AMERICAN PSYCHOLOGICAL ASSOCIATION

Journal of Experimental Psychology:
Learning, Memory, and Cognition

ISSN: 0278-7393

2023, Vol. 49, No. 10, 1572–1587
https://doi.org/10.1037/xlm0001265

More Is Not Necessarily Better: How Different Aspects of Sensorimotor Experience Affect Recognition Memory for Words

Agata Dymarska[1], Louise Connell[1, 2], and Briony Banks[1]
[1] Department of Psychology, Lancaster University
[2] Department of Psychology, Maynooth University

Semantic richness theory predicts that words with richer, more distinctive semantic representations should facilitate performance in a word recognition memory task. We investigated the contribution of multiple aspects of sensorimotor experience—those relating to the body, communication, food, and objects—to word recognition memory, by analyzing megastudy data in a series of hierarchical linear regressions. We found that different forms of sensorimotor experience produced different effects on memory. While stronger grounding in object- and food-related experience facilitated word memory performance as expected for semantic richness, experience relating to communication did not. Critically, sensorimotor experience relating to the body impaired rather than facilitated recognition memory by inflating false alarms, which was not consistent with the idea that semantically richer representations are more memorable. Additionally, we found that pure imageability (i.e., consciously generating mental imagery, distinct from sensorimotor experience) contributes to semantic richness effects on word memory but with much smaller effect sizes than previously reported, once sensorimotor grounding was taken into account. These results suggest that word recognition memory is often but not consistently facilitated by rich semantic representations and that it is essential to separately consider distinct forms of sensorimotor experience rather than assuming more information is always better. The findings have implications for the use of semantic variables in memory research.

Keywords: word memory, sensorimotor information, imageability, semantic richness

Conceptual representations of word meaning consist of multidimensional information about their referents from a range of different sources (Barsalou et al., 2008; Connell & Lynott, 2014; Vigliocco et al., 2009). Many semantic variables capture information about word meaning by measuring different aspects of our experience with and/or understanding of the concept, such as its number of features (McRae et al., 2005; Pexman et al., 2003), emotional valence (Warriner et al., 2013), and strength of sensory experience (Lynott et al., 2020; Lynott & Connell, 2009, 2013; see also Juhasz & Yap, 2013) or action experience (Lynott et al., 2020; Tillotson et al., 2008). Semantic richness theory (Buchanan et al., 2001; Pexman et al., 2008) proposes that encountering a word whose referent concept has a richer semantic representation makes it easier to process and ultimately respond to the word. Semantic richness effects have mainly been demonstrated in visual word recognition research, where variables such as the number of conceptual features, density of distributional neighborhood, sensory experience, and body–object interaction ratings have all predicted performance in word reading tasks, such as lexical decision and word naming (Pexman et al., 2008; Recchia & Jones, 2012; Yap et al., 2011; Zdrazilova & Pexman, 2013). Not all semantic richness variables affect all tasks equally, but where effects appear they are consistently facilitatory (Yap et al., 2011). That is, some word meanings are more richly varied than others in their semantic representation—more features or neighbors, stronger sensory or bodily interaction—and are processed more quickly and easily as a result.

This semantic richness theory has been extended to memory, where many studies show that words with a richer semantic representation are remembered better (Bourassa & Besner, 1994; Khanna & Cortese, 2021; Lau et al., 2018; Sidhu & Pexman, 2016; Snefjella & Kuperman, 2016). Specifically, a word richer in semantic information elicits stronger semantic activation (Pexman et al., 2013) and therefore produces a stronger memory trace (Hargreaves et al., 2012; Sidhu & Pexman, 2016). Such rich representations are more likely to be correctly recognized as old (higher hit rates), because it is less likely that their memory trace will fade or be replaced by interfering information. In support of this idea, many semantic variables such as higher imageability (Cortese et al., 2010, 2015; Groninger, 1974), higher body–object interaction (BOI, Sidhu & Pexman, 2016), and higher animacy and perceived threat (Bonin et al., 2014; Leding, 2020) all lead to higher hit rates. Moreover, higher imageability and arousal have also been found to reduce false alarm rates (Cortese et al., 2010, 2015; Lau

---

Agata Dymarska https://orcid.org/0000-0003-0931-3081
Louise Connell https://orcid.org/0000-0002-5291-5267
Briony Banks https://orcid.org/0000-0003-3666-0222

This work was supported by the European Research Council (ERC) under the European Union's Horizon 2020 Research and Innovation Program (Grant agreement 682848) to Louise Connell.

Correspondence concerning this article should be addressed to Agata Dymarska, Department of Psychology, Lancaster University, Fylde Ave, Bailrigg, Lancaster LA1 4YF, United Kingdom. Email: a.dymarska@lancaster.ac.uk

et al., 2018; but cf., Ballot et al., 2021 regarding imageability); the classic mirror effect often observed for semantic variables in recognition memory (Glanzer & Adams, 1985), which Lau and colleagues have linked to the idea that semantic richness facilitates word distinctiveness. In general, the more distinctive an encountered word is, the less likely it is to be mistaken for a previously seen word (see, e.g., Glanzer & Adams, 1985; Zechmeister, 1972). Hence, words that are low-frequency or have small orthographic neighborhoods are associated with lower false alarms (Cortese et al., 2015; Glanc & Greene, 2007), as are words with high imageability or arousal, because their distinctive representations are not easily confused with the memory trace of similar items. However, not all semantic richness variables facilitate distinctiveness in this straightforward manner. Lau et al. (2018) found that words with a higher number of senses (i.e., polysemous words with multiple meanings) have a *higher* false alarm rate, which they suggest is because their ambiguous representations are less distinctive and therefore easier to confuse with previously seen words. Accordingly, they conclude that semantic richness effects on recognition memory are constrained by distinctiveness, such that semantically rich words only make good retrieval cues when their representations are also distinctive and are otherwise prone to create false alarms due to mistaken overlap with the memory trace.

Our focus in this paper is how sensorimotor experience, as a form of semantic richness, contributes to the recognition memory for words. The sensorimotor experience underlying word meaning is varied and multidimensional (Fernandino et al., 2022; Kiefer & Pulvermüller, 2012; Lynott et al., 2020), but its effects on word memory have only been examined in a relatively limited way to date, such as via imageability ratings (i.e., the ease of generating a mental image for a word: Cortese et al., 2010, 2015; Khanna & Cortese, 2021; Lau et al., 2018; see also Paivio, 1971) or via a composite score of perceptual or action strength (Khanna & Cortese, 2021). Each of these measures, however, offers only a flawed approximation of the perception and action underpinnings of a concept. People are generally unable to condense all perceptual experience into a single rating (e.g., imageability, sensory experience ratings) without neglecting and distorting all but the dominant modality (Connell & Lynott, 2016a). Imageability, in particular, is heavily visually biased (Connell & Lynott, 2012; Speed & Brysbaert, 2022). Composite measures that aggregate individual perceptual modalities or action effectors into a single score, weighted by dominance, offer a methodological improvement (e.g., Minkowski-3 perceptual strength, employed as a predictor of word memory by Khanna & Cortese, 2021, is a composite of six individually rated perceptual modalities in Lynott et al.'s norms), but necessarily lose a lot of information and explain less variance in word reading than considering multiple dimensions separately (Speed & Brysbaert, 2022). While all forms of sensorimotor experience should in principle contribute to semantic richness—and representational distinctiveness—and thereby enhance word memory, it remains an open question whether they actually do so. Given that, for example, taste and/or smell experience appears to contribute little to lexical decision and word naming tasks (Dymarska et al., 2023a; Speed & Brysbaert, 2022), it is important to establish whether different forms of sensorimotor experience differentially affect word memory.

On a final note, it is also important to establish how imageability—a theoretically distinct construct that is concerned with the ease of generating mental imagery—affects word memory independently of different forms of sensorimotor experience. Although the effects of imageability on word memory are long established (e.g., Paivio, 1971; Paivio & Csapo, 1969, 1973; Rubin & Friendly, 1986), recent research has raised issues regarding the reliability of imageability effects and potential issues with construct validity. As mentioned above, there is conflicting evidence for the effects of imageability on recognition memory, with higher imageability sometimes reducing false alarm rates (e.g., Cortese et al., 2010, 2015) and sometimes increasing them (Ballot et al., 2021; see also Peterson & McGee, 1974). These discrepant findings may be due to methodological differences, such as the fact that the studies sampled different items or did not use the same control variables alongside imageability. Alternatively (or additionally), the discrepancy may be due to the fact that imageability ratings in different studies came from different participant groups. In recent work, we have found that imageability ratings from different sources—despite using identical instructions and scale—vary enormously in the conceptual information that they capture, likely due to inconsistent strategies in how participants interpret and rate imageability in different norming studies (Dymarska et al., 2023a). As a result, different sources of imageability ratings fail to consistently and reliably facilitate word reading once lexical predictors and sensorimotor grounding have already been taken into account. This research raises a possibility that the construct of imageability itself—that is, the ease of generating mental imagery for a word—is not what imageability ratings actually measure (i.e., imageability may have poor construct validity). Alternatively, it is possible that ease of generating mental imagery is indeed captured by imageability ratings, but it does not usefully contribute in its own right to semantic richness effects on word reading. If the former possibility is true and imageability has poor construct validity, then (as with lexical decision and word naming) imageability ratings from different sources will fail to facilitate word recognition memory over and above sensorimotor grounding. On the other hand, if the latter possibility is true and imageability has adequate construct validity despite its unreliable effects on some tasks, then imageability ratings from different sources could still contribute to semantic richness effects on recognition memory (i.e., increasing hit rates and reducing false alarms), over and above sensorimotor grounding, due to the different task demands inherent in retrieving a studied list of words. Concurrently examining the effects of both sensorimotor grounding and multiple sources of imageability on word recognition memory performance will allow us to disentangle their contributions.

Current Studies

In the current series of studies, our aim was to examine the effects of various aspects of sensorimotor experience and imageability on word memory. We first explored whether different forms of sensorimotor experience in meaning representation consistently facilitate word memory (i.e., higher hit rates, lower false alarms), as would be predicted by semantic richness theory. We investigated this issue in Study 1 using a megastudy (Balota et al., 2012) hierarchical regression analysis of 5,305 words from the combined word recognition memory datasets of Cortese et al. (2010, 2015), predicted by a broad range of lexical and semantic information including sensorimotor grounding in multiple forms of perception and action experience. Additionally, given the unstable nature of imageability effects in word reading (lexical decision and word naming, Dymarska et al., 2023a), we wanted to investigate whether the effect of imageability on word recognition memory is independent of sensorimotor grounding and is stable across different sources of imageability, as would be expected for a semantic richness variable with sound construct validity. We addressed this question in Study 2 using imageability ratings from six different sources (all based on the same norming instructions and scale) as predictors of word recognition memory over and above the lexical and sensorimotor effects of Study 1.

Study 1: Different Forms of Sensorimotor Experience

The aim of Study 1 was to test whether different forms of sensorimotor experience all influence word recognition memory performance. We used data from existing megastudies of word recognition memory (Cortese et al., 2010, 2015) as dependent measures at the item level (hits, false alarms, hit rate minus false alarm rate, d', c). Sensorimotor information was based on the Lancaster Sensorimotor Norms (Lynott et al., 2020), which comprise ratings of experiential strength in 11 sensorimotor dimensions (six perceptual modalities and five action effectors). Since these 11 sensorimotor dimensions are highly correlated with each other (Lynott et al., 2020) and with many lexical and lexico-semantic variables such as frequency, length, and age of acquisition (Dymarska et al., 2023a; Lynott & Connell, 2013), we opted not to analyze the individual dimensions as predictors due to the risk of multicollinearity and suppression effects that can make individual contributions uninterpretable. Rather, we used the principal component analysis (PCA) of Dymarska et al. (2023a) that collapsed this large number of variables into the most important components of lexical and sensorimotor information that, critically, were uncorrelated with one another. Specifically, four distinct sensorimotor components emerged that related to the experience of the body, communication, food, and objects, which we used as critical predictors of memory performance in the present study.

Based on semantic richness theory (Buchanan et al., 2001; Pexman et al., 2008), we predicted that stronger sensorimotor experience, regardless of component, would improve performance on a word recognition memory task, although we were neutral as to whether their effects would be of comparable size. Specifically, following patterns of effects previously observed for imageability (Cortese et al., 2010, 2015; Lau et al., 2018), BOI (Sidhu & Pexman, 2016), and perceptual strength (Khanna & Cortese, 2021) in similar tasks, we expected stronger sensorimotor experience in the referent concept of a target word would make its representation more distinctive and therefore lead to higher hit rates, lower false alarm rates, better overall memory performance (i.e., higher overall hit rate minus false alarms), and better discrimination of old and new items independent of response bias (i.e., higher d'). We had no specific predictions regarding semantic richness effects on response bias but report analysis of c for completeness.

Method

All stimuli, data, and code, as well as any additional materials containing full statistical results, are available at https://osf.io/r8fmb/. The study was not preregistered.

Materials

Items comprised a total of 5,305 words that represented the overlap of items for which dependent and predictor variables were available. Dependent measures came from two megastudies of word recognition memory, one focusing on monosyllabic words (Cortese et al., 2010) and one on disyllabic words (Cortese et al., 2015). In both studies, participants were asked to study lists of 50 words at a time for a later recognition task, which took the form of an old/new judgment on each target word presented (i.e., half the targets were old and half new). These data provided five measures of memory performance per word: hit rate (HR: how many items are correctly recognized as previously seen); false alarm rate (FA: how many items are incorrectly recognized as previously seen); hit rate minus false alarm rate (HR-FA: a common composite measure of word memory performance); d' (sensitivity: how well are old items distinguished from new); and c (criterion or response bias: how strong is the overall tendency to respond “old” or “new” to all items). Of the original 5,577 words with recognition memory data from Cortese et al. (2010, 2015), 5,305 words had predictors available for the current analysis; however, d' and c values were missing for four disyllabic words (Cortese et al., 2015), so the analysis of these DVs includes only 5,301 words.

As predictor variables, we used six PCA components previously obtained in another study to consolidate lexical and sensorimotor predictors[1]; full details can be found in Dymarska et al. (2023a) but we summarize the method of extracting the components here for the benefit of the reader. The item set for the PCA was based on 9,796 words used in the analysis of imageability on word reading by Dymarska et al. (2023a). Variables used for the PCA are detailed in Table 1 and included a variety of sublexical (e.g., orthographic and phonological neighborhoods), lexical (e.g., word length and frequency measures), and lexico-semantic (e.g., age of acquisition, linguistic distributional distance) properties that impact on word processing. In addition, the PCA incorporated 11 dimensions of sensorimotor strength from the Lancaster Sensorimotor Norms (Lynott et al., 2020), where each dimension contained a rating of the extent to which the word’s referent was experienced with the specified perceptual modality or by performing an action with the specified action effector, as well as Lynott et al.’s composite measure of all 11 dimensions, Minkowski-3 sensorimotor strength, which was weighted toward the dominant dimension(s). PCA (parallel analysis at 95th percentile, correlation matrix, varimax rotation) reduced the original 24 dimensions to an optimal six orthogonal components that captured 77.4% of the original variance: two components representing lexical characteristics of the word (Frequency and Length) and four components representing sensorimotor experience of the referent concept (Body, Object, Food, Communication). These components were uncorrelated

---

[1] We also considered an alternative analysis where the PCA included attribute ambiguity, that is, standard deviation of ratings as per Brainerd et al. (2022), which were available for all sensorimotor dimensions and age of acquisition. This analysis produced eight components, with similar loading patterns to our original six components. However, on analysing recognition memory performance, we found that models with these new eight components offered a *worse* fit for three out of five DVs than our original models with six components (i.e., Bayesian evidence strongly favored the six-component models). We therefore opted to retain the original six components as our predictors of interest. The alternative analysis with attribute ambiguity is available in additional materials.

Table 1

Variables Used by Dymarska et al. (2023a) in Principal Component Analysis and the Rotated Components (Used as Study 1 Predictors) to Which They Most Strongly Contributed With Positive or Negative Weighting ($r > .3$ *or* $< -.3$)

Original variable	Source	Definition	Component
LgSUBTLWF	ELP	Log word frequency (U.S. English)	+Frequency
LgSUBTLCD	ELP	Log contextual diversity (how many contexts a word appears in; U.S. English)	+Frequency
Zipf frequency	Van Heuven et al. (2014)	Word frequency on Zipf scale (U.K. English)	+Frequency
Prevalence	Brysbaert et al. (2018)	How many people know the word (probit value)	+Frequency
Familiarity	Stadthagen-Gonzalez and Davis (2006), Scott et al. (2018) and Wilson (1988)	How subjectively familiar a word seems (ratings)	+Frequency
Age of acquisition	Kuperman et al. (2012) [a]	Approximate age that the word was learned	−Frequency
Linguistic distributional distance (LDD20)	Dymarska et al. (2023a, 2023b)	Distributional neighborhood (mean cosine distance to closest 20 neighbors, based on vectors of log co-occurrence frequency)	−Frequency
Word length	ELP	Word length in letters	+Length
Number of syllables	ELP	Word length in syllables	+Length
Orthographic Levenshtein distance (OLD20)	ELP	Orthographic neighborhood (mean letter Levenshtein distance to closest 20 neighbors)	+Length
Phonological Levenshtein distance (PLD20)	ELP	Phonological neighborhood (mean phoneme Levenshtein distance to closest 20 neighbors)	+Length
Torso action strength	LSN	Motor strength in torso effector	+Body
Foot/leg action strength	LSN	Motor strength in foot/leg effector	+Body
Hand/arm action strength	LSN	Motor strength in hand/arm effector	+Body, +Object
Composite sensorimotor strength	LSN	Aggregated sensorimotor strength in all dimensions (Minkowski-3 distance of 11-dimension vector from the origin)	+Body, +Object, +Communication, +Food
Head action strength	LSN	Motor strength in head effector	+Communication
Auditory strength	LSN	Perceptual strength in hearing modality	+Communication
Mouth action strength	LSN	Motor strength in mouth effector	+Communication, +Food
Gustatory strength	LSN	Perceptual strength in taste modality	+Food
Olfactory strength	LSN	Perceptual strength in smell modality	+Food
Visual strength	LSN	Perceptual strength in sight modality	+Object
Noun (part of speech)	ELP	Whether or not word is a noun (binary coded: noun = 1, nonnoun = 0)	+Object
Haptic strength	LSN	Perceptual strength in touch modality	+Object, +Body, −Communication
Interoceptive strength	LSN	Perceptual strength in interoceptive (sensations inside the body) modality	−Object, +Body, +Communication

Note. ELP = English Lexicon project (Balota et al., 2007); LSN = Lancaster Sensorimotor Norms (Lynott et al., 2020).

[a] With extended norms from http://crr.ugent.be/archives/806.

and cleanly distinguished between lexical and semantic information, with the exception of the Object component, which included the noun (part of speech) variable in addition to sensorimotor variables (i.e., since object concepts are typically labeled with nouns and tend to be strongly experienced with visual, haptic, and hand/arm action). Table 1 summarizes how each component relates to the original variables in the PCA, and Table 2 shows a sample of the highest- and lowest-scoring words in each component. All components were centered before the analysis.

Design and Analysis

To investigate the extent to which sensorimotor information contributed to word recognition memory, we conducted item-level hierarchical linear regression analyses of the five dependent measures of memory performance: HR, FA, HR-FA, d', and c. Step 1 entered the two lexical components (Frequency, Length) as baseline model predictors, then Step 2 entered the four sensorimotor components (Body, Communication, Food, Object), and Step 3 entered their

Table 2

Top Five (Highest Scoring) and Bottom Five (Lowest Scoring) Words for Each Component in Studies 1 and 2

Component	High scoring	Lowest scoring
Frequency	The, that, and, what, about	Slat, adage, welt, jeer, vise
Length	Friendship, transplant, somewhere, Privilege, threshold	Rap, sang, pun, gab, hum
Body	Move, movement, bathe, strength, pain	Because, about, but, than the
Food	Meal, pizza, pastry, omelet, pasta	Waltz, listen, chase, polka, ballet
Object	Nail, dog, pillow, pistol, cat	Quench, queasy, hungry, nauseous, digest
Communication	Song, concert, joke, word, chat	Dorsal, fertile, which, than, enzyme

interactions with the Frequency component. We included these interactions in order to fully quantify sensorimotor effect sizes because semantic effects in word reading are typically larger for low-frequency words than high-frequency words (e.g., Connell & Lynott, 2016b; Dymarska et al., 2023a; James, 1975). Such interactions have not been studied extensively in memory research, but there is some evidence that the effect of concreteness on word recall is larger for low-frequency words (Miller & Roodenrys, 2009). We therefore considered that similar patterns were likely to appear in the present study.

We ran Bayesian linear regressions in JASP (0.14.1; JASP Team, 2020) with default JZS priors ($r = .354$) on fixed-effect model parameters, and Bernoulli ($p = .5$) model priors, from which we report Bayes factors (BFs) for model comparisons between hierarchical steps and inclusion BFs of coefficients (i.e., relative likelihood of models including a particular predictor compared to models excluding it). Threshold for inference was $\text{BF}_{\text{inclusion}} = 3.00$ or its reciprocal 0.33. In addition, to calculate part (semipartial) correlation coefficients for each predictor (i.e., the unique contribution each predictor makes to the dependent measure in question), we ran null hypothesis significance testing (NHST) linear regression analyses using the same structure as the Bayesian linear regression.

Results

Overall, performance on the memory task was good, with high hit rates, low false alarms, and low response bias (see Table 3 for descriptive statistics). Lexical effects at Step 1 were largely consistent with previous research: lower-frequency words produced higher hit rates and HR-FA and better d' sensitivity (as in Cortese et al., 2010, 2015; Higham et al., 2009; Lau et al., 2018). Lower-frequency words also led to higher false alarms and a more conservative response bias (i.e., tendency to judge target words as new), similar to Cortese et al. (2010, 2015), but unlike Higham et al. (2009) and Lau et al. (2018). Word length produced small effects on word memory performance, but the pattern of results indicated that shorter words elicited lower HR and FA, and higher HR-FA, with no effect on d' or c. Full statistics are available in additional materials on the Open Science Framework (OSF) page.

Sensorimotor Components

Sensorimotor components at Step 2 strongly improved model fit for all measures of word memory performance (see Table 4). They contributed up to 6.8% of variance in word recognition memory performance (HR-FA), or 5.9% of variance in d' sensitivity, which was greater than that previously reported by Khanna and Cortese (2021) for two variables of composite perceptual strength and composite action strength (i.e., 3.5% and 3.1% for HR-FA and d', respectively). However, contrary to the predictions of semantic richness theory, the four kinds of sensorimotor experience affected memory in different ways and therefore we will report their effects separately. Figure 1 illustrates their directional effects as part correlations (i.e., effect size representing unique contribution); full statistics are available in additional materials on the OSF page.

Table 3
Mean Performance on Each Memory Measure in Study 1 With Its Standard Deviation

Dependent variable	*M*	*SD*
Hit rate	0.728	0.097
False alarms	0.202	0.096
HR-FA	0.526	0.132
d'	1.525	0.472
c	0.126	0.243

Note. HR-FA = hit rate minus false alarm rate.

Body. Sensorimotor experience relating to the Body (i.e., involving motor action of the torso, feet/legs, and hand/arms, plus touch and interoceptive experience) elicited complex effects on word memory. Words scoring higher on the Body component did not clearly influence hit rates at Step 2 ($\text{BF}_{\text{inclusion}} = 1.11$, equivocal evidence) but notably elicited *higher* false alarms, meaning that they led to a large liberal response bias (negative c). In other words, participants were more likely to think that a word was "old" (i.e., previously seen) if its referent was strongly grounded in Body experience, even when they had not actually seen the word in the study list. Higher Body scores overall had a negative effect on composite performance measures of word memory (negative HR-FA and d' sensitivity). That is, contrary to expectations, high Body strength hindered, rather than helped, performance on word recognition memory.

Communication. There was no evidence that the Communication component elicited any effects on word memory performance (all $\text{BF}_{\text{inclusion}} < 1$). That is, words relating to Communication experience (i.e., sound and interoceptive experience, as well as mouth and head action), were not easier to remember, nor more likely to be accurately identified as new, nor prone to any particular response bias.

Food. In contrast to Body effects, sensorimotor experience relating to Food (i.e., involving taste, smell, and mouth action) had more consistent effects on word memory, and generally followed our predictions regarding semantic richness. Words scoring higher in the Food component had higher hit rates and lower false alarms, and therefore better HR-FA and d' sensitivity. There was no effect on bias (c).

Object. The Object component (i.e., sensorimotor experience relating to manipulable objects, namely vision, touch, hand/arm movements, and whether a word was a noun) had the strongest effects on word memory. Words rated higher on the Object component had higher hit rates and lower false alarms, which led to a small liberal response bias—participants had a tendency to judge words as previously seen when they were grounded in the experience of Objects, regardless of whether they had actually been presented in the study list. However, overall the Object component had a strong positive effect on composite performance measures of word memory (positive HR-FA and d' sensitivity), meaning that high Object strength facilitated performance on word recognition memory. These findings resemble conventional effects of semantic richness and were in line with our predictions.

Sensorimotor Components × Frequency

The simultaneous addition of the four interaction terms at Step 3 did not improve model fit overall (see Table 4). However, the inclusion BFs for the interaction terms indicated positive evidence for at least some sensorimotor interactions with Frequency (see Figure 1), indicating that the sensorimotor components differed in how Frequency moderated their effects on recognition memory.

Body. There was strong evidence that Body interacted with Frequency for hit rate performance, where words scoring strongly on the Body component led to higher hit rates but the effect was

Table 4

Percentage of Variance in Memory Performance Explained by Each Step of the Study 1 Regression Models (Change in R^2*, With Levels of Bayesian Evidence), and Uniquely Explained by Each Sensorimotor Component in the Step 2 Model and Each Component Plus Its Interaction With Frequency in the Step 3 Model (Squared Part Correlations)*

Model/parameter	HR	FA	HR-FA	d'	c
Step 1: Lexical baseline R^2	26.10 ***	0.30 *	11.80 ***	8.20 ***	11.80 ***
Step 2: Sensorimotor ΔR^2	4.73 ***	2.96 ***	6.80 ***	5.91 ***	1.14 ***
Body	0.08	1.00	0.27	0.36	0.86
Communication	0.06	0.01	0.07	0.06	0.02
Food	1.17	1.10	2.43	2.37	0.01
Objects	3.31	0.77	3.88	2.99	0.20
Step 3: Sensorimotor × Frequency ΔR^2	0.26	0.04	0.13	0.13	0.16
Body	0.17	0.88	0.14	0.21	0.98
Body × Frequency	0.21	0.00	0.08	0.05	0.14
Communication	0.07	0.01	0.07	0.06	0.01
Communication × Frequency	0.04	0.01	0.05	0.04	0.00
Food	1.10	1.02	2.28	2.28	0.01
Food × Frequency	0.00	0.00	0.00	0.01	0.00
Object	2.62	0.74	3.31	2.72	0.14
Objects × Frequency	0.02	0.03	0.00	0.03	0.02
Total R^2	31.10	3.30	18.70	14.20	13.10

Note. HR = hit rate; FA = false alarm rate; HR-FA = hit rate minus false alarm rate; BF = Bayes factor.
* $\text{BF}_{10} \geq 3$, positive evidence. ** $\text{BF}_{10} \geq 20$, strong evidence. *** $\text{BF}_{10} \geq 150$, very strong evidence.

attenuated for higher-Frequency words. That is, the apparently equivocal effect of Body on HR at Step 2 was superseded by the Step 3 evidence in favor of Body effects that varied with Frequency. There was also evidence for the interaction on response bias (*c*), where the overall liberal bias induced by stronger Body scores was weaker at higher Frequency, but there were no further interaction effects on false alarms, HR-FA, nor d' sensitivity.

Communication. The Communication component did not interact with Frequency, nor did it elicit any effects on word memory performance when the interaction with Frequency was analyzed (consistent with Step 2; all $\text{BF}_{\text{inclusion}} < 0.40$).

Food. There was no interaction of the Food component with word frequency (all $\text{BF}_{\text{inclusion}} < 0.33$), where the pattern of effects remained as per Step 2.

Object. There was no positive evidence for any interactions between Object component scores and Frequency (all $\text{BF}_{\text{inclusion}} < 0.33$), where effects on word memory remained as per Step 2.

Cross-Validation

As a final check to ensure that the estimates of effect size were not overfitted, we conducted *k*-fold cross-validation[2] (10 folds, repeated 200 times) of the above linear regression analyses for all DVs (using the Caret package in R; Kuhn, 2022). The cross-validated models showed very small differences in overall fit (within ±.002 of the original R^2 value), with identical coefficients at Step 3 for cross-validated and original models. We therefore concluded that our original regression models were appropriately fitted, and the contribution of each predictor to each DV was appropriately estimated. The analysis code and results of this cross-validation are available in additional materials on the OSF page.

Discussion

The present study showed that different forms of sensorimotor experience contributed independently to performance in word recognition memory, but often in unexpected ways. Some sensorimotor components (i.e., Object, Food) facilitated memory performance as expected, where stronger experience increased the likelihood of a studied word being correctly remembered (i.e., higher hit rate) and decreased the likelihood of a new word being mistaken for a study item (i.e., lower false alarms). These Object and Food effects were consistent with the general pattern of semantic richness effects (Sidhu & Pexman, 2016; Snefjella & Kuperman, 2016) and with our predictions regarding the influence of perceptual and action experience on word memory: stronger sensorimotor experience of a word's referent allows it to act as an effective retrieval cue to the memory trace of the studied items. Sensorimotor experience relating to Communication had no effect on any aspect of word recognition memory, which did not follow our predictions but was nonetheless consistent with other null effects observed for semantic richness variables in word recognition memory (e.g., body–object interaction and affective valence in Lau et al., 2018). It is possible that words scoring high in Communication (e.g., *chat*, *joke*) lacked the same degree of distinctiveness as those scoring high in the Object (e.g., *pillow*, *dog*) or Food (e.g., *meal*, *pasta*) components, which therefore quashed any facilitatory effects on word memory; we return to this point in the General Discussion section.

Critically, however, the Body component produced effects in the opposite direction to predictions by increasing false alarms and impairing overall HR-FA and sensitivity to the difference between old and new items (d').[3] In particular, high Body scores led

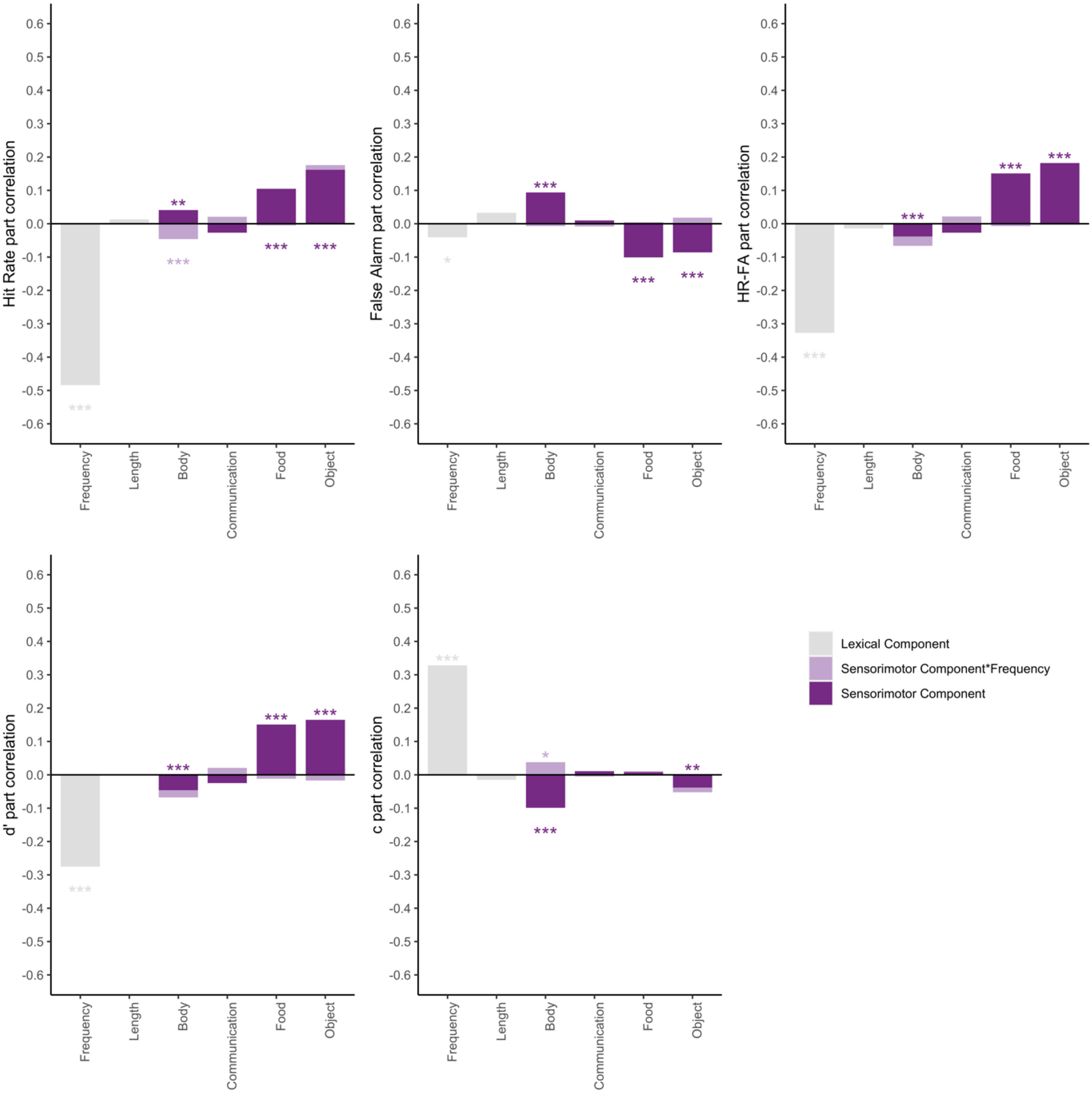
---

[2] We thank Marc Brysbaert for this suggestion.

[3] We note that Khanna and Cortese (2021) also reported that a composite variable of action strength unexpectedly decreased HR–FA and d' (they do not analyze HR and FA separately). However, this effect is difficult to interpret because the action strength variable was correlated with other, stronger predictors (including composite perceptual strength and word frequency) and therefore may be prone to suppression effects in regression. The orthogonal sensorimotor components we use in the present study do not share this problem.

Figure 1

Part Correlations of Each Component Predictor in Study 1



Note. Stacked bars represent the combined effect of each sensorimotor component (darker shade) and the respective Component × Frequency interaction (lighter shade). Asterisks indicate the inclusion Bayes factor (BF) of each predictor: *** $BF_{inclusion} \geq 150$, constituting very strong evidence; ** $BF_{inclusion} \geq 20$, strong evidence; * $BF_{inclusion} \geq 3$, positive evidence. HR-FA = hit rate minus false alarm rate. See the online article for the color version of this figure.

participants to mistakenly judge words that had *not* been studied as "old" and did so to a greater extent than they enhanced recognizing words that *had* in fact been presented in the study phase (i.e., negative HR-FA). No such behavior was observed for other components. In other words, experiencing a referent concept with the Body gave a newly encountered word an illusion of being a studied word with a strong memory trace, regardless of whether or not it actually had been studied, to the point of creating a liberal response bias (i.e., negative *c*). This pattern of effects could not be explained by lack of distinctiveness (see Lau et al., 2018), because words scoring highly on the Body component, referring to a specific bodily experience such as *cuddle* or *fitness*, are likely to be distinct in their representations and should not be easily confused with one another. An alternative possibility is that Body-related words uniquely attract and direct attention in a way that influences their memorability. For instance, some researchers have argued that there is an adaptive advantage in attending to stimuli that are relevant to survival and allowing them to spread activation to other survival-related knowledge, which leads to a stronger, more distinctive memory trace for studied items but also increases false alarms (Bonin et al., 2014; Howe & Derbish, 2010; Leding, 2020). Other work has shown that directing attention to interoception—a perceptual modality that loads strongly on the Body component—causes mistaken sensations of touch (Mirams et al., 2012; see also Mirams et al., 2013), which could make the memory trace less distinctive and prone to false alarms. We return to these possible explanations in the General Discussion section.

The finding that different aspects of perceptual and motor experience affect memory in different ways is an important one, because it does not support the broader theoretical position that semantic richness of word meaning facilitates word recognition memory (e.g., Madan, 2020; Sidhu & Pexman, 2016; Yap et al., 2011) so long as the particular semantic richness variable does not reflect lower distinctiveness (Lau et al., 2018). Our findings instead suggest that sensorimotor grounding of word meaning could either facilitate or inhibit recognition memory, depending on the type of experience it represents.

Study 2: Conscious Imagery

Imageability is a theoretically distinct construct from sensorimotor grounding, as it is specifically concerned with the conscious generation of mental imagery for a word rather than the automatic and unconscious activation of sensorimotor information upon reading a word (Connell & Lynott, 2016a; Dymarska et al., 2023a). It should thereby enhance the distinctiveness of the memory trace independently of different forms of sensorimotor experience and produce the typical pattern of semantic richness effects on word recognition memory by increasing hit rates while reducing false alarms (i.e., overall facilitating HR-FA and d'). Many studies have indeed found such effects (e.g., Cortese et al., 2010, 2015; Groninger, 1974; Khanna & Cortese, 2021; Paivio & Csapo, 1969). However, higher imageability sometimes unexpectedly increases false alarms instead of reducing them (Ballot et al., 2021; Peterson & McGee, 1974), and recent work on word reading has cast doubt on the construct validity of imageability by showing that it elicits heterogenous, unreliable effects according to the source of norms employed (Dymarska et al., 2023a). The goal of the present study was therefore to disentangle the contributions of imageability and sensorimotor grounding (in various forms) on recognition memory performance and to determine whether imageability has sufficient construct validity to produce reliable semantic richness effects on word recognition memory.

If imageability has poor construct validity (i.e., cannot properly measure ease of generating mental imagery), then imageability ratings from different sources would fail to produce reliable semantic richness effects on word recognition memory above and beyond sensorimotor grounding. Conversely, if imageability itself has good construct validity but simply does not usefully contribute to word reading (as per Dymarska et al., 2023a), then we would still expect imageability ratings from different sources to produce reliable semantic richness effects on recognition memory by increasing hit rates, reducing false alarms, and facilitating overall memory performance (i.e., higher HR-FA and d'). We had no predictions regarding response bias (c) but, as in Study 1, we report its analysis for completeness. Previously, imageability ratings examined in word recognition memory studies (e.g., Khanna & Cortese, 2021; Lau et al., 2018; Sidhu & Pexman, 2016) were primarily taken from three main sources: Chiarello et al. (1999); Cortese and colleagues (Cortese & Fugett, 2004; Schock et al., 2012); and the MRC database (Coltheart, 1981; Wilson, 1988) that combines a number of earlier norms. Here, we aim to consider a wider range of sources and examine their effects over and above a lexical and sensorimotor baseline model, in order to evaluate the unique contribution of imageability ratings to semantic richness effects on word memory.

Method

All stimuli, data, code, and full results are available at https://osf.io/r8fmb/. The study was not preregistered.

Materials

Items were identical to Study 1. As our predictor of interest in the present study, we collated imageability ratings from six different sets of imageability norms, each of which used the same instructions and scale to collect ratings from participants: the Bird norms (Bird et al., 2001); Bristol norms (Stadthagen-Gonzalez & Davis, 2006); Chiarello norms (Chiarello et al., 1999); Cortese norms (Cortese & Fugett, 2004; Schock et al., 2012)[4]; Glasgow norms (Scott et al., 2018); and the widely used MRC norms (Coltheart, 1981; Wilson, 1988; featuring imageability ratings from Gilhooly & Logie, 1980; Paivio et al., 1968; Toglia & Battig, 1978). Critically, all six sets of norms used the same imageability scale and rating instructions, originating with Paivio et al. (1968). Because each set of norms covered a different sample of words with varying overlap, and because previous work found large differences in predictive ability of different imageability norms (Dymarska et al., 2023a), we analyzed each separately.

Design and Analysis

Due to high correlations between imageability ratings and the individual components (see additional materials) that are likely to produce suppression effects, and due to the automatic activation of lexical and sensorimotor information, which occurs prior to conscious generation of mental imagery (Connell & Lynott, 2016a; Pecher et al., 2009), we opted to analyze the residuals of the regression on lexical and sensorimotor components. We first calculated the residual values for the entire word sample from Study 1, which was covered by ratings from the Cortese norms, by running a linear regression in R (using lm function; R Core Team, 2021; Venables & Ripley, 2002) with all the Study 1 predictors (i.e., lexical and sensorimotor components, as well as the interaction between sensorimotor components and the Frequency component) for each memory

---

[4] Cortese and Fugett (2004) and Schock et al. (2012) norms were combined into a single variable since they came from the same laboratory and were used in the recognition memory studies of Cortese et al. (2010, 2015), respectively, which we analyze as a single dataset (see also Khanna & Cortese, 2021).

DV. We then calculated residuals for the subsets of words covered by the remaining five sets of imageability norms using the same method; see additional materials for full details. These residuals therefore represent variance in word memory performance from which effects of lexical characteristics and sensorimotor grounding had been removed, meaning that any effect of imageability would represent a pure effect of the construct (i.e., ease of generating mental imagery, independent of sensorimotor grounding) that is comparable across different sources of imageability ratings.

We then conducted linear regressions on the residuals from each of the five dependent variables from Study 1: HR, FA, HR-FA, d' (sensitivity), and c (bias). In Step 1, we entered imageability (centered) as a predictor, and in Step 2, we entered the interaction of Imageability × Frequency component. The interaction term was included to ensure we fully quantified imageability effects, following previous findings that concreteness effects on word memory may be stronger for low-frequency words (Miller & Roodenrys, 2009), and was calculated by multiplying each centered imageability variable with the Frequency component from Study 1 (we did not enter Frequency as a parameter because it was already partially out in the regression that generated the residuals). There were 30 regression models in total, representing six different imageability norms (Bird, Bristol, Chiarello, Cortese, Glasgow, MRC) by the residuals of five different dependent variables (HR, FA, HR-FA, d', c). We ran Bayesian regressions as per Study 1 from which we report BFs for model comparisons between hierarchical steps and inclusion BFs of coefficients. We also ran NHST regressions with the same parameters to obtain part correlation coefficients. We report effect sizes as the variance of the original DV, which we calculated by scaling each R^2-change and squared part correlation by the size of the relevant residual variance [scaled $\Delta R^2 = \Delta R^2 \times (1 - \text{original } R^2)$], which allows comparison with the results of Study 1.

Results

Performance on the memory task was reasonably consistent across the subsets of words covered by each set of norms, with high hit rates and overall memory performance measures (HR-FA and d'), and low false alarms and bias (see Table 5).

Imageability

Overall, imageability effects on word recognition memory were relatively consistent across different sets of norms when the variance associated with lexical and sensorimotor characteristics had already been removed (see Figure 2 and Table 6). Higher imageability predicted higher HR, HR-FA, and d' sensitivity in all analyses at Step 1. That is, in line with semantic richness theory, regardless of which set of imageability norms provided ratings, words high in imageability were more likely to be correctly recognized as "old" when they had featured in the study list (increased HR), and were better differentiated as old versus new (i.e., higher HR-FA and d').

The effects on FA were less consistent, however, where imageability reduced FA on four out of six analyses (Bristol, Cortese, Glasgow, MRC) but had no effect on the remaining two (equivocal evidence for Bird and Chiarello norms). That is, words rated higher in imageability were often—but not consistently—less likely to be mistaken for "old" when they were actually new (unseen) words. Finally, the measure of response bias (c) was least affected by imageability ratings. Higher imageability led to a more liberal response bias for the Cortese norms only, such that high imageability words were more likely to be considered as "old," regardless of whether they had appeared in the study list. There was no further evidence in favor of imageability effects on the bias from the other norms.

Imageability × Frequency

Imageability effects varied little with Frequency at Step 2. There was no positive evidence for the interaction in the analysis of HR or FA, nor—with a single exception each—in the analysis of HR-FA, d', or c (see additional materials for full details). The exception in two cases was the Chiarello norms, which produced a positive interaction on both HR-FA and d': that is, the facilitatory effect of imageability was larger for higher-frequency words than for lower-frequency words. The final exception was the Cortese norms, which had a positive interaction effect on c, where the liberal response bias induced by imageability ratings from the Cortese norms was attenuated for high-Frequency words.

Discussion

Imageability norms overall elicited a facilitation effect on word recognition memory, in line with semantic richness theory. When lexical and sensorimotor variance was accounted for, ease of consciously generating mental imagery played an independent role in word recognition memory. While the direction of results was consistent with the findings of Cortese et al. (2010, 2015; see also Khanna & Cortese, 2021), in that higher imageability facilitated hit rate and at least sometimes led to lower false alarms, the *magnitude* of the effects differed markedly. For example, in Cortese et al. (2010, 2015), imageability predicted 14%–24% of variance in HR when lexical variables were accounted for. In the present study, using the same Cortese imageability norms and memory dataset but a different baseline model that accounted for both lexical *and* sensorimotor variables, imageability (including its Frequency interaction) explained 3.6% variance in HR, and even less variance when most of the other imageability norms were used. Notably, the present effects were also much smaller than those found by Khanna and Cortese (2021), who also used the Cortese norms alongside composite measures of perceptual and action strength (i.e., where individual perceptual modalities were aggregated into a single weighted score, and likewise for individual action effectors). They found that imageability predicted 11% of variance in HR-FA and 13% of variance in d', above and beyond lexico-semantic baseline variables and composite perceptual strength and action strength. Again, using the same norms and memory dataset but a baseline model that included different forms of sensorimotor experience, we found that imageability overall explained only 4.3% and 3.5% of variance in HR-FA and d', respectively. Such differences in the magnitude of imageability effects are likely due to us employing a comprehensive lexico-semantic-sensorimotor baseline across all analyses and isolating pure imageability from this baseline by residuals analysis, which suggests that a large part of the imageability effect size reported in the word memory literature is actually due to sensorimotor grounding of word meaning rather than a pure effect of the imageability construct itself (i.e., ease of generating mental imagery).

Table 5

Mean Performance on Each Memory Measure (With Standard Deviations in Parentheses) for the Words Covered by Each Set of Imageability Norms in Study 2

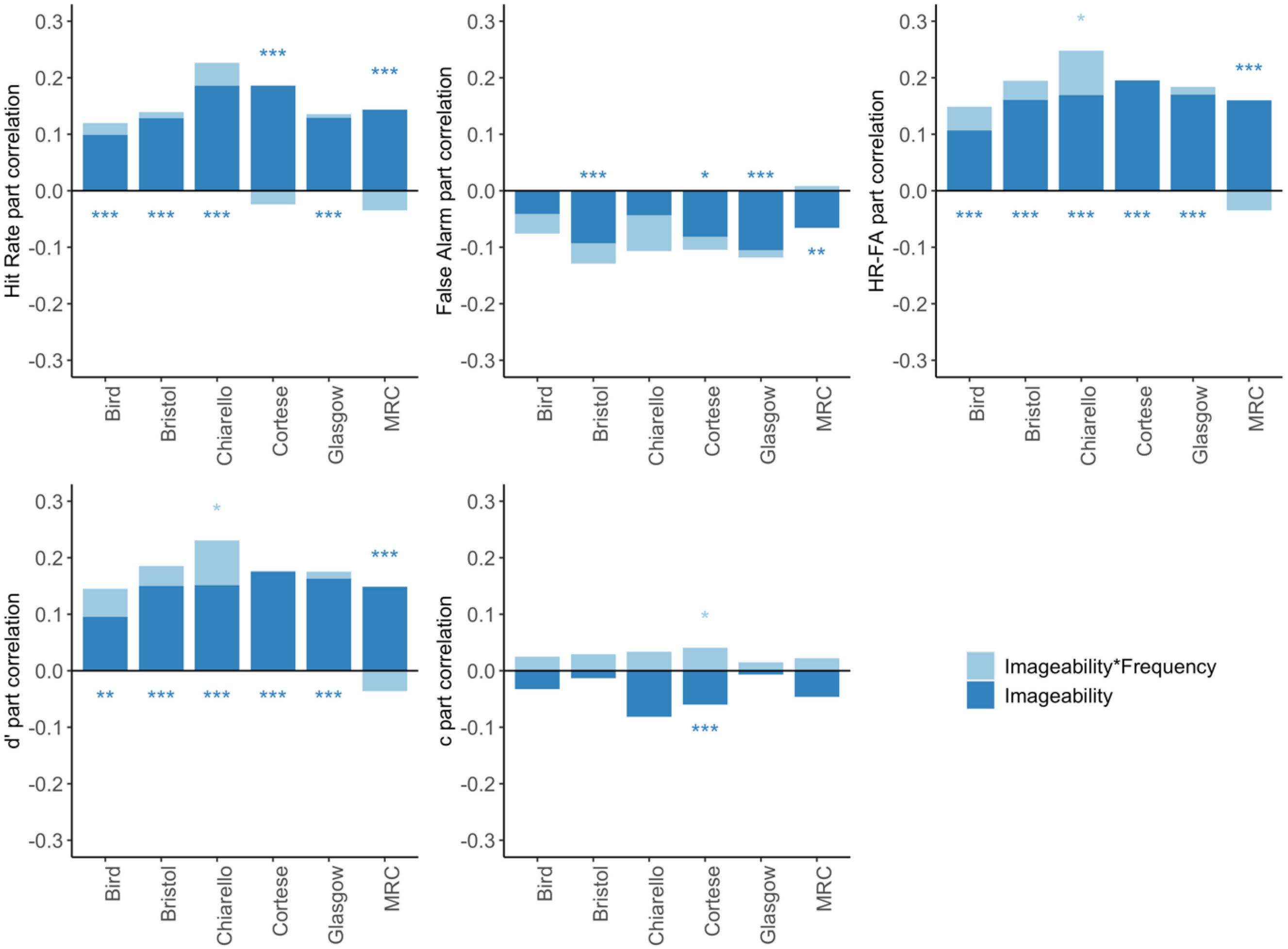
Dependent variable	Bird	Bristol	Chiarello	Cortese	Glasgow	MRC
Hit rate	0.687 (0.102)	0.727 (0.088)	0.723 (0.092)	0.729 (0.097)	0.722 (0.094)	0.712 (0.099)
False alarms	0.212 (0.098)	0.223 (0.096)	0.218 (0.094)	0.202 (0.096)	0.207 (0.094)	0.198 (0.091)
HR-FA	0.475 (0.133)	0.505 (0.126)	0.505 (0.124)	0.526 (0.132)	0.515 (0.130)	0.515 (0.126)
d′	1.358 (0.457)	1.438 (0.442)	1.438 (0.419)	1.525 (0.472)	1.481 (0.459)	1.487 (0.444)
c	0.169 (0.247)	0.090 (0.228)	0.104 (0.235)	0.126 (0.243)	0.127 (0.235)	0.158 (0.244)
N	810/808	1,223	989	5,305/5,301	2,593	2,563

Note. The Bird norms and the Cortese norms show lower *N* for *d′* and *c*, due to missing values in the memory dataset, as indicated in Study 1. HR-FA = hit rate minus false alarm rate.

Nonetheless, imageability was reasonably consistent in facilitating recognition memory regardless of the ratings source, as all sets of imageability norms led to increased HR, HR-FA, and *d′*. The effect on FA was more unstable, but where it appeared (for four out of six norms) it was in the predicted direction and reduced FA. These findings show that imageability impacts on word memorability above and beyond sensorimotor grounding, which suggests that imageability ratings do indeed measure the ease of consciously

Figure 2

Part Correlations of Imageability Effect on Memory Performance Residuals in Study 2



Note. Stacked bars represent the combined effect of imageability (darker shade) and the Imageability × Frequency interaction (lighter shade) in the regression model. Effect sizes are scaled to reflect variance of the original DV rather than variance of the residuals only. Asterisks indicate the inclusion Bayes factor (BF) of each predictor: *** $\mathrm{BF}_{\mathrm{inclusion}} \geq 150$, constitutes very strong evidence; ** $\mathrm{BF}_{\mathrm{inclusion}} \geq 20$, strong evidence; * $\mathrm{BF}_{\mathrm{inclusion}} \geq 3$, positive evidence. HR-FA = hit rate minus false alarm rate. See the online article for the color version of this figure.

Table 6

Percentage Variance in Memory Performance Explained by Each Step of the Study 2 Regression Models (R^2*, With Levels of Bayesian Evidence), and Uniquely Explained by Imageability at Step 1 and Imageability Plus Its Interaction With Frequency at Step 2 (Squared Part Correlations), for Each Set of Imageability Norms*

Norms	Model/parameter	HR	FA	HR-FA	d'	c
Bird	Step 1: Imageability R^2	1.94 ***	0.61	2.69 ***	2.46 ***	0.05
	Imageability parameter (sr^2)	1.93	0.61	2.68	2.46	0.05
	Step 2: Imageability × Frequency ΔR^2	0.04	0.12	0.17	0.25	0.06
	Imageability parameter (sr^2)	0.97	0.17	1.15	0.92	0.10
	Imageability × Frequency (sr^2)	0.04	0.12	0.17	0.24	0.06
Bristol	Step 1: Imageability R^2	2.51 ***	1.78 ***	4.49 ***	4.01 ***	0.00
	Imageability parameter (sr^2)	2.52	1.77	4.47	4.01	0.00
	Step 2: Imageability × Frequency ΔR^2	0.01	0.13	0.11	0.12	0.09
	Imageability parameter (sr^2)	1.65	0.87	2.58	2.25	0.02
	Imageability × Frequency (sr^2)	0.01	0.13	0.12	0.13	0.09
Chiarello	Step 1: Imageability R^2	4.38 ***	0.45	4.19 ***	3.47 ***	0.57
	Imageability parameter (sr^2)	4.39	0.45	4.21	3.48	0.57
	Step 2: Imageability × Frequency ΔR^2	0.16	0.40	0.60*	0.63*	0.11
	Imageability parameter (sr^2)	3.45	0.19	2.88	2.30	0.67
	Imageability × Frequency (sr^2)	0.16	0.39	0.61	0.62	0.11
Cortese	Step 1: Imageability R^2	3.57 ***	0.89 ***	4.28 ***	3.47 ***	0.24 **
	Imageability parameter (sr^2)	3.58	0.89	4.26	3.47	0.24
	Step 2: Imageability × Frequency ΔR^2	0.06	0.05 ***	0.00	0.00	0.17 *
	Imageability parameter (sr^2)	3.46	0.67	3.83	3.06	0.36
	Imageability × Frequency (sr^2)	0.06	0.05	0.00	0.00	0.17
Glasgow	Step 1: Imageability R^2	2.61 ***	1.86 ***	4.69 ***	4.26 ***	0.00
	Imageability parameter (sr^2)	2.61	1.86	4.68	4.27	0.00
	Step 2: Imageability × Frequency ΔR^2	0.00	0.02	0.02	0.01	0.02
	Imageability parameter (sr^2)	1.66	1.11	2.90	2.67	0.00
	Imageability × Frequency (sr^2)	0.00	0.02	0.02	0.01	0.02
MRC	Step 1: Imageability R^2	2.66 ***	0.64 ***	3.43 ***	2.82 ***	0.17
	Imageability parameter (sr^2)	2.65	0.64	3.44	2.82	0.17
	Step 2: Imageability × Frequency ΔR^2	0.12	0.01	0.12	0.13	0.05
	Imageability parameter (sr^2)	2.07	0.43	2.57	2.21	0.21
	Imageability × Frequency (sr^2)	0.12	0.01	0.12	0.13	0.05

Note. The summed sr^2 for imageability and Imageability × Frequency terms in the final model can be greater than the corresponding imageability R^2 due to mutual suppression of parameters. Effect sizes are scaled to reflect variance of the original DV rather than variance of the residuals only. HR = hit rate; FA = false alarm rate; HR-FA = hit rate minus false alarm rate; BF = Bayes factor.

*** $BF_{10} \geq 150$ very strong evidence. ** $BF_{10} \geq 20$ strong evidence. * $BF_{10} \geq 3$ positive evidence. No symbol, $BF_{10} < 3$, evidence against inclusion.

generating mental imagery for a word as well as reflecting sensorimotor (primarily visual) information, and for this reason contribute independently to semantic richness effects on word recognition memory. That is, even though participants from different imageability norming studies vary enormously in the strategies and information they use to rate imageability, despite receiving the same instructions and rating scales (Dymarska et al., 2023a), they still manage to encode information in their imageability ratings that is functionally useful to word recognition memory. The present results suggest that the ease of generating mental imagery enhances the distinctiveness of a studied word's memory trace independently of different forms of sensorimotor grounding and hence produces the typical pattern of semantic richness effects on word recognition memory.

General Discussion

The aim of the paper was to examine the effects of various aspects of sensorimotor experience and imageability on word recognition memory within the framework of semantic richness theory (Buchanan et al., 2001; Pexman et al., 2008). We used a novel method of investigating how sensorimotor grounding affects word memory by examining how orthogonal measures of different forms of sensorimotor experience (Body, Communication, Food, Objects) can predict recognition memory performance. We found that different forms of sensorimotor experience vary in their effects on word recognition memory, contrary to the proposal that the words with a richer semantic representation are remembered better when they make representations more distinctive (Study 1). While sensorimotor experience relating to Objects and Food did produce effects consistent with semantic richness (i.e., increased hit rates, reduced false alarms, and overall facilitated memorability HR-FA and d'), experience relating to Communication had no effects on word memory performance. Most strikingly, experience relating to the Body *impaired* memory performance rather than facilitating it, by inflating false alarms, inducing an overly liberal response bias, and overall worsening performance (HR-FA and d'). Nonetheless, we found that imageability (i.e., the ease of generating mental imagery) produced effects consistent with semantic richness independently of sensorimotor grounding (Study 2), although the effect was smaller than suggested in previous literature.

These results reveal unexpectedly complex effects of sensorimotor information on word recognition memory performance and raise questions about the mechanisms behind those effects. The lack of Communication effects could potentially be attributed to the lack of distinctiveness of words that score highly on the Communication component. Many of the words in our study that relate strongly to Communication experience appear to cluster with other words of rather similar meanings that could easily be confused with one other (e.g., *scream*, *yell*, *shout*; *chat*, *talk*, *speak*), making it difficult for participants to confidently distinguish between items which were previously seen and items which were not. Such items, despite being presented in the study phase, are easily confused with similar new items presented in the test phase. Therefore, words from such clusters of similar concepts have poor diagnostic value at retrieval, leading to a lack of discrimination between old and new items. However, other strongly Communication-related words seem relatively distinct in meaning (e.g., *song*, *pun*, *lecture*, *sneeze*), meaning low distinctiveness is not endemic among high Communication scores. Nonetheless, if stronger Communication experience does not systematically increase the distinctiveness of a word's representation, it could explain why semantically richer representations (in terms of Communication experience) do not necessarily facilitate word memory. Future research should examine how different forms of sensorimotor experience may differentially influence the distinctiveness of meaning representation.

The impairment effects of Body experience, however, require a different explanation. Unlike Communication, words scoring highly on the Body component do not appear to lack distinctiveness. Apart from occasional exceptions (e.g., *move*, *movement*; *strength*, *strong*), words involving the body appear to have distinct meanings that are not easily confusable (e.g., *bathe*, *climb*, *dance*, *massage*). Moreover, unlike previous demonstrations of impairment effects on memory (e.g., composite action strength variable in Khanna & Cortese, 2021), in the present study, the use of orthogonal PCA components allowed us to conclude that Body experience is indeed inflating FA (to the point of inducing an overall response bias), and thereby impairing HR-FA and d' in word recognition memory. There are two theoretical accounts regarding attentional mechanisms that may offer some explanation for these unexpected effects of Body experience. First is the adaptive explanation: previous work has demonstrated that stimuli important to survival facilitate hit rates while simultaneously increasing false alarm rates (Bonin et al., 2014; Howe & Derbish, 2010; Leding, 2020), contrary to the mirror pattern (Glanzer & Adams, 1985) that many other semantic variables have on recognition memory (e.g., imageability: Study 2; Cortese et al., 2010, 2015; Lau et al., 2018). According to this account, there is an adaptive advantage in how survival-related words automatically capture attention and spread activation to networks of other interconnected concepts that may increase the chance of survival (Howe & Derbish, 2010). While this process leads to a stronger, more distinctive memory trace for studied words (thereby increasing hit rates), the activation of other related items has the side effect of increasing the likelihood of false alarms. Specifically, spreading activation to representations of other, similar concepts leads to generating a memory trace for items that were not presented in the study phase, which reduces discriminability between the memory trace of the studied word and that of related concepts when a cue word is presented in the test phase. That is, it reduces the diagnostic value of the cue (Goh & Lu, 2012; Nairne, 2002) by making it harder to discriminate between words that were actually studied and words that were activated as related concepts. If sensorimotor experience of the Body is important to survival—and bodily function and integrity are core to an individual's survival—then strongly Body-related words may automatically spread activation to other related words in the study phase, thereby increasing hit rates (at least for low-Frequency words: Study 1) but inflating false alarms as new Body-related words are mistakenly matched to those granted a memory trace via spreading activation.

Second is the somatic attention explanation: previous work in perception has found that directing attention to interoception (i.e., heartbeat; Mirams et al., 2012), to the hand (Mirams et al., 2010), or to locations within peripersonal space (Mirams et al., 2017) all cause increased false alarms on tactile stimulation detection without consistently producing a corresponding increase in hit rates. In other words, directing attention toward the body in various ways leads people to mistakenly believe they are perceiving touch sensations. Several other perceptual phenomena have been found to reappear in the semantic processing of sensorimotor information, including modality switching costs (Pecher et al., 2003) and the tactile disadvantage in stimulus detection (Connell & Lynott, 2010), supporting grounded theories that the conceptual system has co-opted the sensorimotor system for the purposes of representation (e.g., Barsalou, 1999; Connell & Lynott, 2014). If attending to the body in one perceptual modality activates a false sense of touch, then the presence of interoceptive and hand/arm action experience in Body-related words may similarly cause additional, irrelevant tactile activation that renders the total representation less distinctive and thereby prone to false alarms. That is, the meaning representation of a strongly Body-related word might be distinctive in its own right, but its tendency to activate other modalities of Body-related experience (particularly touch) lowers its distinctiveness and leads to a more confusable memory trace and/or cue representation, which reduces the diagnostic value of the cue (Goh & Lu, 2012; Nairne, 2002) and increases false alarms without necessarily affecting hit rates. While the present data cannot distinguish between the adaptive and somatic attentional explanations of Body's impairment effects, future work should seek to determine precisely how Body-related sensorimotor information affects memory for words.

It is important to note that the adaptive advantage and somatic attentional accounts are restricted to the effects of the Body component. While one might reason that interacting with objects could conceivably relate to survival or somatic experience, closer examination shows that the mechanisms that we outlined above as possible explanations for the pattern of results for the Body component are unlikely to extend to the Object component. For instance, although some objects may be related to survival if encountered in the real world (e.g., *pistol*, *tiger*), the vast majority of the words scoring highly on the Object component are everyday object concepts that have nothing to do with survival at all (e.g., *cat*, *pen*, *jar*, *arcade*). Conversely, there are words that score very low on the Object component that seem strongly related to survival (e.g., *virus*, *starve*, *immune*). The lack of a systematic relationship between Object component scores and relevance to survival eliminates the possibility that the adaptive account could extend to the Object component. Additionally, although both Body and Object concepts share some sensorimotor dimensions, the unique loading profiles of each component (see Table 1) mean they reflect experience with different types of concepts. Somatic experience is particularly associated with interoception (i.e., sensations inside the body), which loads positively on the Body component but negatively on the Object component, while other forms of somatic experience (e.g., action of the torso and foot/leg) load positively on the Body component while having a negligible impact on the Object component. As a result, because component rotation during PCA rendered the Object and Body components orthogonal, it means that somatically relevant concepts tend to score highly on the Body component while having no systematic relationship with the Object component. For example, while there are a few concepts that score highly on the Object component *and* may plausibly be the subject of somatic attention (e.g., *finger*, *face*), the vast majority of words scoring highly on the Object component have nothing to do with somatic attention at all (e.g., *cat*, *pen*, *jar*, *arcade*). Conversely, some words with extremely low Object scores are indeed somatic by their relevance to interoception (e.g., *breathe*, *bladder*, *hungry*). That is, due to how the components are constructed, somatic experience is overwhelmingly loaded on the Body component, meaning the somatic account does not extend to the Object component. Hence, the adaptive and somatic accounts both predict impairment effects for the Body component without concomitant effects for the Object component.

We also considered whether pure imageability—that is, the ease of consciously generating mental imagery, as a separate theoretical construct to sensorimotor grounding—played an additional role in word recognition memory, over and above automatically activated sensorimotor information. We found that different sets of imageability norms had relatively consistent facilitatory effects on word memory once lexical and sensorimotor information has been taken into account, in contrast to previous findings for imageability in word reading (Dymarska et al., 2023a). Notably, imageability effect sizes on word memory were much smaller than those previously reported in the literature, when examined on top of multiple, distinct forms of sensorimotor experience. These findings suggest that imageability ratings have adequate construct validity—that is, they do indeed reflect the ease of generating a mental image—and that this construct facilitates word recognition memory independently of sensorimotor information in a pattern consistent with semantic richness. A word can be remembered better if it has rich semantic detail in its representation, because it will have a strong memory trace that can be effectively retrieved when the word is presented again later (e.g., Hargreaves et al., 2012; Lau et al., 2018; Sidhu & Pexman, 2016). Consciously generating imagery can contribute to a stronger, more distinctive memory trace in the study phase, independently of any effects of sensorimotor grounding, and can also contribute to a more distinctive and discriminable retrieval cue in the test phase, thus facilitating memory performance in the classic mirror pattern.

Moreover, the present findings have implications for research on word reading. Dymarska et al. (2023a) found that some sets of imageability norms (e.g., Bird) performed much better than others (e.g., Bristol and Glasgow) in predicting latency and accuracy of lexical decision and word naming tasks, with very weak effects overall, once lexical and sensorimotor information has been taken into account. Since the same patterns did not occur in Study 2 (i.e., all norms consistently facilitated word recognition memory), it means that the construct of imageability itself—the ease of generating a mental image—is simply not a useful predictor of word reading performance. According to the models of word recognition in a lexical decision task (e.g., Coltheart et al.'s, 2001, dual-route cascading model; Harm & Seidenberg's, 2004, triangle model) when a written word is presented, its orthographic representation spreads activation to semantic content, which in turn feeds activation back to the orthographic and/or phonological representation and facilitates a relatively rapid task response. While sensorimotor grounding contributes to this semantic feedback process, at least in part due to attentional modulation of sensory systems during word reading (Connell & Lynott, 2014), the ability to consciously generate mental imagery does not. The present findings also suggest that the best-performing imageability norms for lexical decision and word naming, such as the Bird norms, succeed in predicting word reading for a reason *other than ease of generating mental imagery*. That is, as speculated by Dymarska et al. (2023a), some—but not all—imageability ratings reflect some hitherto unidentified semantic construct that is *not* sensorimotor grounding and *not* the ease of generating mental imagery, but is nonetheless useful to semantic feedback in the process of recognizing words. In summary, even though pure imageability contributes to performance in some tasks such as word recognition memory (Study 2), it does not necessarily extend to other types of conceptual processing due to their different processes and posited roles for semantic information.

In conclusion, the current study provides important theoretical and methodological insights into the study of word recognition memory. First, it shows that semantic richness—where words that are richer in semantic information are remembered better because they elicit stronger semantic activation (Pexman et al., 2013) and therefore a stronger memory trace (Hargreaves et al., 2012; Sidhu & Pexman, 2016), so long as they enhance distinctiveness and discriminability (Lau et al., 2018)—does not fully explain word memory performance. Detailed analysis of sensorimotor information does not conform to this "more is better" view of semantic richness effects on word memory, because while some forms of sensorimotor experience (relating to Objects or Food) do indeed facilitate recognition memory as expected, other forms of sensorimotor experience either have no effect (when relating to Communication) or actually impair rather than facilitate performance (when relating to the Body). Second, although imageability produced much smaller effects than indicated in previous research, it nonetheless consistently facilitated recognition memory performance independently of sensorimotor grounding, suggesting that generating conscious imagery contributes to semantic richness effects when remembering a list of studied words. Together, these findings indicate that semantic richness effects on word memory are more complex than previously suggested and encompass multiple constraints regarding how different forms of sensorimotor experience differentially influence the distinctiveness of the memory trace and/or the effectiveness of the cue. In order to fully understand these complex effects, a multidimensional approach to semantic information is needed when investigating its effects on memory. Word recognition memory is often, but not consistently, facilitated by rich semantic representations, and it is essential to separately consider distinct forms of sensorimotor experience rather than assuming more information is always better.

References

Ballot, C., Mathey, S., & Robert, C. (2021). Word imageability and orthographic neighbourhood effects on memory: A study in free recall and recognition. *Memory*, *29*(6), 829–834. https://doi.org/10.1080/09658211.2021.1921216

Balota, D. A., Yap, M. J., Hutchison, K. A., & Cortese, M. J.. (2012). Megastudies: What do millions (or so) of trials tell us about lexical processing? In J. S. Adelman (Ed.), *Visual word recognition volume 1: Models and methods, orthography and phonology*. Psychology Press.

Balota, D. A., Yap, M. J., Hutchison, K. A., Cortese, M. J., Kessler, B., Loftis, B., 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. (1999). Perceptual symbol systems. *Behavioral and Brain Sciences*, *22*(4), 577–660. https://doi.org/10.1017/S0140525X99002149

Barsalou, L. W., Santos, A., Simmons, W. K., & Wilson, C. D. (2008). Language and simulation in conceptual processing. In M. De Vega, A. M. Glenberg, & A. C. Graesser (Eds.), *Symbols, embodiment, and meaning* (pp. 246–283). Oxford University Press.

Bird, H., Franklin, S., & Howard, D. (2001). Age of acquisition and imageability ratings for a large set of words, including verbs and function words. *Behavior Research Methods, Instruments, & Computers*, *33*(1), 73–79. https://doi.org/10.3758/BF03195349

Bonin, P., Gelin, M., & Bugaiska, A. (2014). Animates are better remembered than inanimates: Further evidence from word and picture stimuli. *Memory & Cognition*, *42*(3), 370–382. https://doi.org/10.3758/s13421-013-0368-8

Bourassa, D. C., & Besner, D. (1994). Beyond the articulatory loop: A semantic contribution to serial order recall of subspan lists. *Psychonomic Bulletin & Review*, *1*(1), 122–125. https://doi.org/10.3758/BF03200768

Brainerd, C. J., Bialer, D. M., & Chang, M. (2022). Memory effects of semantic attributes: A method of controlling attribute contamination. *Behavior Research Methods*, *Advanced online publication*, 1–30. https://doi.org/10.3758/s13428-022-01945-x

Brysbaert, M., Mandera, P., & Keuleers, E. (2018). The word frequency effect in word processing: An updated review. *Current Directions in Psychological Science*, *27*(1), 45–50. https://doi.org/10.1177/0963721417727521

Buchanan, L., Westbury, C., & Burgess, C. (2001). Characterizing semantic space: Neighborhood effects in word recognition. *Psychonomic Bulletin & Review*, *8*(3), 531–544. https://doi.org/10.3758/BF03196189

Chiarello, C., Shears, C., & Lund, K. (1999). Imageability and distributional typicality measures of nouns and verbs in contemporary English. *Behavior Research Methods, Instruments, & Computers*, *31*(4), 603–637. https://doi.org/10.3758/BF03200739

Coltheart, M. (1981). The MRC psycholinguistic database. *Quarterly Journal of Experimental Psychology*, *33*(4), 497–505. https://doi.org/10.1080%2F14640748108400805

Coltheart, M., Rastle, K., Perry, C., Langdon, R., & Ziegler, J. (2001). DRC: A dual route cascaded model of visual word recognition and reading aloud. *Psychological Review*, *108*(1), Article 204. https://doi.org/10.1037/0033-295X.108.1.204

Connell, L., & Lynott, D. (2010). Look but don't touch: Tactile disadvantage in processing modality-specific words. *Cognition*, *115*(1), 1–9. https://doi.org/10.1016/j.cognition.2009.10.005

Connell, L., & Lynott, D. (2012). Strength of perceptual experience predicts word processing performance better than concreteness or imageability. *Cognition*, *125*(3), 452–465. https://doi.org/10.1016/j.cognition.2012.07.010

Connell, L., & Lynott, D. (2014). I see/hear what you mean: Semantic activation in visual word recognition depends on perceptual attention. *Journal of Experimental Psychology: General*, *143*(2), 527–533. https://doi.org/10.1037/a0034626

Connell, L., & Lynott, D. (2016a). Do we know what we're simulating? Information loss on transferring unconscious perceptual simulation to conscious imagery. *Journal of Experimental Psychology: Learning Memory and Cognition*, *42*(8), 1218–1232. https://doi.org/10.1037/xlm0000245

Connell, L., & Lynott, D. (2016b). Embodied semantic effects in visual word recognition. In M. H. Fischer & Y. Coello (Eds.), *Foundations of embodied cognition* (Vol. 2, pp. 71–92). Psychology Press.

Cortese, M. J., & Fugett, A. (2004). Imageability ratings for 3,000 monosyllabic words. *Behavior Research Methods, Instruments, & Computers*, *36*(3), 384–387. https://doi.org/10.3758/BF03195585

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

Dymarska, A., Connell, L., & Banks, B. (2023a). Weaker than you might imagine: Determining imageability effects on word recognition. *Journal of Memory and Language*, *129*(9), Article 104398. https://doi.org/10.1016/j.jml.2022.104398

Dymarska, A., Connell, L., & Banks, B. (2023b). *Word memory*. https://osf.io/r8fmb

Fernandino, L., Tong, J. Q., Conant, L. L., Humphries, C. J., & Binder, J. R. (2022). Decoding the information structure underlying the neural representation of concepts. *Proceedings of the National Academy of Sciences of the United States of America*, *119*(6), Article e2108091119. https://doi.org/10.1073/pnas.2108091119

Gilhooly, K. J., & Logie, R. H. (1980). Age-of-acquisition, imagery, concreteness, familiarity, and ambiguity measures for 1,944 words. *Behavior Research Methods & Instrumentation*, *12*(4), 395–427. https://doi.org/10.3758/BF03201693

Glanc, G. A., & Greene, R. L. (2007). Orthographic neighborhood size effects in recognition memory. *Memory & Cognition*, *35*(2), 365–371. https://doi.org/10.3758/BF03193457

Glanzer, M., & Adams, J. K. (1985). The mirror effect in recognition memory. *Memory & Cognition*, *13*(1), 8–20. https://doi.org/10.3758/BF03198438

Goh, W. D., & Lu, S. H. X. (2012). Testing the myth of the encoding–retrieval match. *Memory & Cognition*, *40*(1), 28–39. https://doi.org/10.3758/s13421-011-0133-9

Groninger, L. D. (1974). Imagery and subjective categorization effects on long-term recognition and retrieval. *Bulletin of the Psychonomic Society*, *3*(4), 261–263. https://doi.org/10.3758/BF03333461

Hargreaves, I. S., Pexman, P. M., Johnson, J. C., & Zdrazilova, L. (2012). Richer concepts are better remembered: Number of features effects in free recall. *Frontiers in Human Neuroscience*, *6*, Article 73. https://doi.org/10.3389/fnhum.2012.00073

Harm, M. W., & Seidenberg, M. S. (2004). Computing the meanings of words in reading: Cooperative division of labor between visual and phonological processes. *Psychological Review*, *111*(3), Article 662. https://doi.org/10.1037/0033-295X.111.3.662

Higham, P. A., Perfect, T. J., & Bruno, D. (2009). Investigating strength and frequency effects in recognition memory using type-2 signal detection theory. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *35*(1), 57–80. https://doi.org/10.1037/a0013865

Howe, M. L., & Derbish, M. H. (2010). On the susceptibility of adaptive memory to false memory illusions. *Cognition*, *115*(2), 252–267. https://doi.org/10.1016/j.cognition.2009.12.016

James, C. T. (1975). The role of semantic information in lexical decisions. *Journal of Experimental Psychology: Human Perception and Performance*, *1*(2), 130–136. https://doi.org/10.1037/0096-1523.1.2.130

JASP Team. (2020). *JASP* (Version 0.14.1) [Computer software]. https://jasp-stats.org/

Juhasz, B. J., & Yap, M. J. (2013). Sensory experience ratings for over 5,000 mono- and disyllabic words. *Behavior Research Methods*, *45*(1), 160–168. https://doi.org/10.3758/s13428-012-0242-9

Khanna, M. M., & Cortese, M. J. (2021). How well imageability, concreteness, perceptual strength, and action strength predict recognition memory, lexical decision, and reading aloud performance. *Memory*, *5*(5), 1–15. https://doi.org/10.1080/09658211.2021.1924789

Kiefer, M., & Pulvermüller, F. (2012). Conceptual representations in mind and brain: Theoretical developments, current evidence and future directions. *Cortex*, *48*(7), 805–825. https://doi.org/10.1016/j.cortex.2011.04.006

Kuhn, M. (2022). *Caret: Classification and regression training* (R package version 6.0-93). https://CRAN.R-project.org/package=caret

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

Lau, M. C., Goh, W. D., & Yap, M. J. (2018). An item-level analysis of lexical-semantic effects in free recall and recognition memory using the megastudy approach. *Quarterly Journal of Experimental Psychology*, *71*(10), 2207–2222. https://doi.org/10.1177/1747021817739834

Leding, J. K. (2020). Animacy and threat in recognition memory. *Memory & Cognition*, *48*(5), 788–799. https://doi.org/10.3758/s13421-020-01017-5

Lynott, D., & Connell, L. (2009). Modality exclusivity norms for 423 object properties. *Behavior Research Methods*, *41*(2), 558–564. https://doi.org/10.3758/BRM.41.2.558

Lynott, D., & Connell, L. (2013). Modality exclusivity norms for 400 nouns: The relationship between perceptual experience and surface word form. *Behavior Research Methods*, *45*(2), 516–526. https://doi.org/10.3758/s13428-012-0267-0

Lynott, D., Connell, L., Brysbaert, M., Brand, J., & Carney, J. (2020). The Lancaster sensorimotor norms: Multidimensional measures of perceptual and action strength for 40,000 English words. *Behavior Research Methods*, *52*(3), 1–21. https://doi.org/10.3758/s13428-019-01316-z

Madan, C. R. (2020). Exploring word memorability: How well do different word properties explain item free-recall probability? *Psychonomic Bulletin & Review*, *28*(2), 1–13. https://doi.org/10.3758/s13423-020-01820-w

McRae, K., Cree, G. S., Seidenberg, M. S., & McNorgan, C. (2005). Semantic feature production norms for a large set of living and nonliving things. *Behavior Research Methods*, *37*(4), 547–559. https://doi.org/10.3758/BF03192726

Miller, L. M., & Roodenrys, S. (2009). The interaction of word frequency and concreteness in immediate serial recall. *Memory & Cognition*, *37*(6), 850–865. https://doi.org/10.3758/MC.37.6.850

Mirams, L., Poliakoff, E., Brown, R. J., & Lloyd, D. M. (2010). Vision of the body increases interference on the somatic signal detection task. *Experimental Brain Research*, *202*(4), 787–794. https://doi.org/10.1007/s00221-010-2185-7

Mirams, L., Poliakoff, E., Brown, R. J., & Lloyd, D. M. (2012). Interoceptive and exteroceptive attention have opposite effects on subsequent somatosensory perceptual decision making. *Quarterly Journal of Experimental Psychology*, *65*(5), 926–938. https://doi.org/10.1080/17470218.2011.636823

Mirams, L., Poliakoff, E., Brown, R. J., & Lloyd, D. M. (2013). Brief body-scan meditation practice improves somatosensory perceptual decision making. *Consciousness and Cognition*, *22*(1), 348–359. https://doi.org/10.1016/j.concog.2012.07.009

Mirams, L., Poliakoff, E., & Lloyd, D. M. (2017). Spatial limits of visuotactile interactions in the presence and absence of tactile stimulation. *Experimental Brain Research*, *235*(9), 2591–2600. https://doi.org/10.1007/s00221-017-4998-0

Nairne, J. S. (2002). The myth of the encoding-retrieval match. *Memory*, *10*(5–6), 389–395. https://doi.org/10.1080/09658210244000216

Paivio, A. (1971). Imagery and language. In S. J. Segal (Ed.), *Imagery: Current cognitive approaches* (pp. 7–32). Academic Press.

Paivio, A., & Csapo, K. (1969). Concrete image and verbal memory codes. *Journal of Experimental Psychology*, *80*(2, Pt.1), 279–285. https://doi.org/10.1037/h0027273

Paivio, A., & Csapo, K. (1973). Picture superiority in free recall: Imagery or dual coding? *Cognitive Psychology*, *5*(2), 176–206. https://doi.org/10.1016/0010-0285(73)90032-7

Paivio, A., Yuille, J. C., & Madigan, S. A. (1968). Concreteness, imagery, and meaningfulness values for 925 nouns. *Journal of Experimental Psychology*, *76*(1), 1–12. https://doi.org/10.1037/h0025327

Pecher, D., van Dantzig, S., & Schifferstien, H. N. J. (2009). Concepts are not represented by conscious imagery. *Psychonomic Bulletin & Review*, *16*(5), 914–919. https://doi.org/10.3758/PBR.16.5.914

Pecher, D., Zeelenberg, R., & Barsalou, L. W. (2003). Verifying different-modality properties for concepts produces switching costs. *Psychological Science*, *14*(2), 119–124. https://doi.org/10.1111/1467-9280.t01-1-01429

Peterson, M. J., & McGee, S. H. (1974). Effects of imagery instructions, imagery ratings, and number of dictionary meanings upon recognition and recall. *Journal of Experimental Psychology*, *102*(6), 1007–1014. https://doi.org/10.1037/h0036378

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., 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., Siakaluk, P. D., & Yap, M. J. (2013). Introduction to the research topic meaning in mind: Semantic richness effects in language processing. *Frontiers in Human Neuroscience*, *7*, Article 723. https://doi.org/10.3389/fnhum.2013.00723

R Core Team. (2021). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing. http://www.R-project.org/

Recchia, G., & Jones, M. N. (2012). The semantic richness of abstract concepts. *Frontiers in Human Neuroscience*, *6*, Article 315. https://doi.org/10.3389/fnhum.2012.00315

Rubin, D. C., & Friendly, M. (1986). Predicting which words get recalled: Measures of free recall, availability, goodness, emotionality, and pronunciability for 925 nouns. *Memory & Cognition*, *14*(1), 79–94. https://doi.org/10.3758/BF03209231

Schock, J., Cortese, M. J., & Khanna, M. M. (2012). Imageability estimates for 3,000 disyllabic words. *Behavior Research Methods*, *44*(2), 374–379. https://doi.org/10.3758/s13428-011-0162-0

Scott, G. G., Keitel, A., Becirspahic, M., Yao, B., & Sereno, S. C. (2018). 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

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://doi.org/10.3389/fpsyg.2016.01010

Snefjella, B., & Kuperman, V. (2016). It's all in the delivery: Effects of context valence, arousal, and concreteness on visual word processing. *Cognition*, *156*, 135–146. https://doi.org/10.1016/j.cognition.2016.07.010

Speed, L. J., & Brysbaert, M. (2022). Dutch sensory modality norms. *Behavior Research Methods*, *54*(3), 1306–1318. https://doi.org/10.3758/s13428-021-01656-9

Stadthagen-Gonzalez, H., & Davis, C. J. (2006). The Bristol norms for age of acquisition, imageability, and familiarity. *Behavior Research Methods*, *38*(4), 598–605. https://doi.org/10.3758/BF03193891

Tillotson, S. M., Siakaluk, P. D., & Pexman, P. M. (2008). Body–object interaction ratings for 1,618 monosyllabic nouns. *Behavior Research Methods*, *40*(4), 1075–1078. https://doi.org/10.3758/BRM.40.4.1075

Toglia, M. P., & Battig, W. F. (1978). *Handbook of semantic word norms*. Lawrence Erlbaum Associates.

van Heuven, W. J. B., Mandera, P., Keuleers, E., & Brysbaert, M. (2014). SUBTLEX-UK: A new and improved word frequency database for British English. *Quarterly Journal of Experimental Psychology*, *67*(6), 1176–1190. https://doi.org/10.1080/17470218.2013.850521

Venables, W. N., & Ripley, B. D. (2002). *Modern applied statistics with S*. Springer.

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

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

Wilson, M. (1988). MRC psycholinguistic database: Machine-usable dictionary, Version 2.00. *Behavior Research Methods, Instruments, & Computers*, *20*(1), 6–10. https://doi.org/10.3758/BF03202594

Yap, M. J., Tan, S. E., Pexman, P. M., & Hargreaves, I. S. (2011). Is more always better? Effects of semantic richness on lexical decision, speeded pronunciation, and semantic classification. *Psychonomic Bulletin & Review*, *18*(4), 742–750. https://doi.org/10.3758/s13423-011-0092-y

Zdrazilova, L., & Pexman, P. M. (2013). Grasping the invisible: Semantic processing of abstract words. *Psychonomic Bulletin & Review*, *20*(6), 1312–1318. https://doi.org/10.3758/s13423-013-0452-x

Zechmeister, E. B. (1972). Orthographic distinctiveness as a variable in word recognition. *The American Journal of Psychology*, *85*(3), 425–430. https://doi.org/10.2307/1420843

Received January 26, 2023
Revision received April 14, 2023
Accepted April 18, 2023 ■