstract concepts: Is "ponder" different from "dissolve"? *Psychological Research*, *86*(8), 2478–2494. https://doi.org/10.1007/s00426-020-01398-x

Newcombe, P., Campbell, C., Siakaluk, P., & Pexman, P. (2012). Effects of Emotional and Sensorimotor Knowledge in Semantic Processing of Concrete and Abstract Nouns. *Frontiers in Human Neuroscience*, *6*. https://www.frontiersin.org/articles/10.3389/fnhum.2012.00275

Paivio, A., Walsh, M., & Bons, T. (1994). Concreteness effects on memory: When and why? *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *20*, 1196–1204. https://doi.org/10.1037/0278-7393.20.5.1196

Peirce, J., Gray, J. R., Simpson, S., MacAskill, M., Höchenberger, R., Sogo, H., Kastman, E., & Lindeløv, J. K. (2019). PsychoPy2: Experiments in behavior made easy. *Behavior Research Methods*, *51*(1), 195–203. https://doi.org/10.3758/s13428-018-01193-y

Pexman, P. M. (2012). Meaning-level influences on visual word recognition. In J. Adelman (Ed.), *Visual Word Recognition: Meaning and context, individuals and development* (pp. 24–43). Psychology Press.

Pexman, P. M. (2020). How Does Meaning Come to Mind? Four Broad Principles of Semantic Processing. *Canadian Journal of Experimental Psychology*, *74*(4), 275–283. https://doi.org/10.1037/CEP0000235

Pexman, P. M., Diveica, V., & Binney, R. J. (2023). Social semantics: The organization and grounding of abstract concepts. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *378*(1870), 20210363. https://doi.org/10.1098/rstb.2021.0363

Pexman, P. M., Hargreaves, I. S., Siakaluk, P. D., Bodner, G. E., & Pope, J. (2008). There are many ways to be rich: Effects of three measures of semantic richness on visual word recognition. *Psychonomic Bulletin & Review*, *15*(1), 161–167. https://doi.org/10.3758/PBR.15.1.161

Pexman, P. M., Heard, A., Lloyd, E., & Yap, M. J. (2017). The Calgary semantic decision project: Concrete/abstract decision data for 10,000 English words. *Behavior Research Methods*, *49*(2), 407–417. https://doi.org/10.3758/s13428-016-0720-6

Pexman, P. M., Holyk, G. G., & Monfils, M.-H. (2003). Number-of-features effects and semantic processing. *Memory & Cognition*, *31*(6), 842–855. https://doi.org/10.3758/bf03196439

Pexman, P. M., Lupker, S. J., & Hino, Y. (2002). The impact of feedback semantics in visual word recognition: Number-of-features effects in lexical decision and naming tasks. *Psychonomic Bulletin & Review*, *9*(3), 542–549. https://doi.org/10.3758/BF03196311

Pexman, P. M., & Yap, M. J. (2018). Individual differences in semantic processing: Insights from the Calgary Semantic Decision Project. *Journal of Experimental Psychology: Learning Memory and Cognition*, *44*(7), 1091–1112. https://doi.org/10.1037/XLM0000499

R Core Team. (2022). *R: The R Project for Statistical Computing* [Computer software]. https://www.r-project.org/

Reilly, J., Peelle, J. E., Garcia, A., & Crutch, S. J. (2016). Linking somatic and symbolic representation in semantic memory: The dynamic multilevel reactivation framework. *Psychonomic Bulletin & Review 2015 23:4*, *23*(4), 1002–1014. https://doi.org/10.3758/S13423-015-0824-5

Siakaluk, P. D., Pexman, P. M., Aguilera, L., Owen, W. J., & Sears, C. R. (2008). Evidence for the activation of sensorimotor information during visual word recognition: The body–object interaction effect. *Cognition*, *106*(1), 433–443. https://doi.org/10.1016/j.cognition.2006.12.011

Siakaluk, P. D., Pexman, P. M., Sears, C. R., Wilson, K., Locheed, K., & Owen, W. J. (2008). The Benefits of Sensorimotor Knowledge: Body–Object Interaction Facilitates Semantic Processing. *Cognitive Science*, *32*(3), 591–605. https://doi.org/10.1080/03640210802035399

Sidhu, D. M., Kwan, R., Pexman, P. M., & Siakaluk, P. D. (2014). Effects of relative embodiment in lexical and semantic processing of verbs. *Acta Psychologica*, *149*, 32–39. https://doi.org/10.1016/j.actpsy.2014.02.009

Sidhu, D. M., & Pexman, P. M. (2016). Is Moving More Memorable than Proving? Effects of Embodiment and Imagined Enactment on Verb Memory. *Frontiers in Psychology*, *7*. https://www.frontiersin.org/articles/10.3389/fpsyg.2016.01010

Singmann, H., Bolker, B., Westfall, J., & Aust, F. (2017). *Packacke "afex"* (1.1-1) [Computer software]. https://github.com/singmann/afex

Strik-Lievers, F., Bolognesi, M., & Winter, B. (2021). The linguistic dimensions of concrete and abstract concepts: Lexical category, morphological structure, countability, and etymology. *Cognitive Linguistics*, *32*(4), 641–670. https://doi.org/10.1515/COG-2021-0007/ASSET/GRAPHIC/J_COG-2021-0007_FIG_005.JPG

Taikh, A., Hargreaves, I. S., Yap, M. J., & Pexman, P. M. (2015). Semantic classification of pictures and words. *Quarterly Journal of Experimental Psychology*, *68*(8), 1502–1518. https://doi.org/10.1080/17470218.2014.975728

Taylor, J. E., Beith, A., & Sereno, S. C. (2020). LexOPS: An R package and user interface for the controlled generation of word stimuli. *Behavior Research Methods*, *52*(6), 2372–2382. https://doi.org/10.3758/s13428-020-01389-1

Taylor, R. S., Francis, W. S., Borunda-Vazquez, L., & Carbajal, J. (2019). Mechanisms of word concreteness effects in explicit memory: Does context availability play a role? *Memory & Cognition*, *47*(1), 169–181. https://doi.org/10.3758/s13421-018-0857-x

Tousignant, C., & Pexman, P. (2012). Flexible recruitment of semantic richness: Context modulates body-object interaction effects in lexical-semantic processing. *Frontiers in Human Neuroscience*, *6*. https://www.frontiersin.org/articles/10.3389/fnhum.2012.00053

Troche, J., Crutch, S. J., & Reilly, J. (2017). Defining a Conceptual Topography of Word Concreteness: Clustering Properties of Emotion, Sensation, and Magnitude among 750 English Words. *Frontiers in Psychology*, *8*. https://doi.org/10.3389/fpsyg.2017.01787

Troche, J., Crutch, S., & Reilly, J. (2014). Clustering, hierarchical organization, and the topography of abstract and concrete nouns. *Frontiers in Psychology*, *5*, 360. https://doi.org/10.3389/FPSYG.2014.00360

Vasishth, S., Nicenboim, B., Beckman, M. E., Li, F., & Kong, E. J. (2018). Bayesian data analysis in the phonetic sciences: A tutorial introduction. *Journal of Phonetics*, *71*, 147–161. https://doi.org/10.1016/j.wocn.2018.07.008

Vigliocco, G., Meteyard, L., Andrews, M., & Kousta, S. (2009). Toward a theory of semantic representation. *Language and Cognition*, *1*(2), 219–247. https://doi.org/10.1515/LANGCOG.2009.011

Villani, C., Lugli, L., Liuzza, M., & Borghi, A. M. (2019). Varieties of abstract concepts and their multiple dimensions. *Language and Cognition*, *11*(3), 403–430. https://doi.org/10.1017/LANGCOG.2019.23

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

Wiemer-Hastings, K. K., & Xu, X. (2005). Content Differences for Abstract and Concrete Concepts. *Cognitive Science*, *29*(5), 719–736. https://doi.org/10.1207/S15516709COG0000_33

Yap, M. J., Lim, G. Y., & Pexman, P. M. (2015). Semantic richness effects in lexical decision: The role of feedback. *Memory & Cognition*, *43*(8), 1148–1167. https://doi.org/10.3758/s13421-015-0536-0

Yap, M. J., & Pexman, P. M. (2016). Semantic Richness Effects in Syntactic Classification: The Role of Feedback. *Frontiers in Psychology*, *7*. https://www.frontiersin.org/articles/10.3389/fpsyg.2016.01394

Yap, M. J., & Seow, C. S. (2014). The influence of emotion on lexical processing: Insights from RT distributional analysis. *Psychonomic Bulletin & Review*, *21*(2), 526–533. https://doi.org/10.3758/s13423-013-0525-x

Yap, M., Pexman, P., Wellsby, M., Hargreaves, I., & Huff, M. (2012). An Abundance of Riches: Cross-Task Comparisons of Semantic Richness Effects in Visual Word Recognition. *Frontiers in Human Neuroscience*, *6*. https://www.frontiersin.org/articles/10.3389/fnhum.2012.00072

Yarkoni, T., Balota, D., & Yap, M. (2008). Moving beyond Coltheart's N : A new measure of orthographic similarity. *Psychonomic Bulletin & Review*, *15*(5), 971–979. https://doi.org/10.3758/PBR.15.5.971

Yee, E., & Thompson-Schill, S. L. (2016). Putting concepts into context. *Psychonomic Bulletin & Review 2015 23:4*, *23*(4), 1015–1027. https://doi.org/10.3758/S13423-015-0948-7

Zdrazilova, L., Sidhu, D. M., & Pexman, P. M. (2018). Communicating abstract meaning: Concepts revealed in words and gestures. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *373*(1752), 20170138. https://doi.org/10.1098/rstb.2017.0138