When abstract becomes concrete: naturalistic encoding of concepts in the brain

Kewenig, V.[1], Vigliocco, G.[1], Skipper, J.[1]

1. Experimental Psychology, University College London

Abstract

Humans learn and process language in complex and dynamic contexts, involving simultaneous processing of connected speech, faces, bodies, objects and including, at times, touch, odor, and taste. How words and their associated concepts are encoded in the brain during real world processing is unclear. We investigated the representational structure of conceptual encoding in a naturalistic setting and asked to what extent brain responses dynamically depend on context. Across different contexts, we find evidence that concrete and abstract concepts encode experience-based information in roughly differential sets of brain regions. However, these differences reduce when contextual variance is considered. Specifically, the response profile of abstract concepts becomes more concrete-like when they are processed in context highly related to their meaning. Conversely, when context is unrelated to a given concrete concept, the activation pattern resembles more that of abstract conceptual processing. Our results suggest that while concepts encode habitual experiences on average, the underlying neurobiological organization is not fixed but depends dynamically on available contextual information at the time of processing.

Introduction

Words are said to encode conceptual knowledge, for example the word “Moon” encodes the concept *the only planet orbiting the earth*. In some cases, the same concept can be encoded by

distinct words, as in case of “Moon” and “Luna”. Similarly, the same word can encode various concepts; “Moon” can also mean *white round silhouette in the sky*. In this sense, what concept is encoded by a word in a given situation determines its meaning (Margolis & Laurence, 2019). Traditionally, concepts are separated into concrete and abstract. Concrete concepts have a physical referent in the world (e.g. *table*) and are encoded in terms of experience with that referent. Abstract concepts on the other hand lack a physical referent because they are defined in a negative way (e.g. *freedom*) or are dissociated from experiential information (e.g. *democracy*). Since it is consequently difficult to see how abstract concepts could be encoded in terms of experience with a given referent (Dove 2022), they have been defined in terms of their associations with other words (Crutch et al., 2009; Crutch & Warrington, 2005) However, behavioral data suggest that encoding of abstract concepts might also retain experiential information related to internal experience, for example emotion or interoception (Kousta et al., 2011).

Studies of the neurobiological basis of conceptual processing tend to support this dichotomy. That is, concrete and abstract concepts are thought to engage separate brain regions involved in processing different types of information. (Bedny & Thompson-Schill, 2006; Binder et al., 2009; Sabsevitz et al., 2005; Wang et al., 2010). Concrete words engage experiential processing regions related to their referent (Barsalou et al., 2003; Pulvermuller & Fadiga, 2010). For example, motor cortices activate during the processing of concrete action verbs like *throw* (Hauk & Pulvermüller, 2004) or action related nouns like *hammer* (Kiefer & Pulvermüller, 2012), auditory cortex for concrete sound-related words like *telephone* (Goldberg et al., 2006; Kiefer et al., 2008) and visual cortex for concrete color-related words like *yellow* (Hsu et al., 2011; Simmons et al., 2007). Consistent with the view that abstract concepts are encoded in relation to other words rather than experience, they correlate with greater activation in brain

regions associated with general linguistic processing (J. R. Binder, 2005; Mellet, 1998; Noppeney et al., 2004; Sabsevitz et al., 2005). However, neurobiological data also support the view that abstract conceptual encoding retains experiential information (Fernandino et al., 2022). Due to their lack of external referents, this is said to be characterized mainly by internal experience. For example, studies have found that activation in brain regions associated with emotional (Oosterwijk et al., 2015; Vigliocco et al., 2014), and interoceptive processing (Connell et al., 2018) correlate with abstract concepts.

Overall, previous studies suggest that conceptual encoding in the brain is the product of a stable set of fixed regions processing different types of information depending on whether a concept is concrete or abstract (with some ambiguity regarding the latter). However, the idea of a stable dichotomy between concrete and abstract concepts might be flawed for real-world contextualized processing (Lebois et al., 2015). Instead, what information is used for conceptual encoding may depend entirely on the contextual information available. For example, a situational encoding of, say, the concrete concept *chair* could be linked to many internal aspects such as goals ("I want to rest"), motivation ("I have been standing for 2 hours"), emotion ("I would like to feel comfortable") and theory of mind ("are those older people more in the need of the chair than me?") - which clearly have abstract elements. Conversely, an abstract concept like *truth* no longer seems so abstract when used as reference to a perceived physical situation (such as "it is snowing") that matches the utterance's meaning (that it is snowing). Here, *truth* refers to a quite concrete state of the world (Barsalou et al., 2018).

Previous behavioral data support the view that contextual information can affect conceptual encoding (Chambers et al., 2004; Cooper, 1974; Eberhard et al.) For example, when an object is depicted in a context consistent with its use, the action associated with using the object is

more readily available than when the context is more consistent with picking the object up (Kalenine et al., 2014). There is also neurobiological evidence that objects visually present in a situation can influence conceptual encoding (Hoffman et al., 2013; Yee & Thompson-Schill, 2016). For example, a study found that task-related color-congruency of objects correlated with less activation of brain regions involved in color perception during conceptual encoding – likely because less retrieval of detailed color knowledge was necessary (Hsu et al., 2011). However, no data exist on the extent to which context can modulate the representational nature of conceptual encoding. The present study investigates the hypothesis that naturalistic conceptual encoding in the brain, as measured during movie watching, is not fixed. Rather, the supporting brain regions depend on the information available in context at the time of processing. Here, we test two corollaries of this view.

The first concerns the traditional debate about the representational nature of encodings. When concepts are investigated across different contexts, their neurobiology will partly reflect the average experiential information of the type of situational context in which they are usually processed. Thus, we predict that both concrete and abstract conceptual encoding retain experiential information. More specifically, we suggest that the neurobiological encoding of concrete concepts retains experiences related to their physical referent, which is predominantly characterized by sensorimotor information and associated brain regions (Pulvermuller, 2018; Willems et al., 2010). Abstract concepts on the other hand do not necessarily have such referents. Therefore, we expect their encoding to mostly reflect information related to internal experience as well as more general linguistic processing, implying brain regions associated with emotional interoceptive and general linguistic processing (Reinboth & Farkaš, 2022).

The second corollary concerns the role of context in modulating neurobiological encoding of concepts. An investigation of different contexts may show that the neurobiological structures underlying concrete and abstract conceptual encoding dynamically switch alliance depending on whether the situation renders an abstract concept concrete or vice versa. We predict that when the situation contains little information related to a concrete concept's physical referent, response profiles will include regions related to general linguistic and internal processing. On the other hand, when there is a lot of related information present in the situation (for example the word "love" processed in a situation of two people kissing), abstract conceptual encoding will more heavily draw on regions involved in sensorimotor processing. We propose that this erosion of the concrete/abstract dichotomy for contextualized processing shows that both concrete and abstract concepts draw on information related to experience as well as linguistic association. In other words, conceptual encoding is determined by both, reference as well as conceptual role (Block, 1998). Which associated neurobiological structures are expressed during processing depends dynamically on the contextual information available. This new way of thinking about the representational nature of conceptual knowledge may help synthesize contradicting evidence concerning the involvement of different brain areas during conceptual processing.

Results

Conceptual Encoding Across Context

First, we investigate naturalistic conceptual processing across different contexts. We predict that across contexts, concrete and abstract concepts are processed in a separable set of brain regions in line with previous studies. To test this prediction, we use an amplitude modulated deconvolution regression (Chen et al., 2015) to analyse the data from 86 participants who watched 1 of 10 movies during functional magnetic resonance imaging (fMRI) (Aliko et al.,

2020). This model has the advantage that it does not assume an ideal shape of the hemodynamic response function, which is instead derived from each individual participant's timeseries. Specifically, we estimate the response over a 20s window starting at word-onset for all non-function words in the movie, which were additionally matched for frequency and length in order to avoid confounds from these lexical variables. We also estimate the modulation of the amplitude of these responses using independent ratings of concreteness and abstractness of each word (Brysbaert et al., 2014).

Confounds in our naturalistic movie task are visual and auditory activation from the stimulus. Therefore, our model includes a measure of the average luminance (low level visual variable) and volume (low level auditory variable) for each word as nuisance regressors. Luminance as well as Loudness for each frame in the movie was measured across the full duration of each word using the "Librosa" package for music and audio analysis in Python (McFee et al, 2015). The final amplitude modulated βs of this analysis represent the BOLD response beyond the average response across concrete and abstract words. We conduct a group level linear mixed effects model and subsequent general linear tests and correct all results for multiple comparisons using an alpha level of 0.01.

Post hoc contrasts between modulation of brain response to concrete and abstract words revealed where activation for concreteness was greater than abstractness (Figure 1, red). In the frontal lobes, concreteness was greater in the right posterior IFG and the precentral sulcus and gyrus. In parietal and temporal cortices, activity was greater for concreteness in the transverse temporal gyrus and sulcus, planum polare and temporale, and superior temporal gyrus (STG) and sulcus (STS). Finally, a very large swaths of the occipital lobe (running into the inferior temporal lobe) and precuneus were activated more for concreteness than abstractness,

bilaterally. Subcortically, the dorsal and posterior medial cerebellum was more active bilaterally for concreteness.

Conversely, activation for abstract word-onsets was greater than concreteness in a grossly different set of regions (Figure 1, blue). In the frontal lobes, this includes bilateral cingulate and lateral and medial aspects of the superior frontal gyrus. More left frontal activity for abstractness was in the prefrontal cortices and more right in the orbital gyrus, posterior inferior frontal gyrus (IFG), and precentral sulcus and gyrus. In the parietal and temporal lobes, activity was greater for abstractness in the angular gyri (AG) and inferior parietal lobes, including the postcentral gyrus and aspects of the STS. Subcortically, activity was greater bilaterally in anterior thalamus and nucleus accumbens as well as left lateralised in the amygdala.

Finally, there was overlap in activity between the contrasts of concreteness and abstractness, indicating some regions are involved in processing both sets of words but on different time scales (Figure 1, yellow). In frontal, parietal, and temporal regions this was primarily left lateralized and included the IFG, AG, and STG, respectively. In the occipital lobe (OCC), processing overlapped primarily around the calcarine sulcus. In all cases, abstract words were processed in overlapping regions at a longer delay (i.e. more slowly than concrete words despite being matched on length and frequency to concrete words).

Partially consistent with prior research, these separable response profiles support the hypothesis that across contexts, the central tendency for concrete and abstract concepts is to activate regions associated with different experiential processes. Specifically, concreteness is associated with multiple regions involved in sensorimotor processing (Mkrtychian et al., 2019), including most of the dorsal and ventral visual systems (Gao et al., 2019) and the right frontal

motor system (Pulvermuller, 2005). In contrast, abstractness engages regions typically associated with 'internalised' processing like interoception (Harpaintner et al., 2018; Villani et al., 2021) (anterior thalamus, somatosensory cortex), autobiographical memory (Conca, Borsa, et al., 2021) (anterior medial prefrontal regions), and emotional processing and regulation (Conca, Catricala, et al., 2021) (anterior medial prefrontal regions, orbital prefrontal cortex, dorsolateral prefrontal cortex, nucleus accumbens and amygdala). That said, concrete as well as abstract concepts combine visual and linguistic encoding (Tang et al., 2021). On the one hand, abstract concepts also clearly engage left sensorimotor regions though to a much smaller extent, including visual and left motor regions. On the other hand, overlap engages regions associated with language processing (like the left superior temporal plane and IFG) and early visual processing regions.

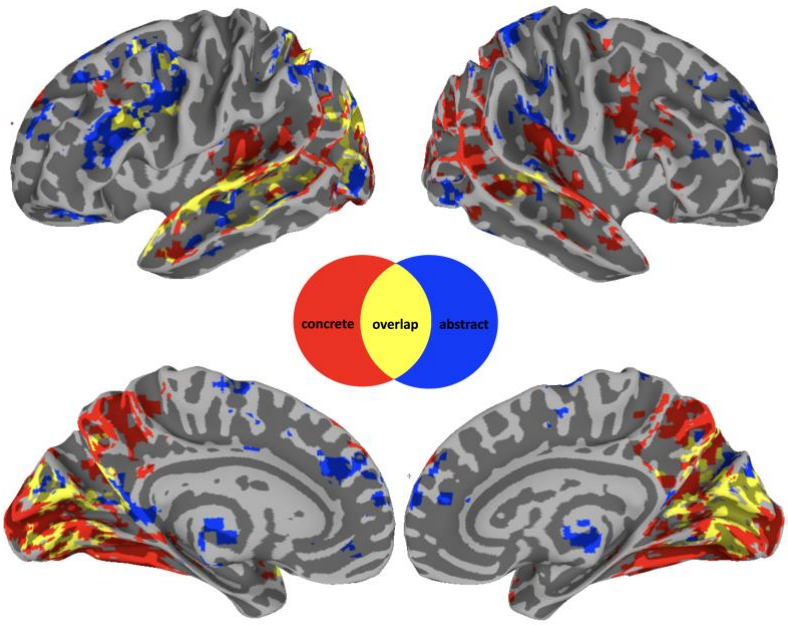


Figure 1. *Group level results from linear mixed effect models and post hoc general linear tests contrasting activity at concrete- and abstract word onset. Activation associated with concrete word onsets are coloured in light blue, brain responses associated with concrete word onsets are coloured in yellow. Overlap of activity is displayed in dark blue. Overlaps still indicate differences between abstact and concrete words at delays. Results are corrected for multiple comparisons at* $\alpha = 0.01$ *and displayed with a cluster-size >= 20.*

Meta-Analytic Descriptions of Clusters

The interpretation of results in the prior section are based on informal reverse inference. To more formally and quantitatively evaluate the putative distinction between concrete and abstract words, this and the subsequent section will employ a meta-analytic and reverse correlation analysis, respectively. In both analyses we examined the hypothesis that brain regions involved in concrete and abstract conceptual processing would reflect different types of habitual experience with their associated referents. Namely, concrete concepts are more associated with sensorimotor information whereas abstract concepts are more associated with 'internalised' (e.g., interoceptive and emotional) information (Harpaintner et al., 2018; Pexman et al., 2017).

To test these hypotheses, we extracted the coordinates of the centres of mass of each cluster from each condition in Figure 1. We then input these coordinates into the neurosynth package (Yarkoni et al, 2011), which reports term-based meta-analytic coactivation maps. These maps reflect coactivation of brain regions seeded with a 6 mm sphere centered on the selected location across studies in the Neurosynth database (14,371 peer-reviewed neuroscience publications; accessed in January 2022). Maps are created using a two-way Anova quantifying the strength of association between terms used in the abstract or title of publications in the Neurosynth Database and voxel activation at the specified coordinates. We extracted terms with a z-score greater than 3.12 ($p < 0.001$).

Comparing counts of associated key words for concrete and abstract activation clusters with a Kruskal Wallis test, we find that significantly more abstract clusters (coloured in blue) are related to internal processing compared to concrete clusters (coloured in red). Out of 35 abstract clusters in total, there were (in order of frequency): 9 abstract clusters associated with the term

"Autobiographical Memory", compared to 1 concrete cluster ($H(2) = 4$, $p = 0.05$); 8 clusters with the term "Valence", compared to 0 concrete clusters ($H(2) = 5.6$, $p = 0.01$); 6 clusters to the term "Theory of Mind" compared to 0 concrete clusters ($H(2) = 4.8$, $p = 0.03$); 6 clusters to the term "Nausea", compared to 0 concrete clusters ($H(2) = 4.8$, $p = 0.03$); 5 clusters to the term "Pain", compared to 0 concrete clusters ($H(2) = 4$, $p = 0.05$). The same test reveals that significantly more (11 out of a total of 20) concrete clusters are related to "Movement" (coloured in red) compared to abstract clusters ($H(2) = 12.4$, $p < 0.001$). No significant difference in association was found for the key words "Reward/Motivation", "Social/Empathy", "Touch", "Speech" and "Language". Finally, "Language" was the only term associated more with overlapping clusters than both concreteness ($H(2) = 4$, $p = 0.045$) and abstractness ($H(2) = 7$, $p < 0.001$), (10 out of 16 clusters, colored in yellow).

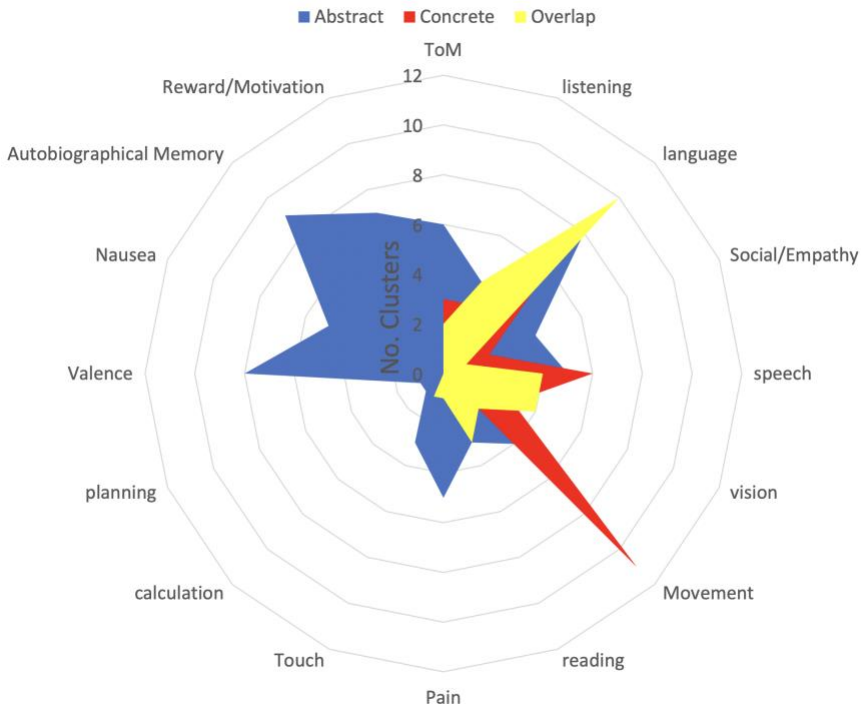


Figure 2. *Meta-analytic description of concrete and abstract activation clusters. We used Neurosynh to find the terms associated with the centers of mass for each activation cluster from each condition. Displayed is the number of activation clusters associated with each meta-analytic term. We find that significantly more abstract clusters are associated with 'Autobiographical Memory' and 'Valence' compared to 'Movement' for concrete clusters.*

Peak and Valley Analysis

The meta-analytic approach suggests that brain activity correlated with concrete and abstract word-onset relates to separable experiential domains. However, this approach is potentially biased toward high frequency and relatively nonspecific terms in the existing neuroimaging literature and, more importantly, are not drawn from the specific stimuli as experienced by each participant. Thus, we did a variant of reverse correlation (Hasson et al., 2004) called a "peak and valley" analysis (Skipper et al., 2009) in which the sensorimotor and emotional properties of words processed at peaks (when the brain is processing information) is compared to valleys (when the brain is either not processing information or actively inhibiting it).

First, we extract all participants' average timeseries from each concrete and abstract activation cluster across voxels. Next, we determine peaks and valleys by calculating the discrete difference along each timeseries (Figure 3, step 1) to determine whether the signal is rising (peak) or falling (valley). We then average these values across participants and delay the timeseries to the words mentioned in the subtitle according to the canonical model of the hemodynamic response function (which peaks at around 6s). This way, we can extract words that were mentioned in the movie during peaks and valleys in the averaged timeseries of each cluster (Figure 3, step 2). We then use the Lancaster sensorimotor norms (Lynott et al., 2020) and the norms for valence and arousal (Warriner et al., 2013) to determine a 12-dimensional experience-based representation for each word (Figure 3, step 3). For each dimension, we create two categorical arrays, one for peaks and one for valleys, noting down 0 if the word mentioned at a peak or valley was not highly rated on the dimension and 1 if it was highly rated (larger than one standard deviation from the dimension's mean rating). We then conduct a Kruskal Wallis test between these arrays to determine whether a given experiential dimension occurred significantly more with peaks than valleys in the averaged timeseries of a cluster

(Figure 3, step 4). Finally, we compare the number of concrete and abstract activation clusters with significantly more peaks than valleys for each dimension using a further Kruskal-Wallis test.

For concrete clusters, we expect significantly more sensorimotor dimensions (in the corpora these are “Haptic”, “Foot_Leg”, “Hand_Arm”, “Torso”) to be associated with peaks rather than valleys in the timeseries compared to abstract clusters. Conversely, we expect significantly more experiential dimensions related to internal processing (in the corpora these are “Autobiographical Memory”, “Valence”, “Arousal” and “Interoception”) to be associated with excitatory peaks for abstract word onsets relative to valleys compared to concrete words.

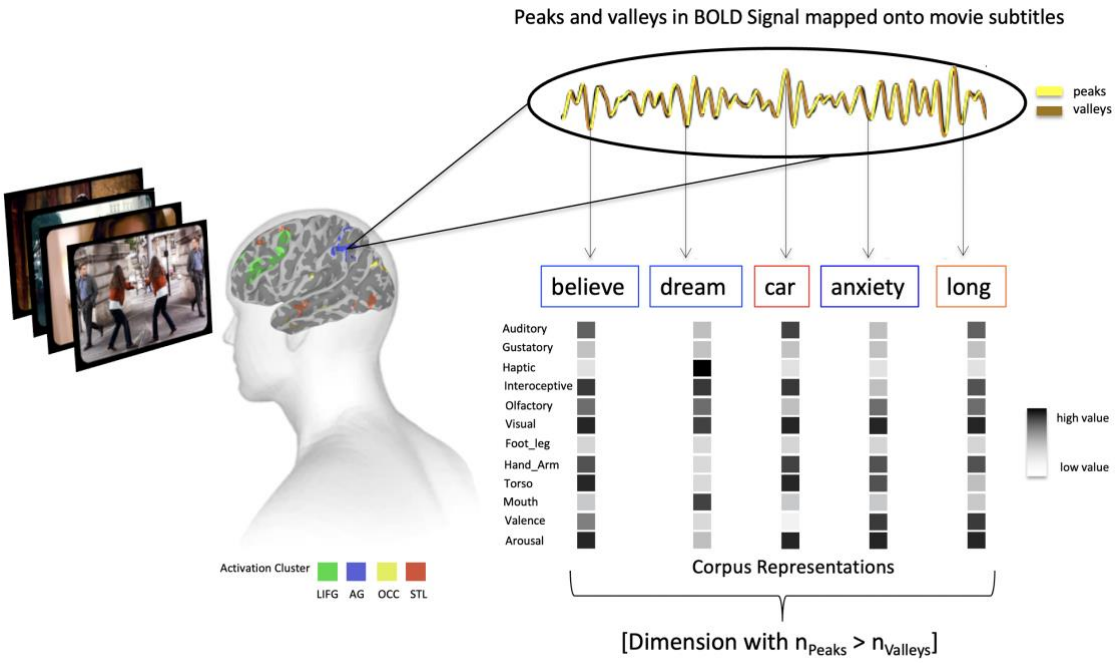


Figure 3. *Peak and Valley Analysis to understand cluster wise sensorimotor and emotional encodings. First, we average the fMRI timeseries for each participant, for abstract, concrete and overlap clusters of activity from Figure 1. Then we label peaks and valleys (1) and average across participants. Next, we map the averaged peaks and valleys onto word on- and off-set times (2). Finally, we estimate sensorimotor as well as valence and arousal representations for each word (3) and determine which dimensions are associated with significantly more peaks than valleys in each cluster using a Kruskal Wallis test (4).*

Results show that peaks associated with four clusters processing concrete words (Figure 4, red) are correlated significantly with “Torso” compared to valleys. This is significantly higher than for abstract clusters ($H(2)=7$, $p<0.001$). Three clusters were correlated with “Haptic” and “Mouth”, which is also significantly more than for abstract clusters ($H(2)=5.2$, $p=0.02$). Two clusters with “Foot_Leg” were not significant but the means were in the correct direction more than for abstract clusters ($H(2)=3.4$, $p=0.06$). As concerns abstract words, associated peaks in eight clusters (Figure 4, blue) are significantly correlated with the dimension “Valence”. This number is significantly higher than for concrete clusters ($H(2)=5.6$, $p=0.02$). Three clusters are associated with the dimension (“Auditory”), which is not significantly more than for concrete clusters ($H(2)=1.9$, $p=0.17$). Peaks compared to valleys in five clusters of overlap activity (figure 4, yellow) correlated more with the dimension “Mouth” (see the appendix for this analysis with 4 and 5s lag). Overall, our results confirm that abstract clusters process valence-based information, while concrete clusters process sensorimotor information.

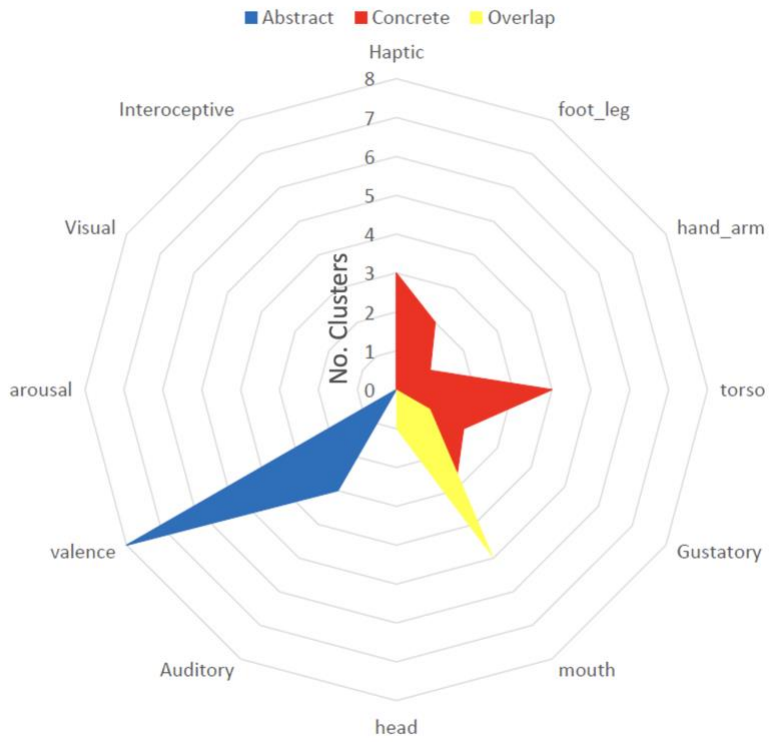


Figure 4. *Peak and valley analysis. We extract the type of information processed in each activation cluster by looking at experience-based features of those words that correlated with significantly more peaks than valleys (see figure 3 for details). Words highly rated on sensorimotor dimensions correlated significantly with peaks in concrete clusters (red). Mentioning of words highly rated on emotion-related dimensions correlated significantly with peaks in abstract clusters (blue). Finally, mentioning of words highly rated on the dimension "Mouth" correlated significantly with peaks in overlap clusters.*

Conceptual Encoding in Context

Across contextual information, we observe a central tendency for concrete and abstract words to be processed in separable brain regions encoding different types of experiential information. However, we hypothesized that these response profiles are not completely fixed but instead dynamically change depending on contextual information.

We develop a new procedure for testing this hypothesis. First, we automatically extract labels for all objects visually present in a 2s (or 60 frames) window before a given abstract or concrete word was mentioned, using two pre-trained visual recognition models: faster R-CNN (He et al., 2015)and OmniSource (Duan et al., 2020) (Figure 5, step 1). Next, we average GloVe vector representations (Pennington et al., 2014)for all labels within this frame, which represent

the meaning of each label based on global word-word co-occurrence statistics. We then calculate the cosine similarity between the average object representation and the Glove vector representation of the word mentioned in the movie (Figure 5, step 2).

We take this distance to measure how contextually embedded a given word is within its visual environment at the time of processing. Concrete and abstract words that are highly embedded (cosine distance one standard deviation below the median, < 0.37) or have a low embeddedness score (cosine distance one standard deviation above the median, > 0.63) are included for further analysis (for some examples see Figure 5, step 3). We use the same amplitude modulation approach as in the analysis previously conducted across context on a subset of highly and lowly embedded concrete and abstract words. In addition to low-level visual and auditory regressors, estimates included our measure of contextual embeddedness. A group level linear mixed effects model and subsequent general linear tests between high and low context conditions was conducted and all results were corrected for multiple comparisons using an alpha level of 0.01.

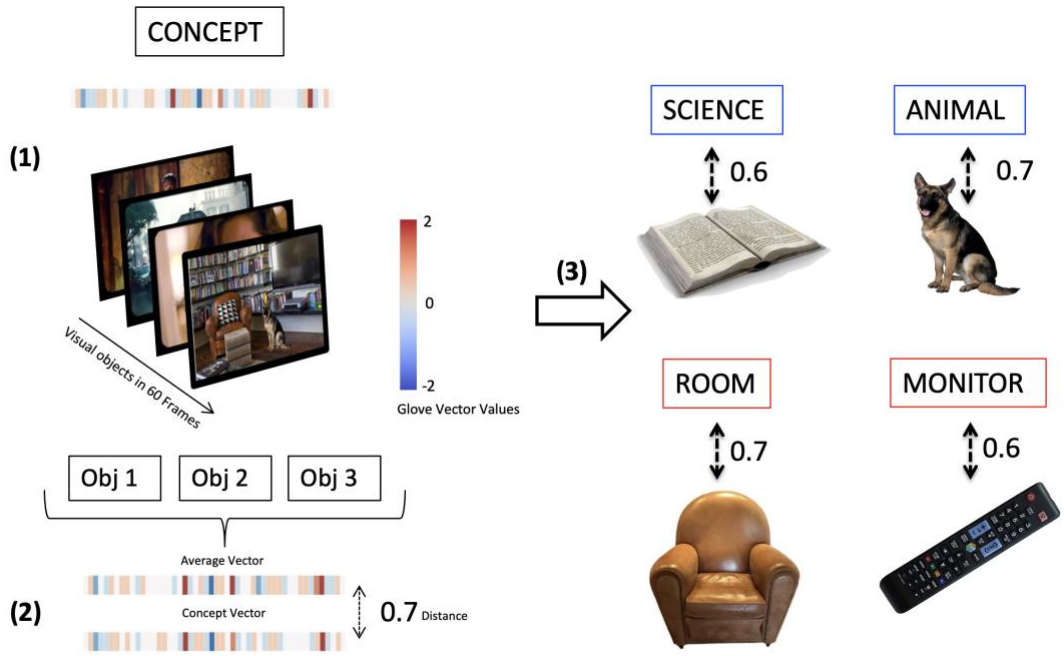


Figure 5. *Estimating Contextual embeddedness for each concrete and abstract word to model context-dependent modulation of conceptual encodings. We use visual recognition models for automatically extracting labels that were visually present in the scene (60 frames, ~2 seconds) before a given concept was mentioned in the movie (step 1). We then correlate an average Glove Vector representation of all these labels with a Glove Vector representation of the mentioned word to estimate how closely related the labels of objects in the scene are to the concept (step 2). Displayed are four randomly extracted measures of contextual embeddedness for both abstract (blue frame) and concrete (red frame) words (step 3).*

The interaction between word-type (abstract/concrete) and contextual embeddedness (high/low) results in activity in regions that partially overlaps with both concrete and abstract activation clusters from the previous analysis (Figure 6). These include the middle and superior frontal gyrus, cingulate, precuneus, and AG. An automated meta-analysis of the unthresholded interaction effect using Neurosynth (xx) reveals that the default mode network or "DMN" ($r(1500) = 0.194$, $p < 0.001$), "Default Mode" ($r(1500) = 0.219$, $p < 0.001$) and "Default" ($r(1500) = 0.226$, $p < 0.001$) are the most frequent terms with this response profiles.

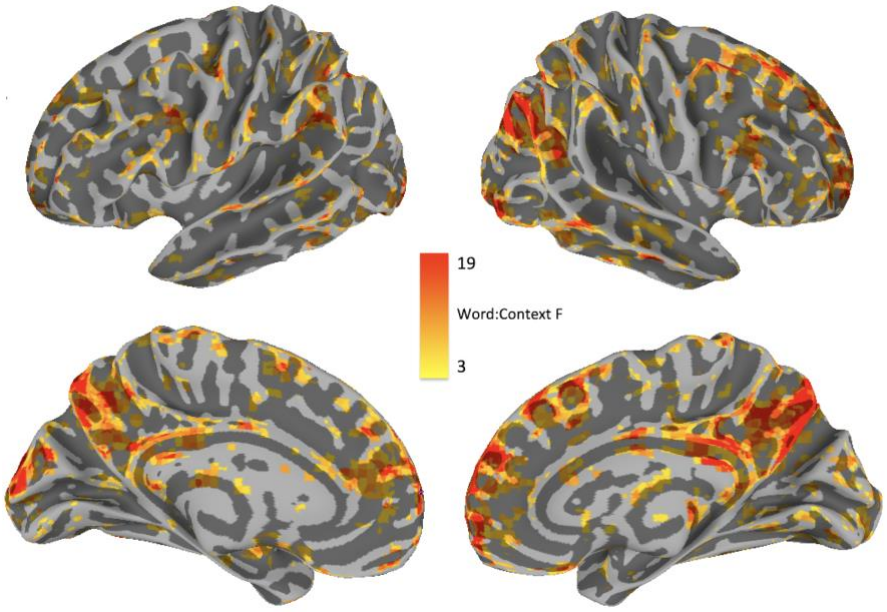


Figure 6. *Main effect of the interaction between contextual embeddedness (high/low) and conceptual processing, results corrected for multiple comparisons at α = 0.01.*

Post hoc general linear tests contrasting concrete and abstract words in low and high contextual embeddedness yield no significant differences in the low context condition that survive correction for multiple comparisons. In the high context condition, concrete words resulted in more activity in visual and primary auditory areas (Figure 7, red) whereas abstract words yielded more activity in the temporal lobes, AG, Precuneus as well as superior and medial frontal regions (Figure 7, blue). Using the same meta-analytic approach as above in high contextual embeddedness, concrete words are associated with inferior parietal and occipital regions and abstract words are correlated with activation in DMN nodes.

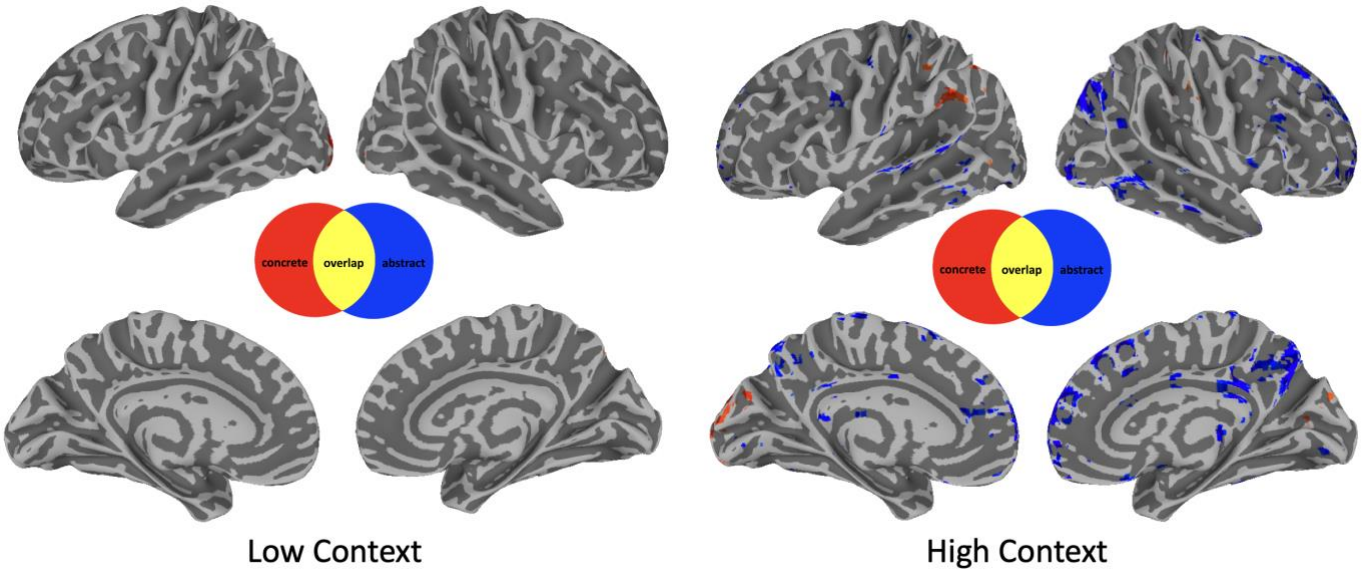
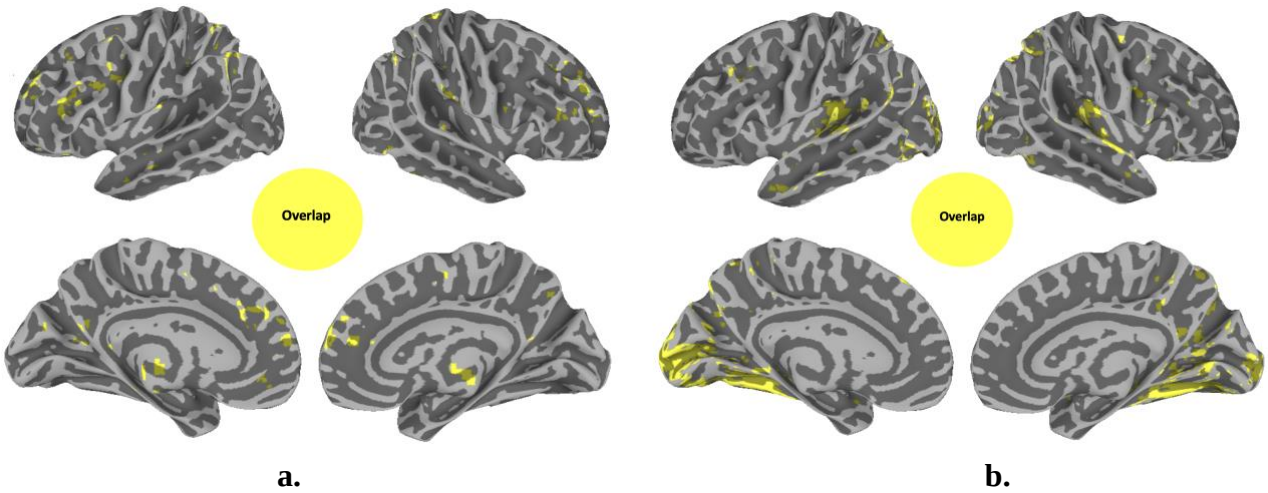


Figure 7. *Contrasts between concrete and abstract concepts in high and low contextual embeddedness. No differential activation found in low context. Differential activation in high context is found for abstract words (coloured in blue) in the Angular Gyrus (cluster-size = 56, z = 2.9, p = 0.03) Precuneus (cluster-size = 33, z = 2.3, p < 0.03), Posterior Cingulate cortex (cluster-size = 28, z = 2.6, p < 0.03) Middle Prefrontal Cortex (cluster-size = 26, z = 2.6, p < 0.03) and Middle Frontal Gyrus (cluster-size = 22, z = 2.1, p = 0.04). For concrete words, differential activation in high context (red) is found in primary auditory (cluster-size = 26, p = 0.01) and primary visual (cluster-size = 31, p = 0.02) areas. All p-values corrected for multiple comparisons at α = 0.01.*

In a final analysis, we use cosine similarity to spatially correlate the unthresholded low and high context results with the unthresholded results obtained in the first set of analyses for concrete and abstract words collapsing across context. This shows that in low context conditions, the spatial pattern of activity for concrete words more resembles that of abstract words ($r(72964) = 0.49$, $p < 0.001$) than concrete words averaged across context ($r(72964) = 0.12$, $p < 0.001$). In particular, low context concrete words overlap bilaterally with abstract words (across context) sub-cortically in ACC, Thalamus, Precuneus and medial prefrontal regions. Cortically, overlap is found bilaterally in primary auditory regions and AG. Left lateralized overlap is found in the superior and anterior temporal lobe (Figure 8a).

Conversely, in the high context condition, the organization of the brain during processing abstract words resembles more that of contextually averaged concrete ($r(72964) = 0.62$, $p <$

0.001) than abstract words (r(72964) = 0.25, p = 0.476). Overlap can be found subcortically in fusiform, ACC and motor regions bilaterally. Both premotor and motor cortex as well as anterior temporal lobe are also engaged by abstract words processed in high context as well as concrete words across context.



Figures 8. *Spatial overlap between thresholded statistical brain images of concrete and abstract conceptual processing obtained from the original contextually averaged model and from the model with an added regressor for contextual embeddedness. Results show overlap in subcortical and prefrontal regions during processing of concrete concepts in low cortex with contextually averaged processing of abstract concepts (a). Conversely, patterns of activation during processing of abstract concepts in high context overlap with activation during contextually averaged concrete conceptual processing in visual, auditory and motor regions (b).*

Discussion

In sets of analyses in and out of context, we estimated brain activity derived from participants watching movies in order to better understand the neurobiological organization of real-world conceptual encoding. First, we investigated where in the brain concrete and abstract concepts are generally processed independent of context as well as the type of information they encode. Given the hypothesis that conceptual encoding reflects contextual information at the time of processing, we expected a set of regions that process relevant average contextual information

to activate across different contexts. That is, we expected concrete conceptual encoding to activate regions more involved in sensorimotor processing and abstract conceptual encoding to activate regions associated with more 'internal' (emotional and interoceptive) as well as general linguistic processing.

Indeed, contrasting estimated activity for concrete and abstract conceptual encoding, we find response profiles consistent with this prediction. Whereas concrete word onsets are more associated with regions and terms related to sensorimotor processing, abstract words are more associated with regions and meta-analytic terms related to internal processing such as valence, nausea and pain. Similarly, a peak and valley analysis showed that concrete word regions were more highly associated with sensorimotor properties (i.e., "Torso", "Haptic", and "Hand_Arm") whereas abstract word regions were associated with "Valence". Together, these results are evidence that, on average, concrete and abstract conceptual encoding includes regions associated with sensorimotor and internal experience, respectively.

That said, it cannot be claimed that there is a strong dichotomy or boundary between concrete and abstract conceptual processing in the brain. Nor can this be the only encoding format for these word types. First, regions involved in sensorimotor processing are clearly active for both word types despite the preponderance for concrete words. Second, overlap between abstract and concrete word activation was found in regions often associated with language processing. This is evidence supporting that both concrete and abstract conceptual encoding make use of shared experiential information as well as linguistic associations between the words themselves – that is, their "conceptual role" (Piantadosi & Hill, 2022). Finally, depending on available context, regions associated with concrete or abstract conceptual encoding in the brain may switch alliance (Barsalou & Wiemer-Hastings, 2005).

Consistent with this last point, a review of behavioral and neuroimaging data suggests that conceptual encodings are not stable but may vary as a function of contextual information (Cai & Vigliocco, 2018). In a second set of analyses, we aimed to start understanding the dynamics of conceptual encoding during more naturalistic language processing investigating contextual modulation. We compared the activation profiles of concrete and abstract concepts in high and low contextual embeddedness with brain images obtained when averaging across contexts. Our results show that concrete concepts can become more abstract-like in contexts with less relevant visual information. Conversely, abstract concepts become more concrete-like when they are highly embedded in relevant visual information. We propose that this is because an abstract concept processed in highly related visual context relates more to extrinsic situational variables – which are usually encoded by concrete concepts. Conversely, when a concrete concept is processed in displaced context, its encoding will relate more to intrinsic variables and linguistic associations, which are usually encoded by abstract concepts. Thus, when context is taken into account, the fixed neurobiological boundaries between concrete and abstract concepts begin to erode.

The Mechanism behind Contextual Modulation

What is the neurobiological mechanism behind this contextualized modulation of conceptual encoding in the brain? Our results indicate that variance in visual contextual embeddedness correlates with regions commonly defined as the default mode network (DMN). A recent review of the DMN concludes that it is an active and dynamic 'sense-making' network. Accordingly, nodes in the DMN activate for processing incoming extrinsic information (external variables such as objects present in a situation) together with intrinsic information (internal variables like prior beliefs or feelings) (Yeshurun et al., 2021).

The fact that important nodes in the DMN activate during conceptual encoding in situations of high and low contextual embeddedness fits with this view as follows: if activation of intrinsic and extrinsic components of a given stimulus happens in the DMN, activity necessarily varies depending on how much relevant extrinsic information is available (i.e. situations of high or low contextual embeddedness). This is further supported by the fact that contrasts show no differential activity in conditions of low context (as no extrinsic information can be integrated). However, DMN differentially activates in high context for abstract concepts because their encoding depends on intrinsic information and therefore requires more integration when a lot of relevant extrinsic information is present. For concrete concepts, DMN is not activated in high context because no intrinsic information needs to be integrated. Instead, early visual and auditory regions orchestrate the integration of relevant external features. Together, these results suggest that the DMN plays a pivotal role in contextual modulation of conceptual encoding via integrating external and internal streams of information.

Neurobiological Models of Language

The present findings are at odds with classic and contemporary models of language in neuroscience (Fedorenko et al., 2011; Geschwind, 1970; Hickok, 2022). Roughly, these are based on two assumptions: (i) language is processed in a stable set of regions, (ii) located primarily on the left hemisphere. Our results suggest instead that language in the brain is (i) not limited to this set of regions but is distributed across the whole brain and (ii) not processed in a stable network but activates different regions in a dynamic and contextually determined way (Skipper & Willems, 2015). Rather than the classic neurobiological model of language, this fits with an emergentist view of conceptual encoding (Elman, 2004) as the accumulation of knowledge about the contexts in which a given word and its associated concept usually

occur. On this view, there is no representational distinction between knowledge of a referent and knowledge of the contexts in which the reference occurs.

One explanation for contradicting neurobiological evidence with classic models is suggested by the fact that regions found active for both concrete and abstract concepts near completely overlap with the more classical 'language' regions. Specifically, when categories of stimuli (i.e. different word-meanings) and contexts are averaged across, differences in conceptual encoding average out, once thresholded (Wang et al, 2022). Averaging both word-meaning and contextual information leaves only brain regions involved in either a subset of aspects of language processing (like auditory and phonological processing), encoding of linguistic associations, or general regions of high connectivity coordinating more distributed processes (Skipper et al., 2022). Future work should carefully define what aspects of language can be studied with different ways of averaging across word- and context-types.

Limitations

Our method of quantifying context in naturalistic stimuli presents some drawbacks. For one, there may be general matching issues for comparing high vs low context. For example, high contextual embeddedness may correlate with more visual input and therefore "naturally" engage more sensorimotor regions. A priori, there is no reason to believe that the presence of more objects should drive up semantic correlation measures, however. An analysis of object-counts can be found in the appendix.

A concern regarding our data analysis is the re-use of a subset of our original data for the second, contextual analysis. This was done to maximize the initial limited number of exemplars in order to successfully model word categories. To analyze this possible issue of double-

dipping, we reran the original analysis excluding high and low context words used in the second analysis and confirmed that results were comparable.

Implications

Generally, it is clear from our study that contextual information modulates the neurobiological basis of conceptual encoding and language processing in a profound way. Given this result, future work should not only attempt to analyze specific linguistic categories and processes in context. Instead, they should quantify other types of contexts (for example, semantic, emotional, etc.) in similar ways and examine how these can affect conceptual encoding in the brain.

In addition to implications for the nature of the neurobiological organization underlying conceptual encoding, these results also suggest a new challenge for computational models of natural language processing (NLP) in the field of artificial intelligence: namely that of understanding and modeling the dynamic influences of contextual information on emergent representational structures. The recent innovation of transformer based artificial neural nets in NLP has pushed performance towards almost-human level (Devlin et al., 2018; Vaswani et al., 2017) and advances in this field promise to have a profound impact on human existence. The key factor differentiating transformer-based models is that they define word-embeddings dynamically depending on discourse context. Even though transformer-based models seem to share this computational principle with the human brain (Caucheteux & King, 2022; Goldstein et al., 2022), brain-model comparisons have been limited to discourse context only.

Yet, models trained on multimodal datasets show better match to human brain activity than those trained on text alone (Lu et al., 2022). In combination with the present set of results, this

suggests that including multimodal context effects (e.g. visual and discourse context) could help to push models towards more human-like representations and efficiency. However, more research is needed on both the model and the brain side to understand how conceptual encoding changes depending on the type of context in which words are processed. As this line of research continues, model states will come into closer alignment with human brains, and each such improvement will enrich both the model's conceptual encoding as well as our understanding of the underlying neurobiology.

Materials and Methods

The present study analysed the 'Naturalistic Neuroimaging Database' (NNDb) (Aliko et al., 2020).

Materials

Participants and Task

The Naturalistic Neuroimaging Database (Aliko et al., 2020, https://openneuro.org/datasets/ds002837/versions/1.0.6) includes 86 right-handed participants (42 females, range of age 18–58 years, M = 26.81, SD = 10.09 years) undergoing fMRI scans while watching one of 10 full-length movies selected across a range of genres. All had unimpaired hearing and (corrected) vision. None had any contraindication for magnetic resonance imaging (MRI), history of psychiatric or neurological disorder, or language-related learning disabilities. All participants gave informed consent, and the study was approved by the University College London Ethics Committee.

Word Annotation

Words were annotated in the movies using automated approaches with a machine learning based speech-to-text transcription tool from Amazon Web Services (AWS;

https://aws.amazon.com/transcribe/). The resulting transcripts contained on and offset timings for individual words. However, as not all words were transcribed or accurately transcribed, timings were subsequently hand corrected.

Variables of Interest

Concreteness

We used measures of abstractness and concreteness for each word in the movies (excluding function words), taken from existing norms (Brysbaert et al, 2013). In the norms, 40,000 words are rated on a scale from 0 (not experience based) to 5 (experience based) by 4,000 participants. We only included content words from the movies (concrete or abstract, based on a median split on the concreteness scale) matched for frequency and length to avoid confounds from these lexical variables. After this matching process, we were left with 880 words (half concrete, half abstract). Mean concreteness rating for abstract words was 1.83, compared to 3.22 for concrete words. The mean log frequency and mean length for abstract words was 4.81 and 5.27 compared to 5.21 and 4.91 for concrete words. T-tests between the ratings of both groups revealed no significant differences.

Control Variables

Confounds in our naturalistic movie task included visual and auditory activation from the stimulus. Therefore, our model included a measure of the average luminance (low level visual variable) and volume (low level auditory variable) for each word as nuisance regressors. Luminance as well as Loudness for each frame in the movie was measured across the full duration of each word using the “Librosa” package for music and audio analysis in Python (McFee et al, 2015).

Regional Analysis across Contexts

In this analysis we aimed to test our prediction that when contextual information is averaged together, the neurobiological organization of conceptual processing will reflect experiential

information broadly in line with previous studies. Specifically, sensorimotor information for concrete concepts and interoceptive and emotional information for abstract concepts. All statistical analyses on the preprocessed NiFTI files were carried out in AFNI (Cox, 1995, 1997). Individual AFNI programs are indicated parenthetically or in italics in subsequent descriptions.

Deconvolution Analysis

We used an amplitude modulated deconvolution regression to estimate brain activity. In contrast to a standard convolved regression analysis, this type of model does not assume the canonical hemodynamic response function. Instead, the response function is estimated over a time-window at stimulus onset through a multiple basis function for each participant individually at word onset. This promises a better understanding of shape differences and achieves higher statistical power at both individual and group level (Chen et al, 2015). We chose a 20s time window because this should be sufficient to capture the hemodynamic response function for the average word length. We selected Csplin over the Tent function to deconvolve the BOLD signal because this function offers more interpolation between time points, which might result in a more precise estimate of the individual response function (but is computationally more costly). The final amplitude modulated βs of this analysis represent the BOLD response beyond the average response across words. Since we included two nuisance regressors as amplitude modulators in the model and estimated BOLD response at each timepoint over a 20s window for both concrete and abstract words, this resulted in (20 * 2 * 2) 80 regressors overall.

Group Analysis

Group analysis was performed on the deconvolved brain images with linear mixed effect modelling (LMM) using *'3dLME'*. The within-subject factors were concreteness with two levels (concrete and abstract) and time with twenty levels (one for each time-point). Next to

the interaction between factors, the model included a centered variable for age of participant (xx levels), gender (2 levels) as well as movie ID (10 levels) as covariates. A random intercept for participant and an autocorrelation structure for modeling the effect estimates of the multiple basis function were also included.

Correcting for Multiple Comparisons

To correct for multiple comparisons in all LMMs, we used a multi-threshold approach rather than choosing an arbitrary *P* value at the individual voxel level as is customary. In particular, we used a cluster simulation method to estimate the probability of noise-only clusters using the spatial autocorrelation function from the residuals in each LMM (*3dFWHMx* and *3dClustSim*). This resulted in the cluster sizes to achieve a corrected alpha value of 0.01 at 9 different *P* values (i.e., 0.05, 0.02, 0.01, 0.005, 0.002, 0.001, 0.0005, 0.0002, and 0.0001). We thresholded each map at the corresponding z-value for each of these nine p-values and associated cluster size. We then combined the resulting maps, leaving each voxel with its original z-value.

Estimating Regional Clusters of Activation

We used *3dMerge* with the "1clust" to form clusters with a 1mm connection distance and minimum size of 20 voxels and clip off all other voxels. We determined the center of mass for each of these clusters using *'3dCM'*. The resulting spatial activation map is displayed in figure (3).

Meta-Analysis

The resulting coordinates of the center of mass of each cluster were input into Neurosynth (https://neurosynth.org/), an online tool which includes activation maps of 14371 neuroscientific studies. Neurosynth automatically mines all words in titles and abstracts of these articles and performs a two-way ANOVA, testing for the presence of a non-zero association between terms used in abstracts or titles reporting activation that overlaps with the input location. We web-scraped all terms with z scores above 3 ($p < 0.001$) and repeated this

procedure for each cluster to determine whether any terms would be repeatedly associated with concrete and abstract activation clusters (see figure 4). Then we tested whether certain dimensions were more important for concrete or abstract clusters using a Kruskal Wallis test (not correcting for multiple comparisons).

Peaks and Valleys Analysis

We extracted the averaged timeseries for each activation cluster across voxels using *3dMerge*. Next, we determine peaks and valleys by calculating the discrete difference along the timeseries using the "Numpy" Python package (Harris et al, 2020) (step 1 in figure 3). This is negative if the next number is larger (rising timeseries, therefore peak) and positive if the next number is smaller (falling timeseries, therefore valley). Given that the canonical model of the hemodynamic response function peaks at around 6s after stimulus onset, we extracted what words were mentioned at each peak and each valley in a given cluster's timeseries with a 5- and 6-seconds lag (step 2 in figure 3). We then use the Lancaster sensorimotor norms (Lynott et al, 2019) and norms for valence and arousal Warriner (2013) to determine a 12-dimensional experience-based representation for each word (step 3 in figure 3), which included the dimensions: "Auditory", "Gustatory", "Haptic", "Interoception", "Visual", "Hand_Arm", "Foot_Leg", "Torso", "Mouth", "Head", "Valence" and "Arousal". For each dimension, we create two categorical arrays, one for peaks and one for valleys, noting down 0 if the word mentioned at a peak or valley was not highly rated on the dimension and 1 if it was highly rated. Extreme ratings were defined as deviating at least 1 standard deviation from the mean. Given the distributional nature of this data, we then conduct a Kruskal Wallis test between these arrays to determine whether a given experiential dimension occurred significantly more with peaks than valleys in the averaged timeseries of a cluster (step 4 in figure 3). We repeated this procedure for both, a 5s and a 6s timeseries lag and conducted a cosine-similarity test

between each result (using the “Sklearn” package in Python (Pedregosa et al, 2011) in order to ensure that they were significantly similar.

Analysis of Contextualized Processing

We wanted to see how contextual information modulates brain activity during processing of concrete and abstract concepts. We predicted that when abstract concepts are embedded in a highly related context, they can engage neurobiological structures that are usually involved in processing concrete concepts. Conversely, when concrete concepts are displaced from surrounding context, we predict them to engage more abstract-like regions in the brain. To test these predictions, we utilize two pre-trained visual recognition models, Faster R-CNN (Ren et al., 2015) and OmniSource (Duan et al., 2020), to extract object features using computer vision toolboxes (Chen et al., 2019; MMCV Contributors, 2018; MMAction2 Contributors, 2020), respectively. For each prediction frame (about every four frames, i.e., 4*0.04 = 0.16 s), the object recognition model generated a list of detected object labels and kept those that had a prediction confidence of 90% or greater. Final object features were represented as the average of the vectorized object labels using GloVe (Pennington et al, 2014), which represents the meaning of each label via global co-occurrence statistics. We then estimated a representation of a 2s (or 60 frames) context window (which corresponds to the average length of an utterance) for each word listing all the labels of objects visually present in each frame within that window. Finally, we calculated the cosine similarity between the vector representation of each word and its context average, using the “Sklearn” package in Python (Pedregosa et al, 2011), to estimate contextual embeddedness for each concept (a value between 0 and 1). We then estimate the modulation of the amplitude of the brain response to concrete and abstract concepts as a cause of high and low contextual embeddedness (values of $x < 0.37$ or $x > 0.63$ respectively) by using an amplitude modulated deconvolution regression with an estimated response function.

Deconvolution Analysis

As in the previous analysis, the response function was estimated over 20s through the CSPLIN multiple basis function for each participant individually at word onset. This time we only used a sub-set of all matched abstract and concrete words, namely those that were highly or lowly embedded in context according to our measure (568 concrete and abstract words). The model included the nuisance regressors luminance and volume as amplitude modulators, together with our measure of contextual embeddedness and estimated BOLD response at each timepoint over a 20s window for both concrete and abstract words in high and low context. This resulted in (20 * 2 * 2 * 2) 160 regressors overall.

Group Analysis

Group analysis was performed on the deconvolved brain images with linear mixed effect modelling. The within-subject factors were concreteness with two levels (concrete and abstract), contextual embeddedness with two levels (high and low) and time with twenty levels (one for each time-point). In addition to interactions between each of the factor pairs, the model included a centered variable for age of participant (86 levels), their gender (2 levels) as well as movie ID (10 levels) as covariates. A random intercept for participant and an autocorrelation structure for modeling the effect estimates of the multiple basis function were also included. Finally, we estimated the cosine similarity between the unthresholded activity maps from this analysis and the unthresholded maps obtained from Analysis (1) using the “Scipy” package in Python (Virtanen et al, 2020).

Appendix

Appendix A Peak and Valley Analysis

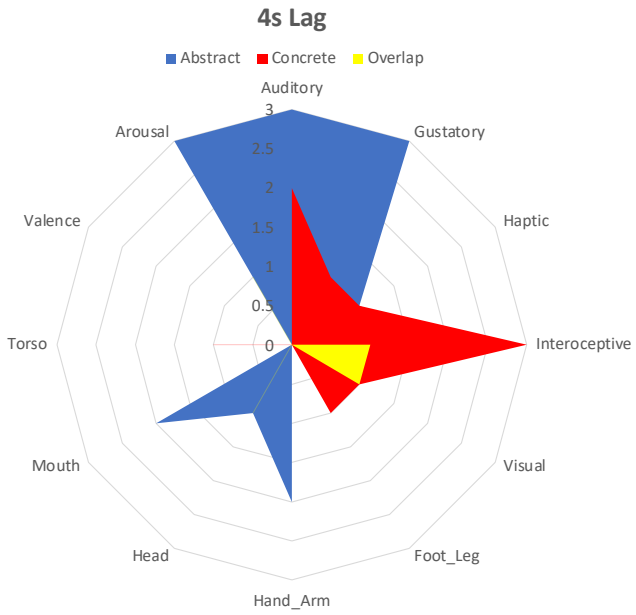


Figure A1. *Peak and Valley Analysis for a 4s lag. A Kruskal-Wallis test shows that the distribution between sensorimotor and interoceptive/emotional dimensions for concrete and abstract words is significantly different to a 5s (H(2)=4,8, p=0.03 and 6s (H(2)=5.3, p=0.02) lag.*

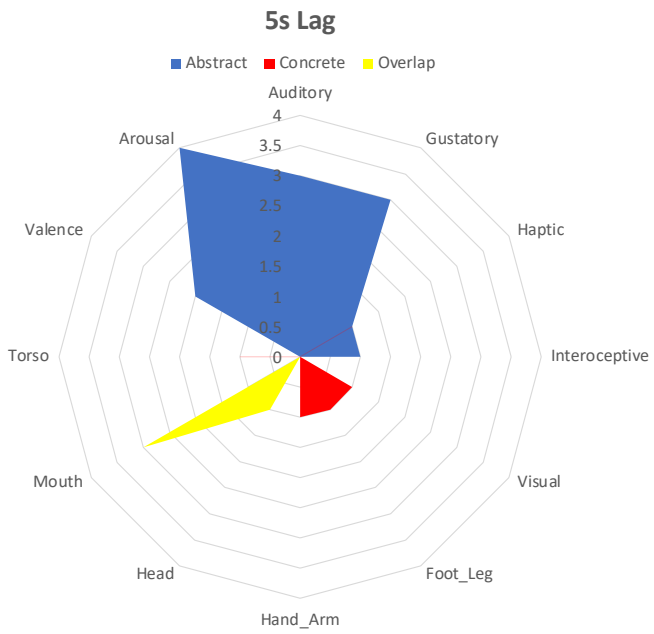


Figure A2. *Peak and Valley Analysis for a 5s lag. A Kruskal-Wallis test shows that the distribution between sensorimotor and interoceptive/emotional dimensions for concrete and abstract words is not significantly different from a 6s lag. (H(2)=1.7, p=0.1)*

Appendix B Context Analysis

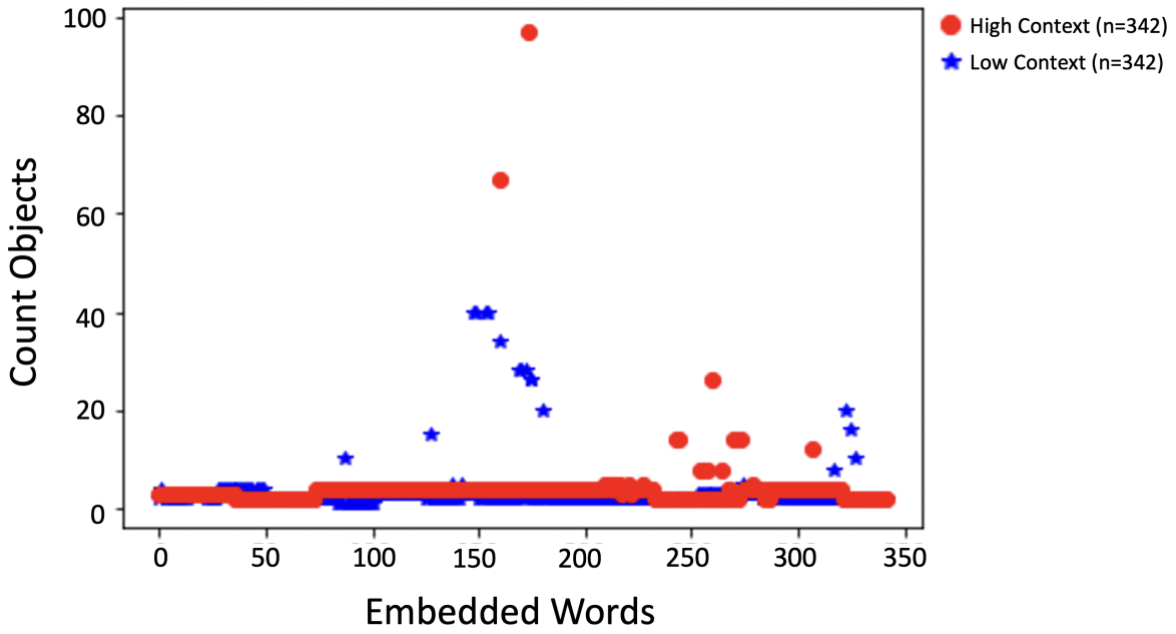


Figure B1. *Comparison between Object Counts for words in high and low context. A Kruskal-Wallis test showed no significant difference between the conditions (H(2)=0.23, p=0.63*

Appendix C Code

We implemented our data analysis in Bash, Python and R. Our code will be provided as online supplemental material upon publication and hosted openly on a dedicated Github repository.

References

Aliko, S., Huang, J., Gheorghiu, F., Meliss, S., & Skipper, J. I. (2020). A naturalistic neuroimaging database for understanding the brain using ecological stimuli. *Sci Data*, *7*(1), 347. https://doi.org/10.1038/s41597-020-00680-2

Barsalou, L. W., Dutriaux, L., & Scheepers, C. (2018). Moving beyond the distinction between concrete and abstract concepts. *Philos Trans R Soc Lond B Biol Sci*, *373*(1752). https://doi.org/10.1098/rstb.2017.0144

Barsalou, L. W., Simmons, W. K., Barbey, A. K., & Wilson, C. D. (2003). Grounding Conceptual Knowledge in Modality Specific Systems. *Trends in Cogntivie Sciences*, *7*(7), 7.

Barsalou, L. W., & Wiemer-Hastings, K. (2005). Situating Abstract Concepts. In *Grounding Cognition* (pp. 129-163). https://doi.org/10.1017/cbo9780511499968.007

Bedny, M., & Thompson-Schill, S. L. (2006). Neuroanatomically separable effects of imageability and grammatical class during single-word comprehension. *Brain Lang*, *98*(2), 127-139. https://doi.org/10.1016/j.bandl.2006.04.008

Binder, J. R., Desai, R. H., Graves, W. W., & Conant, L. L. (2009). Where is the semantic system? A critical review and meta-analysis of 120 functional neuroimaging studies. *Cereb Cortex*, *19*(12), 2767-2796. https://doi.org/10.1093/cercor/bhp055

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

Cai, Z., & Vigliocco, G. (2018). Word Processing. *Language and Thought*, *3*.

Caucheteux, C., & King, J. R. (2022). Brains and algorithms partially converge in natural language processing. *Commun Biol*, *5*(1), 134. https://doi.org/10.1038/s42003-022-03036-1

Chambers, C. G., Tanenhaus, M. K., & Magnuson, J. S. (2004). Actions and affordances in syntactic ambiguity resolution. *J Exp Psychol Learn Mem Cogn*, *30*(3), 687-696. https://doi.org/10.1037/0278-7393.30.3.687

Chen, G., Saad, Z. S., Adleman, N. E., Leibenluft, E., & Cox, R. W. (2015). Detecting the subtle shape differences in hemodynamic responses at the group level. *Front Neurosci*, *9*, 375. https://doi.org/10.3389/fnins.2015.00375

Conca, F., Borsa, V. M., Cappa, S. F., & Catricala, E. (2021). The multidimensionality of abstract concepts: A systematic review. *Neurosci Biobehav Rev*, *127*, 474-491. https://doi.org/10.1016/j.neubiorev.2021.05.004

Conca, F., Catricala, E., Canini, M., Petrini, A., Vigliocco, G., Cappa, S. F., & Della Rosa, P. A. (2021). In search of different categories of abstract concepts: a fMRI adaptation study. *Sci Rep*, *11*(1), 22587. https://doi.org/10.1038/s41598-021-02013-8

Connell, L., Lynott, D., & Banks, B. (2018). Interoception: the forgotten modality in perceptual grounding of abstract and concrete concepts. *Philos Trans R Soc Lond B Biol Sci*, *373*(1752). https://doi.org/10.1098/rstb.2017.0143

Cooper, R. M. (1974). The Control of Eye Fixation by the Meaning of Spoken Language. *Cognitive Psychology*, *6*, 23.

Cox, R. W. (1995). AFNI: Software for Analysis and Visualization of Functional

Magnetic Resonance Neuroimages. *Computers and Biomedical Research*, *29*(14), 11.

Cox, R. W. (1997). Software tools for analysis and visualization of fMRI data. *Nmr in Biomedicine*, *10*, 7.

Crutch, S. J., Connell, S., & Warrington, E. K. (2009). The different representational frameworks underpinning abstract and concrete knowledge: evidence from odd-one-out judgements. *Q J Exp Psychol (Hove)*, *62*(7), 1377-1388, 1388-1390. https://doi.org/10.1080/17470210802483834

Crutch, S. J., & Warrington, E. K. (2005). Abstract and concrete concepts have structurally different representational frameworks. *Brain*, *128*(Pt 3), 615-627. https://doi.org/10.1093/brain/awh349

Devlin, J., Chang, M. W., Lee, K., & Toutanova, K. (2018). BERT: Pre-training of Deep Bidirectional Transformers for

Language Understanding. *Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*, *1*. https://doi.org/10.18653/v1/N19-1423

Duan, H., Zhao, Y., Xiong, Y., Liu, W., & Lin, D. (2020). Omni-source Webly-supervised Learning for Video Recognition. *Bioarchive (Tech Report)*. https://doi.org/ arXiv:2003.13042

Eberhard, K. M., Spivey-Knowlton, M. J., Sedivy, J. C., & Tanenhaus, M. K. Eye Movements as a Window into Real-Time Spoken Language Comprehension in Natural Contexts.

Elman, J. L. (2004). An alternative view of the mental lexicon. *Trends Cogn Sci*, *8*(7), 301-306. https://doi.org/10.1016/j.tics.2004.05.003

Fedorenko, E., Behr, M. K., & Kanwisher, N. (2011). Functional specificity for high-level linguistic processing in the human brain. *Proc Natl Acad Sci U S A*, *108*(39), 16428-16433. https://doi.org/10.1073/pnas.1112937108

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. *Proc Natl Acad Sci U S A*, *119*(6). https://doi.org/10.1073/pnas.2108091119

Gao, C., Baucom, L. B., Kim, J., Wang, J., Wedell, D. H., & Shinkareva, S. V. (2019). Distinguishing abstract from concrete concepts in supramodal brain regions. *Neuropsychologia*, 102-110. https://doi.org/10.1016/j.neuropsychologia.2019.05.032

Geschwind, N. (1970). The Organization of Language in the Brain. *Science*, *170*, 5.

Goldberg, R. F., Perfetti, C. A., & Schneider, W. (2006). Perceptual knowledge retrieval activates sensory brain regions. *J Neurosci*, *26*(18), 4917-4921. https://doi.org/10.1523/JNEUROSCI.5389-05.2006

Goldstein, A., Zada, Z., Buchnik, E., Schain, M., Price, A., Aubrey, B., Nastase, S. A., Feder, A., Emanuel, D., Cohen, A., Jansen, A., Gazula, H., Choe, G., Rao, A., Kim, C., Casto, C., Fanda, L., Doyle, W., Friedman, D., Dugan, P., Melloni, L., Reichart, R., Devore, S., Flinker, A., Hasenfratz, L., Levy, O., Hassidim, A., Brenner, M., Matias, Y., Norman, K. A., Devinsky, O., & Hasson, U. (2022). Shared computational principles for language processing in humans and deep language models. *Nat Neurosci*, *25*(3), 369-380. https://doi.org/10.1038/s41593-022-01026-4

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

Hasson, U., Nir, Y., Levy, I., Fuhrmann, G., & Malach, R. (2004). Intersubject Synchronization of Cortical Activity During Natural Vision. *Science*, *303*(5664), 6.

Hauk, O., & Pulvermüller, F. (2004). Somatotopic Representation of Action Words in Human Motor and Premotor Cortex. *Neuron*, *41*, 6. https://doi.org/10.1016/s0896-6273(03)00838-9

He, K., Zhang, Y., Ren, S., & Sun, J. (2015). Deep Residual Learning for Image Recognition. *Bioarchive (Tech Report)*. https://doi.org/https://doi.org/10.48550/arXiv.1512.03385

Hickok, G. (2022). The dual stream model of speech and language processing. *Handb Clin Neurol*, *185*, 57-69. https://doi.org/10.1016/B978-0-12-823384-9.00003-7

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

Hsu, N. S., J.M., D., M. Kraemer, Robyn, Oliver, T., & Schlichting, M. L. T.-S., S. (2011). Color, Context, and Cognitive Style: Variations in Color Knowledge Retrieval as a Function of Task and Subject Variables. *Journal of Cognitive Neuroscience*, *23*(9), 13.

J. R. Binder, C. F. W., K. A. McKiernan, E. T. Possing and D.A. Medler (2005). Distinct Brain Systems for Processing Concrete

and Abstract Concepts. *Journal of Cognitive Neuroscience*, *17*(6), 12.
Kalenine, S., Shapiro, A. D., Flumini, A., Borghi, A. M., & Buxbaum, L. J. (2014). Visual context modulates potentiation of grasp types during semantic object categorization. *Psychon Bull Rev*, *21*(3), 645-651. https://doi.org/10.3758/s13423-013-0536-7
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
Kiefer, M., Sim, E. J., Herrnberger, B., Grothe, J., & Hoenig, K. (2008). The sound of concepts: four markers for a link between auditory and conceptual brain systems. *J Neurosci*, *28*(47), 12224-12230. https://doi.org/10.1523/JNEUROSCI.3579-08.2008
Kousta, S. T., Vigliocco, G., Vinson, D. P., Andrews, M., & Del Campo, E. (2011). The representation of abstract words: why emotion matters. *J Exp Psychol Gen*, *140*(1), 14-34. https://doi.org/10.1037/a0021446
Lebois, L. A., Wilson-Mendenhall, C. D., & Barsalou, L. W. (2015). Are Automatic Conceptual Cores the Gold Standard of Semantic Processing? The Context-Dependence of Spatial Meaning in Grounded Congruency Effects. *Cogn Sci*, *39*(8), 1764-1801. https://doi.org/10.1111/cogs.12174
Lu, H., Zhou, Q., Fei, N., Lu, Z., M., D., Wen, J., Du, C., Zhao, X., Sun, H., He, H., & Wen, J. R. (2022). Multimodal foundation models are better simulators
of the human brain. *arXivLabs (preprint)*. https://doi.org/
https://doi.org/10.48550/arXiv.2208.08263

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. *Behav Res Methods*, *52*(3), 1271-1291. https://doi.org/10.3758/s13428-019-01316-z
Mellet, E., Tzourio, N., Mazoyer, B. (1998). Cortical anatomy of mental imagery of concrete nouns based on their dictionary definition. *NeuroReport*, *9*(5), 5.
Mkrtychian, N., Blagovechtchenski, E., Kurmakaeva, D., Gnedykh, D., Kostromina, S., & Shtyrov, Y. (2019). Concrete vs. Abstract Semantics: From Mental Representations to Functional Brain Mapping. *Front Hum Neurosci*, *13*, 267. https://doi.org/10.3389/fnhum.2019.00267
Noppeney, U., Phillips, J., & Price, C. (2004). The neural areas that control the retrieval and selection of semantics. *Neuropsychologia*, *42*(9), 1269-1280. https://doi.org/10.1016/j.neuropsychologia.2003.12.014
Oosterwijk, S., Mackey, S., Wilson-Mendenhall, C., Winkielman, P., & Paulus, M. P. (2015). Concepts in context: Processing mental state concepts with internal or external focus involves different neural systems. *Soc Neurosci*, *10*(3), 294-307. https://doi.org/10.1080/17470919.2014.998840
Pennington, J., Socher, R., & Manning, C. D. (2014). GloVe: Global Vectors for Word Representation. *Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing*, 11. https://doi.org/10.3115/v1/D14-1162
Pexman, P. M., Heard, A., Lloyd, E., & Yap, M. J. (2017). The Calgary semantic decision project: concrete/abstract decision data for 10,000 English words. *Behav Res Methods*, *49*(2), 407-417. https://doi.org/10.3758/s13428-016-0720-6
Piantadosi, S., T., & Hill, F. (2022). Meaning Without Refernce in Large Language Models. *Archive (preprint)*. https://doi.org/https://doi.org/10.48550/arXiv.2208.02957
Pulvermuller, F. (2005). Brain mechanisms linking language and action. *Nature Reviews Neuroscience*, *6*, 6.

Pulvermuller, F. (2018). Neural reuse of action perception circuits for language, concepts and communication. *Prog Neurobiol*, *160*, 1-44. https://doi.org/10.1016/j.pneurobio.2017.07.001

Pulvermuller, F., & Fadiga, L. (2010). Active perception: sensorimotor circuits as a cortical basis for language. *Nat Rev Neurosci*, *11*(5), 351-360. https://doi.org/10.1038/nrn2811

Reinboth, T., & Farkaš, I. (2022). Ultimate Grounding of Abstract Concepts: A Graded Account. *Journal of Cognition*, *5*(1). https://doi.org/10.5334/joc.214

Sabsevitz, D. S., Medler, D. A., Seidenberg, M., & Binder, J. R. (2005). Modulation of the semantic system by word imageability. *Neuroimage*, *27*(1), 188-200. https://doi.org/10.1016/j.neuroimage.2005.04.012

Simmons, W. K., Ramjee, V., Beauchamp, M. S., McRae, K., Martin, A., & Barsalou, L. W. (2007). A common neural substrate for perceiving and knowing about color. *Neuropsychologia*, *45*(12), 2802-2810. https://doi.org/10.1016/j.neuropsychologia.2007.05.002

Skipper, J. I., Aliko, S., Brown, S., Jo, Y. J., Lo, S., Molimpakis, E., & Lametti, D. R. (2022). Reorganization of the Neurobiology of Language After Sentence Overlearning. *Cereb Cortex*, *32*(11), 2447-2468. https://doi.org/10.1093/cercor/bhab354

Skipper, J. I., Goldin-Meadow, S., Nusbaum, H. C., & Small, S. L. (2009). Gestures orchestrate brain networks for language understanding. *Curr Biol*, *19*(8), 661-667. https://doi.org/10.1016/j.cub.2009.02.051

Skipper, J. I., & Willems, R. M. (2015). The NOLB model: a model of the natural organization of language and the brain. In *Cognitive Neuroscience of Natural Language Use* (pp. 101-134). https://doi.org/10.1017/cbo9781107323667.006

Tang, J., LeBel, A., & Huth, A. G. (2021). Cortical Representations of Concrete and Abstract Concepts in Language Combine Visual and Linguistic Representations. https://doi.org/10.1101/2021.05.19.444701

Vaswani, A., Shazeer, A., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A., Kaiser, L., & Polosukhin, I. (2017). Attention is All You Need. *arXivLabs (preprint)*. https://doi.org/https://doi.org/10.48550/arXiv.1706.03762

Vigliocco, G., Kousta, S. T., Della Rosa, P. A., Vinson, D. P., Tettamanti, M., Devlin, J. T., & Cappa, S. F. (2014). The neural representation of abstract words: the role of emotion. *Cereb Cortex*, *24*(7), 1767-1777. https://doi.org/10.1093/cercor/bht025

Villani, C., Lugli, L., Liuzza, M. T., Nicoletti, R., & Borghi, A. M. (2021). Sensorimotor and interoceptive dimensions in concrete and abstract concepts. *J Mem Lang*, *116*, 104173. https://doi.org/10.1016/j.jml.2020.104173

Wang, J., Conder, J. A., Blitzer, D. N., & Shinkareva, S. V. (2010). Neural representation of abstract and concrete concepts: a meta-analysis of neuroimaging studies. *Hum Brain Mapp*, *31*(10), 1459-1468. https://doi.org/10.1002/hbm.20950

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

Willems, R. M., Hagoort, P., & Casasanto, D. (2010). Body-specific representations of action verbs: neural evidence from right- and left-handers. *Psychol Sci*, *21*(1), 67-74. https://doi.org/10.1177/0956797609354072

Yee, E., & Thompson-Schill, S. L. (2016). Putting concepts into context. *Psychon Bull Rev*, *23*(4), 1015-1027. https://doi.org/10.3758/s13423-015-0948-7

Yeshurun, Y., Nguyen, M., & Hasson, U. (2021). The default mode network: where the idiosyncratic self meets the shared social world. *Nat Rev Neurosci*, *22*(3), 181-192. https://doi.org/10.1038/s41583-020-00420-w