Title
Flexible Conceptual Representations

Permalink
https://escholarship.org/uc/item/902102cv

Journal
Cognitive Science, 48(6)

ISSN
0364-0213

Authors
Truman, Alyssa
Kutas, Marta

Publication Date
2024-06-01

DOI
10.1111/cogs.13475

COGNITIVE SCIENCE
A Multidisciplinary Journal

Cognitive Science 48 (2024) e13475

ISSN: 1551-6709 online
DOI: 10.1111/cogs.13475

Flexible Conceptual Representations

Alyssa Truman, Marta Kutas

Department of Cognitive Science, University of California, San Diego

Received 5 September 2023; received in revised form 14 May 2024; accepted 5 June 2024

Abstract

A view that has been gaining prevalence over the past decade is that the human conceptual system is malleable, dynamic, context-dependent, and task-dependent, that is, flexible. Within the flexible conceptual representation framework, conceptual representations are constructed ad hoc, forming a different, idiosyncratic instantiation upon each occurrence. In this review, we scrutinize the neurocognitive literature to better understand the nature of this flexibility. First, we identify some key characteristics of these representations. Next, we consider how these flexible representations are constructed by addressing some of the open questions in this framework: We review the age-old question of how to reconcile flexibility with the apparent need for shareable stable definitions to anchor meaning and come to mutual understanding, as well as some newer questions we find critical, namely, the nature of relations among flexible representations, the role of feature saliency in activation, and the viability of all-or-none feature activations. We suggest replacing the debate about the existence of a definitional stable core that is obligatorily activated with a question of the *degree* and *probability* of activation of the information constituting a conceptual representation. We rely on published works to suggest that (1) prior featural salience matters, (2) feature activation may be graded, and (3) Bayesian updating of prior information according to current demands offers a viable account of how flexible representations are constructed. This proposal provides a theoretical mechanism for incorporating a changing momentary context into a constructed representation, while still preserving some of the concept's constituent meaning.

Keywords: Concepts; Representations; Flexible semantics; Context-dependent effect; Brain; Neuroscience; Fuzzy concepts; Flexible conceptual representations

---

Correspondence should be sent to Alyssa Truman, Department of Cognitive Science, University of California San Diego, 9500 Gilman Dr., La Jolla, San Diego, CA 92093, USA. E-mail: atruman@ucsd.edu

1. Introduction: Conceptual representations are flexible

Although many of the previous classical theoretical and computational models of human cognitive processes implicitly treat conceptual representations as stable and invariant (e.g., Coltheart, Rastle, Perry, Langdon, & Ziegler, 2001; Dry & Storms, 2010; Machery, 2007; Mahon & Caramazza, 2008; Marocco, Cangelosi, Fischer, & Belpaeme, 2010), an increasing number of neurocognitive publications suggest that the human conceptual system is malleable, dynamic, context-dependent, and task-dependent, in one word, flexible. Within this view, to which we will refer as the flexible conceptual representation framework, conceptual representations are not thought of as static, stable, or fixed holistic units stored in human minds, but rather as "fuzzy" (Rosch, 2011), "fluid" (Yee, 2017), and flexible entities constructed ad hoc (Casasanto & Lupyan, 2015) as information becomes activated, thereby forming a different, idiosyncratic instantiation with each occurrence.

For this review, we adopt Connell and Lynott's (2014) terminology and use *(conceptual) representation* to refer to a specific, situated, contextual instantiation of one or more concepts necessary for the current task, and the term *concept* to refer to a general, aggregated, canonical (i.e., context-free) aspect of experience that has the potential to be the basis of an offline representation. For example, the concept CHERRIES activates a somewhat different representation within each of the following two sentences (in italics): *I could tell from her lips that she had been eating cherries* and *Several rolled to the floor as I grabbed a handful of cherries*. The former is more likely to activate color information (CHERRIES are red), while the latter is more likely to activate shape information (CHERRIES are round).

Given that we are reviewing the neurocognitive literature, our access to behavioral or brain responses is limited to concept use in real time, that is, as a *conceptual representation* is being activated and constructed. We use the term *concept* to refer to the abstract form or the aggregated knowledge (which may or may not be stored in the mind and shared among speakers), to which we have no direct access in the empirical studies constituting the focus of our review. In this paper, we use the word *retrieval* to refer to the process that occurs upon reading a word; within a flexible conceptual representation framework, this process is typically considered *activation*: Words serve as cues that initiate activation and not as mental bookmarks that index an existing internal representation (Lupyan & Lewis, 2019).

According to a flexible conceptual representation framework, minds construct meaning in a flexible manner, perhaps activating only a subset of the aggregated information comprising a concept (Hoenig, Sim, Bochev, Herrnberger, & Kiefer, 2008; Kiefer, 2005; Kiefer & Pulvermüller, 2012). One early finding, introducing this idea of conceptual flexibility, was reported by Barsalou (1993) when the same participants asked to define the same concepts (e.g., "bird") two weeks apart exhibited only about 66% overlap in their definitions. This relatively large variance in definitions raises the possibility that individuals do not activate *all* the information they may possess about a concept every time they access it. Instead, on any given occasion, a person may access only a *subset* of their knowledge. On the view that concepts are represented by sets of features in a multidimensional space (McRae, De Sa, & Seidenberg, 1997; Plaut, 2002), conceptual flexibility could readily be explained by the activation of different subsets of features in different contexts (Barsalou, 1993; Lebois, Wilson-Mendenhall, & Barsalou, 2015; Yee & Thompson-Schill, 2016).

Nevertheless, some stability in conceptual representations (even if illusory) would seem to be necessary to anchor meaning to a concept and allow speakers to reach a mutual understanding of at least what is being said, if not what is meant by language elements: shared, negotiated, and agreed-upon definitions of concepts are stable and fixed by their practical nature. If representations were idiosyncratic and constructed anew according to fluctuating momentary demands and not anchored to any stable definition, how is it possible to ever understand what words mean? How can these seemingly opposing characteristics—flexibility and the apparent need for stable, shared definitions—be reconciled with each other and with experimental findings?

In this paper, we examine this question and some possible answers by reviewing experimental findings regarding concepts from the psychological, neuroscientific, and computational literatures, focusing on psycholinguistic and neurolinguistic findings. Assuming that conceptual representations are indeed flexible, we review published findings to better understand the nature of this flexibility: What characterizes flexible representations? How are they constructed? Which factors determine what representation will be activated? In Part 1, we identify some key characteristics of flexible conceptual representations. In Part 2, we contemplate how these flexible representations are constructed by reviewing some of the challenges, raised by us and others, to the flexible conceptual representations framework. In Part 3, we offer potential solutions to some of these challenges and highlight open issues in need of further testing. We rely on published works to propose that flexibility in representations may be achieved by updating the *degree* and *probability* of specific features' activations according to current processing demands, thereby incorporating the changing momentary context into the activated representation, while still preserving some of the concept's constituent meaning.

2. PART 1

2.1. Flexible representations are task-dependent and context-dependent

In the Introduction, we mentioned that the same concept CHERRIES can activate different representations in different sentences, each sentence highlighting a different feature of cherries. What factors determine which subset of a concept's features are activated on any given occasion? Several studies suggest that the momentary requirements posed by task and/or context can have a substantial influence on the conceptual representation activated and that these are evident relatively early, within the first few hundreds of milliseconds after word onset.

Task requirements seem to influence the representation constructed. For example, when comparing mean reaction times (RTs), words that refer to same-colored objects show priming effects (*emerald* primes *cucumber*)—that is, shorter RTs for similar-color pairs, compared to unrelated word pairs, when the priming task is preceded by a Stroop color naming task, but not when the priming task is succeeded by the Stroop task (Yee, Ahmed, & Thompson-Schill, 2012). This pattern of results presumably derives from the heightened attention to

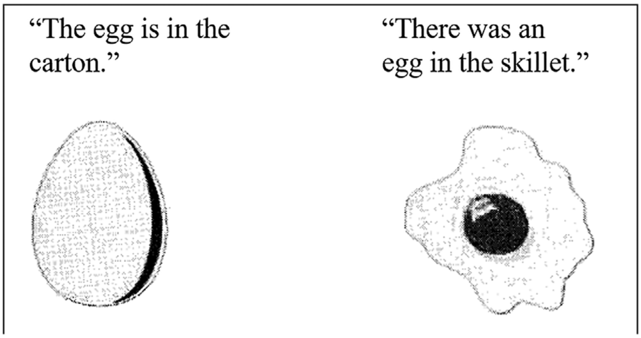

Fig. 1. Sentence–image pairs examples in Zwaan et al. (2002). The sentences imply different shapes of an egg: in a refrigerator versus in a skillet.

color promoted by the Stroop task performance. These results suggest that the color of an *emerald*, for example, is not necessarily activated whenever that conceptual representation is constructed but may be activated given recent experience with a relevant feature of the concept, in this case, color. Similarly, a flash of light, a noise, or a vibration on the skin occurring immediately before a sentence, such as "*broccoli is green*" or "*soup is hot*," leads to faster true/false judgment RTs when the sensory modality to which the sentence refers (visual, auditory, tactile) is the same as that of the preceding presumably attention-directing stimulus (a flash of light, a noise, or a vibration on the skin, accordingly; Van Dantzig, Pecher, Zeelenberg, & Barsalou, 2008).

Other studies have shown a similar influence of context on conceptual representations. For instance, in Zwaan, Stanfield, and Yaxley (2002), participants read a sentence that referenced the same object in different configurations: for example, *Eagle in the sky*, which implied a spread-winged eagle, and *Eagle in a nest*, which implied a seated eagle. Sentences were followed by a line drawing depicting a spread-winged eagle or a seated eagle. Participants judged whether or not the object had been referred to in the prior sentence (e.g., see Fig. 1). Even though a mental image of EAGLE was not required for task performance, the context presumably affected the activated mental representation of the eagle such that responses were faster when the pictured object's shape matched the object's shape implied by the sentence than when it did not.

Context also has been shown to affect part verification RTs (e.g., "Is X a part of Y?") in studies of adopted object perspective. In Borghi, Glenberg, and Kaschak (2004), participants were asked to adopt a perspective with respect to an object ("car"), for example, from the inside ("You are driving a car") versus the outside ("You are washing a car"), and then to make a part verification of a probe item typically found inside (*steering wheel*) or outside (*tires*) of the object. Part verification times for object internal parts were reliably faster when the inside (compared to outside) perspective had been adopted. The significant interaction of perspective by object part on verification times was taken to indicate an effect of perspective on the availability of conceptual knowledge.

Since conceptual representations cannot be meaningfully separated from the contexts in which they occur, a conceptual representation's meaning is intrinsically context-dependent (Yee & Thompson-Schill, 2016). In our example of CHERRIES, it is the context in which each representation occurs that drives the specific, context-dependent representation that is constructed. In sum, within a flexible conceptual representational framework, momentary demands of task and/or context seem to be the main driver for activating flexible representations.

2.2. *Distributed and componential neural organization allows for the activation of flexible representations*

Numerous neuroscientific studies have mapped conceptual representations onto neuronal activity in order to determine how conceptual knowledge is represented in the brain. Overall, conceptual representations seem to be componential in nature and widely distributed across the brain, that is, different components or features of a concept are represented separately in different brain areas. These two characteristics of the neural organization of conceptual knowledge can readily support flexible activation of conceptual representations, by allowing access to a *partial* representation of the concept, through selective activation of a subset of relevant features. If all that someone knew about a concept were stored altogether as an indivisible unit, the entire unit (i.e., all its contents) would need to be activated each and every time that concept was accessed, contra the desiderata for activating a conceptual representation flexibly.

A prevalent view that has emerged over the past two decades is that conceptual knowledge is systematically distributed across the brain, such that different aspects or features of a concept are associated with increased activity in the different sensory-motor brain regions that typically process those features (Allport, 1985; Barsalou, 1999; Damasio, 1989). Distributed Semantic Cognition (McClelland & Rogers, 2003; Rogers & McClelland, 2004), for instance, refers to a neural network in which various types of information about an object are stored in different brain areas, each responsible for the processing of that aspect of the input. On this view, a dog's representation is distributed across multiple brain areas: Its shape and color are stored in visual areas, its bark is stored in auditory areas, and the tactile sensation of its fur is stored in the textural (somatosensory) brain areas. Examination of brain activations with functional magnetic resonance imaging (fMRI) during conceptual processing supports this description. Attending to a specific feature of a concept activates the brain regions that correspond to the type of information that is being accessed: for example, thinking about actions with tools is associated with activity in action-related areas, and similarly information relating to the color, form, movement, and sounds of an object is accompanied by increasing activity in the fMRI blood-oxygen-level-dependent (BOLD) signal in the corresponding brain regions (Kan, Barsalou, Olseth Solomon, Minor, & Thompson-Schill, 2003; Kellenbach, Brett, & Patterson, 2003; Kellenbach, Wijers, & Mulder, 2000; for color: Martin, Haxby, Lalonde, Wiggs, & Ungerleider, 1995; Simmons et al., 2007; for action: Chao & Martin, 2000; for sounds: Kiefer, Sim, Herrnberger, Grothe, & Hoenig, 2008). These sorts of data have afforded maps of the various brain regions involved in the activation of specific features

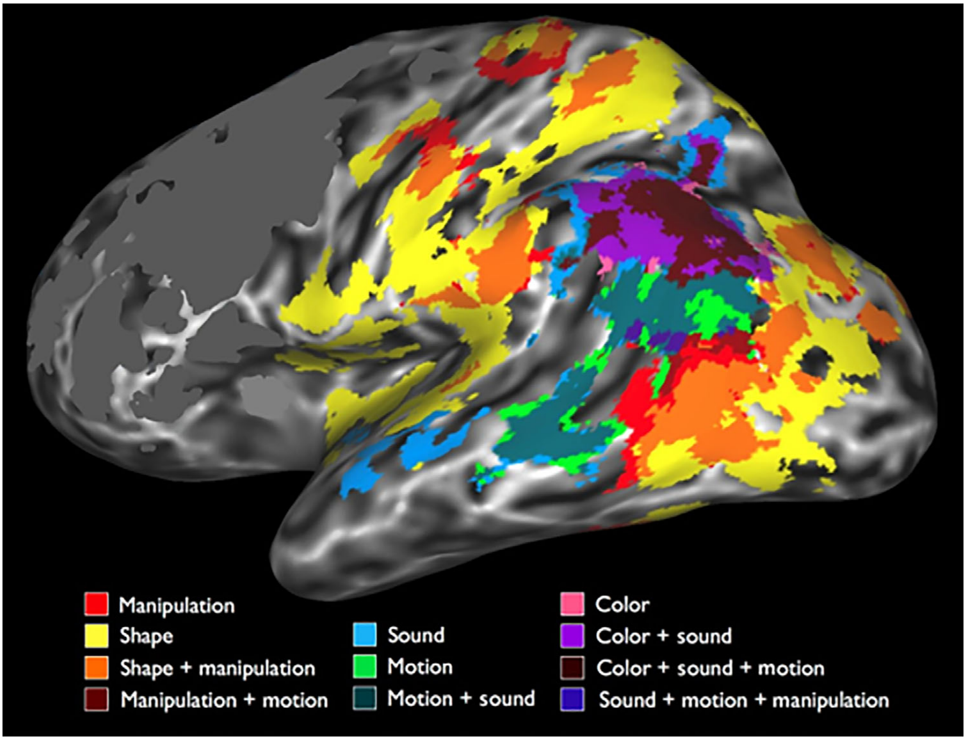

Fig. 2. Activation map of the various brain regions involved in the retrieval of specific features of objects. Adapted from Fernandino et al. (2016).

(sound, shape, color, motion, and others) of concepts, implicating a widely distributed neural representation of conceptual knowledge (see Fig. 2).

Distributed representations of conceptual knowledge in the brain could support flexible activation, by allowing access to only a *partial* representation of a concept, through selective activation of its task-relevant and/or context-relevant features. A distributed conceptual representation, however, does not necessarily require that activation be partial. In principle, all the interconnected information about a concept, distributed throughout the brain, could be activated to the same extent each and every time that concept is encountered and processed to some end. However, case reports of individuals with brain lesions who exhibit attribute-specific semantic impairments support partial activation of semantic information. Attribute-specific semantic impairments refer to an individual's inability to retrieve certain semantic attributes of an object from memory while still being able to retrieve other semantic attributes of the same object with relative ease (Basso, Capitani, & Laiacona, 1988; Coltheart et al., 1998; De Renzi & Lucchelli, 1994; Hart & Gordon, 1992; Sartori & Job, 1988; Silveri & Gainotti, 1988). Basso et al. (1988), for example, describe a person diagnosed with semantic dementia who presented with what was initially believed to be a category-specific impairment for natural things (e.g., animals, fruit, vegetables) together with intact knowledge of inanimate objects (e.g., man-made objects, means of transportation). This patient was able to name pictures of inanimate objects but not pictures of natural objects. Yet, when probed with two distinct semantic attributes of natural objects (the presumed impaired category)—a visual probe ("Does a camel have a smooth back or is it humpbacked?") and a non-perceptual probe ("Does a camel live in Italy?")— the patient was able to correctly answer the non-perceptual probe. This indicates that at least some knowledge about natural objects was intact. The total loss of some attributes of conceptual knowledge together with the preservation of other attributes to varying degrees is in line with the notion of *partial* activation of neurally distributed conceptual representations, which can account for selective partial activations that are created ad hoc and flexibly in real time. We revisit this issue in Part 3.

Functional theories of brain processing, such as *Representational Geometry* (Kriegeskorte & Kievit, 2013), show that selective partial activation may occur in healthy participants as well. In this method, a brain region is represented as a multidimensional space, where the dimensions of the space correspond to neurons, and a point in space corresponds to an activity pattern, that is, each neuron's activity provides the coordinate value for one of the dimensions. Representational geometry enables us to link cognition to brain activity by mapping different patterns of neural activity to different conceptual representations or categories of conceptual representations (Kriegeskorte & Kievit, 2013). In Nastase et al. (2017), participants viewed short video clips of animals in their natural environments. Participants were asked to respond to repetitions of either the same behavior (eating, fighting, running, or swimming) or the same taxonomy (primates, ungulates, birds, reptiles, or insects). The same stimulus generated different patterns of brain activity (inferred from fMRI) under different task demands. The authors proposed that semantic features are encoded in distributed neural populations and that attention may tune a feature space by dynamically altering population encoding and enhancing the discriminability of task-relevant information while collapsing task-irrelevant information, thus leading to different geometric patterns of neuronal activity for the same stimulus under different tasks.

Taken together, these studies suggest that conceptual representations are multidimensionally distributed, componential, and dynamically constructed in real time such that they can manifest partial activation of the information a person seems to have about a concept at any given time and support the construction of flexible, context-relevant conceptual representations. In our example of CHERRIES, a componential and distributed representation—in which color, shape, taste, and other features are represented in different brain regions—affords the construction of flexible representations according to varying tasks and/or context demands. Accordingly, we would expect to see stronger brain activity in color-related brain areas (as inferred from fMRI) when reading *I could tell from her lips that she had been eating cherries* and, alternatively, stronger activation in shape-related brain areas when reading *Several rolled to the floor as I grabbed a handful of cherries*.

2.3. Flexible conceptual representations linger

Conceptual knowledge seems to be modulated not only by the context in which it occurs but also by recent experiences that a person has had with a specific concept. According to

Connell and Lynott (2014), activation of a concept is “not a passive read-only access of memory, but actively alters what is being retrieved, both in strength and content.” On their account, encoding and retrieval are intertwined since “retrieval transfers memory into a state of transient plasticity” (Hardt, Einarsson, & Nader, 2010, p. 151) wherein conceptual content can be altered. The altered representation seems to linger for some time beyond its access such that it affects subsequent processing of the same concept.

Pecher, Zeelenberg, and Barsalou (2004) found that conceptual representations carry over (i.e., linger) from one encounter to the next, with manifest alterations due to the last access of the concept. In a feature verification task, participants were presented with a noun–feature pair: for example, *APPLE–green* (visual feature) or *APPLE–tart* (taste feature). Presumably, each of these noun–feature trials resulted in the construction of two different representations of APPLE: one that highlights visual information and another that highlights taste information. After a number of unrelated word pairs were presented, the noun was repeated but with another of its features, for example, *APPLE–shiny*. Feature verification times for the second feature *APPLE–shiny* (a visual feature) were faster and more accurate when it had been preceded by property verification in the same modality (visual: *APPLE–green*) than in a different modality (taste: *APPLE–tart*). This suggests that the initial conceptual representation of an object formed may linger for some time and facilitate the construction of a subsequent representation of the same concept when the relevant features (in this case, visual) of the recently encountered concept are still activated than when they are not.

Another observation consistent with the lingering effects of conceptual representations was described by Barclay, Bransford, Franks, McCarrell, and Nitsch (1974). Participants were presented with sentences highlighting a specific property of an object in order to determine how that experience with a concept affected its subsequent retrieval. For example, some participants read “*The man lifted the piano*,” which emphasizes the heaviness of pianos, whereas others read *“The man tuned the piano*,” which emphasizes the fact that the piano is a musical instrument. In a subsequent task, participants were asked to recall an object mentioned in the sentences, queried with “something heavy?” or “something with a nice sound?” Although all participants had encountered the word “piano” in one of the contexts, they were more likely to recall it in response to the query when it matched the implied feature of the object in the recently read sentence. These results were taken to suggest that participants formed different conceptual representations of *PIAN* following the different sentence contexts and that these representations lingered after their initial activation during the first task (sentence reading).

An explanation for this result by Connell and Lynott (2014) is that once a neuronal pattern associated with a conceptual representation has been activated in a manner determined by the attended (relevant) aspects of the experience, there is a bias to re-activate the same pattern upon the next encounter with that concept, even after a delay, and even if its activation may have little, if any, utility for the task at hand (Pecher, Zanolie, & Zeelenberg, 2007). In Pecher et al. (2004), verification of same-modality information resulted in shorter times, even with a lag of 12 or 18 trials between repetitions. The authors suggested that when a concept is presented with the first property type, a sensorimotor representation is activated so that participants can carry out the task of verifying the specific property with which it is paired. Focused attention to specific conceptual features could result in the activation of a modality-specific representation that would benefit property verification in the same modality a few trials later (as manifested in shorter verification times) due to a temporary increase in the availability of the conceptual features in the initially activated modality (Pecher et al., 2004, 2007).

Although it may not always be useful to have a recently constructed or recently modified conceptual representation re-activated, it could be beneficial during real-life interactions (e.g., conversations) when the same concepts are reused in temporal or physical proximity, that is, within a short time or essentially the same setting. For example, in an art class, when discussing the different shapes and colors of fruits in a bowl, there is no apparent benefit to activating the fruits' taste upon their every mention. An accessible conceptual representation of APPLE in these cases may be available selectively in a single modality (Pecher et al., 2007) and result in a partial representation that includes only the contextually relevant (attended) features. There could be an advantage to re-activating a readily available partial representation, as it would not require redundant "processing power" to construct the whole representation anew with its every mention. According to Huettig, Guerra, and Helo (2020), this phenomenon of "impoverished simulation" is compatible with the cost-effective Good Enough approach to language processing (Ferreira & Patson, 2007, Ferreira, Bailey, & Ferraro, 2002). On the Good Enough account, language processing is sometimes only partial, and semantic representations are often incomplete, and what is activated are representations just good enough for a comprehender to get the gist of an utterance. This account could explain why the partial representations that were created recently based on task and/or context demands are more easily accessed or are accessible for a longer duration after they are initially constructed: Perhaps this recency effect is more cost-effective during language processing in real time.

2.4. Selective activation of sensorimotor information results in flexible embodied conceptual representations

We mentioned earlier that reading texts or listening to sentences and discourse is accompanied by increased activity in the various brain regions that correspond to the sensorimotor features constitutive of the concept. One of the prominent issues of how concepts are represented in our brains relates to grounded cognition (Barsalou, 2008), which maintains that "the environment, situations, the body, and simulations in the brain's modal systems ground the central representations in cognition" (Barsalou, 2010). On this view, sensorimotor information is part of conceptual representations: Access to conceptual knowledge activates similar sensorimotor neurons as does perception, that is, concepts seem to be embodied. This view, however, remains contentious since embodiment effects seem to be inconsistent: Some studies show embodiment effects related to concepts during reading, but others do not. This inconsistent pattern of results raises the question of whether embodied activation of sensorimotor information occurs each and every time a concept is encountered. Flexible representations, which are context-dependent and characterized by selective activation of only part of the available conceptual information, could explain the inconsistent embodiment effects in the literature, by imbuing the embodiment theory with flexibility.

According to Barsalou (2003), *situated simulation* relies on the sensorimotor activation seen in the brain to words and suggests that conceptual processing uses "reenactments of sensory-motor states," or simulations, to represent conceptual knowledge. Some studies suggest that sensorimotor information is activated during reading, even if the simulation is not required by the task, as seen in the example of the visual features of *eagle in the sky* versus *eagle in the nest* (Zwaan et al., 2002). Other studies, however, suggest that perceptual features are not necessarily activated during reading if the task does not require it. For example, Pecher, Zeelenberg, and Raaijmakers (1998) examined whether words with similarly shaped referents primed each other as inferred from lexical decision times in a word-pair lexical decision task. Lexical decision times to similarly shaped words such as *coin–button* were not significantly different from those to unrelated (in shape) word pairs like *hoe–ball*. However, when participants first made shape judgments about the objects to which the word pairs referred, then there was a reliable facilitation in lexical decision times for similar-shaped words. The contradictory findings between the two studies raise the question of the extent to which sensorimotor information is activated every time a concept is encountered.

Several event-related brain potential (ERP) studies focusing on the N400 (Kutas & Hillyard, 1980) also report conflicting results regarding the grounding of conceptual representations (see text box for an explanation of ERPs and N400). For example, Amsel, DeLong, and Kutas (2015) examined N400s (during the 300-500 ms time window) to sentence-anomalous words that shared perceptual or motor features with the word that was the most likely continuation. Participants read a two-sentence context that set up an expectancy for a critical word: *My date was taking me to a romantic Italian restaurant for dinner tonight. I was worried that afterward I might reek of ___ so I brought gum.* The expected word at the position of the blank is *garlic*, and its contextually highlighted feature is its unpleasant smell. They observed an attenuation of N400 amplitude to an unrelated word (*tobacco*) that shared a perceptual feature with the expected word, relative to an unrelated word (*ice*) that did not show N400 attenuation. This result suggests that sensorimotor information about a concept is activated and affords facilitation of other concepts that share that sensorimotor feature, in line with a grounded cognition view.

Other studies, however, have failed to find a typical N400 modulation in concert with the activation of sensorimotor information. Rommers, Meyer, Praamstra, and Huettig (2013), for example, presented sentences that ended with expected/sensorimotor related/unrelated words: for example, *In 1969, Neil Armstrong was the first man to set foot on the* moon (expected)/*tomato* (shape-related)/*rice* (unrelated). No ERP difference was found between *tomato* and *rice* during the typical N400 time window (300-500 ms), although there was a statistically significant effect in a later window (500–700 ms).

Text box: Event-related brain potentials (ERPs) and N400

ERPs are used as a measure of neocortical processing, primarily summed post-synaptic potentials of pyramidal cells, recorded at the scalp in response to a time-locking signal. The N400 (component) is a negative-going potential relative to behind the ear (mastoid) between 200 and 500 ms peaking ∼400 ms post-stimulus onset, elicited by potentially meaningful stimuli (e.g., written or spoken words, pseudowords, sentences, pictures, video clips); it has been linked to semantic analysis (Kutas & Federmeier, 2011). N400 amplitudes are typically large for semantic anomalies (*I take coffee with cream and* _dog_—where *dog* is neither expected nor sensible), and smaller for semantically congruent stimuli as a function of their predictability and lexical association with anticipated words or words in the context (*I take coffee with cream and* _sugar_—more positive than the less predictable but associated word _saccharine_).

Several theories of the functional significance of the N400 have been proposed (Rabovsky & McRae, 2014), variously suggesting that N400 amplitudes reflect the effort or difficulty involved in semantic memory access (Van Berkum, 2009), the (mis)match between semantic features expected and encountered (Paczynski & Kuperberg, 2012), lexical access (Lau, Phillips, & Poeppel, 2008), semantic integration (Brown & Hagoort, 1993), semantic binding (Federmeier & Laszlo, 2009), or semantic inhibition (Debruille, 2007). For present purposes, we are less interested in the functional significance of N400 than in determining whether there is a difference in mean N400s amplitudes between experimental conditions that may suggest non-identical neural processing of the stimuli.

These, among other findings, have led to the question of whether sensorimotor information is a constituent part of the representation of concrete objects, all of which is activated each time the concept is encountered. Yee and Thompson-Schill (2016) suggest that flexible conceptual representations can allow for the flexible activation of features, and this flexibility is responsible for the seemingly contradicting results in the literature. Others (Lebois et al., 2015; Winkielman, Coulson, & Niedenthal, 2018; Winkielman, Davis, & Coulson, 2023) also suggest that grounded congruency effects are *dynamic* and *context-dependent*. Specifically what sensorimotor information, if any, is activated, seems to depend on task and/or contextual factors. This could explain why Rommers et al. (2013) observed no typical N400 attenuation for words that shared perceptual/motor features with the contextually expected word, whereas Amsel et al. (2015) did. The sentences in the latter study highlighted the relevant shared feature (*reek*, refers to the relevant aspect of *garlic*), whereas Rommers et al. (2013) study did not (no reference to the round shape of the *moon*).

Several studies have shown how context (like task) may modulate the activation of sensorimotor features, such as affordance, in a flexible manner. Affordances are the invitations to act that objects offer to people (Gibson, 1979)—such as the affordance of "sitting" on the object "chair"—and are considered a product of the conjunction of visual and motor experiences (Ellis & Tucker, 2000). Glenberg and Robertson (2000) showed that people could judge uncommon affordances, that were not highly associated with the relevant concept and with which they had no experience, as long as there was an appropriate context. For example, in the context of "*Marissa forgot to bring her pillow on her camping trip*," a substitute for a pillow made from an old sweater filled with leaves is judged as more sensible than an old sweater filled with water. Compared to human performance, a latent semantic analysis (LSA) language model based on the occurrence of words in similar contexts did not differentiate sensible sentences from non-sensible ones. This finding highlights the importance of the human ability to flexibly activate context-related sensorimotor information in order to make sense of novel utterances, a flexibility that the LSA model lacks.

Flexible activation of sensory information does not necessarily undermine a grounded cognition view but rather imbues it with flexibility—conceptual representations may include sensorimotor components, but as with other conceptual attributes, their activation in any given context may be more flexible than was once believed. Studies that have examined functional brain scans support this proposed explanation for the contradictory results on conceptual embodiment. Popp, Trumpp, Sim, and Kiefer (2019), for example, showed that fMRI activations in cortical areas when reading action verbs *(to throw)* and sound verbs (*to ring*) differ depending on the task. When participants were asked to determine the semantic relation between a noun and a verb (*ball–to throw / bell–to ring*), sound and action verbs induced different activations in corresponding sensorimotor areas. However, when participants encountered the same verbs in a lexical decision task, there was no reliable difference in activation between the sound and action verbs. Kemmerer (2015) suggests that although motor features of action-related words may be represented in motor cortices, their recruitment is task-dependent and may not be necessary for word comprehension in tasks that do not require attending to those features.

In a study examining the verbal and motor skills of individuals with motor deficits, Negri, Lunardelli, Reverberi, Gigli, and Rumiati (2007) likewise show that access to sensorimotor information may not be required for concept recognition. Some individuals were unable to perform or mimic specific object-related actions (such as the act of teeth-brushing), suggesting inaccessibility to some of the relevant sensorimotor information (e.g., *how toothbrushes are used*). Nevertheless, those individuals did retain their ability to recognize and name those objects (e.g., *toothbrush*). It seems, therefore, that although brains are populated with grounded conceptual representations, the activation of all features need not be obligatory.

A flexible conceptual representation framework, thus, may explain the inconsistency of activations in sensorimotor brain areas during word reading by suggesting that conceptual representations are flexibly embodied due to selective activation of sensorimotor attributes by context and/or task demands.

In Part 1 of our review, we identified key characteristics of flexible conceptual representations. In short, flexible conceptual representations are context-dependent, task-dependent, componential, and distributed across the brain. They can be activated partially through selective activation of varying subsets of their features. They linger. They manifest embodiment effects in a flexible manner.

3. PART 2

In Part 1, we discussed what characterizes flexible conceptual representations. In this next section, we will raise some of the open questions regarding conceptual flexibility and how flexible representation may emerge.

3.1. Reconciling flexibility with stability: Are there stable concepts, context-independent core features, or "coreless" ad hoc representations?

We began our review with Barsalou (1993) reporting 66% overlap in the definitions provided by the same individuals for the same concepts two weeks apart. Up to this point, our discussion has been focused on explaining the flexibility that allows for the 34% variance observed. In the following, we address the apparent stability of the conceptual system, as there was more overlap in definitions than variance. Which theoretical framework can account for both the evident flexibility and the perceived stability in the conceptual system?

The *degree* of flexibility of conceptual representations has long been a source of extensive debate (Dieciuc & Folstein, 2019; Kiefer & Pulvermüller, 2012; Löhr, 2017; Machery, 2015; Mazzone & Lalumera, 2010; Michel, 2020). Some researchers believe that "the same core information is activated each time an entity is encountered" (Ralph, Sage, Jones, & Mayberry, 2010), whereas others claim that these entities "have no cores at all" (Lebois et al., 2015), being constructed anew each time they are encountered.

Experimental findings and their interpretations are similarly equivocal. One view, sometimes referred to as *invariantalism* (Machery, 2015; Mazzone & Lalumera, 2010; Michel, 2020), proposes that conceptual representations are context-independent mental representations in long-term memory (Machery, 2015; Mazzone & Lalumera, 2010). In order to reconcile the evident conceptual flexibility with the required conceptual stability, one view that follows the outlines of invariantalism, controlled semantic cognition or CSC (Chiou, Humphreys, Jung, & Ralph, 2018; Hoffman, McClelland, & Lambon Ralph, 2018; Lambon Ralph et al., 2017), proposes that the semantic knowledge consists of two parts—a distributed *Semantic Representation System* and a more focal *Semantic Control System* that manages the activation of semantic representations and the division of labor. Conceptual flexibility, on this view, derives primarily from executive control processes, whereas the semantic representation system offers the primary source of conceptual stability. Following Connell and Lynott (2014) terminology (as we have), "semantic representations" in CSC parallels our use of the term "concepts," which refer to the aggregated, canonical aspects of experience stored in people's minds. In CSC, concepts are stored as stable units and thus are not the source of flexibility per se; flexibility is presumed to come from the control system. Data consonant with this proposal can be found in Chiou et al. (2018), who asked participants to pair concepts by their canonical colors (e.g., pairing *ketchup* with *fire extinguishers* because both are red) or by a more typical semantic relationship (e.g., pairing *ketchup* and *mustard* as both are popular condiments). Their fMRI results indicate that the *control system* (mainly in the prefrontal cortex) was more engaged by atypical (color) than typical relationship pairings (function), consistent with the proposal that conceptual flexibility is introduced by the control system. According to this view, the differential activation patterns in response to pairs under different task conditions are the result of task or context effects manifested in the varied activations in the control system and not part of the semantic representation system.

However, it is worth noting that Chiou et al. (2018) study also finds stronger activation in the brain areas considered part of the distributed semantic representation system (e.g., the visual-processing occipitotemporal areas) in the color-pairing task. This result would seem to undermine the assumption that the stored concepts are stable since there are changes in activations outside of the control system, modulated by task factors.

An alternative view along the flexibility–stability continuum maintains that words have conceptual cores that include typical and central information that is activated whenever meaning is processed (Barsalou, 1982, 1989; Whitney, McKay, Kellas, & Emerson, 1985). Barsalou (1982) empirically tested the proposal that concepts have context-independent (CI) properties that are activated on every encounter and context-dependent (CD) properties that are activated only in relevant contexts. On each trial, participants read a sentence in which the noun serving as the subject was underlined (e.g., *skunk; roof*). Several seconds later, a property was presented, and participants were instructed to indicate whether or not the subject noun possessed that property. The property could be CI (*has a smell* for *skunk*) or CD (*can be walked upon* for *roof*). The sentences preceding the property verification task either highlighted the property (*The skunk stunk up the entire neighborhood; The roof creaked under the weight of the repairman)*, or did not highlight the property (*The skunk was under a large willow; The roof had been renovated prior to the rainy season*), or were a control sentence for which the property was irrelevant. Verification times for the CI properties did not differ significantly for sentence-highlighted versus sentence-non-highlighted conditions, but for CD properties, the highlighted sentences yielded significantly faster verification times than the non-highlighted ones. Barsalou concluded that CI properties form the core meanings of words, while CD properties are a source of semantic variability, differing in how frequently they are associated with the word.

It is not clear, however, which features are context-dependent and which are context-independent on this view, or how these features can be classified as such a priori. This gap is especially notable when considering the conflicting evidence for the activation of similar types of features according to task or context demands. Several studies examined whether words whose referents have the same color (*emerald–cucumber*) or the same shape (*pizza–coin*) show priming effects compared to unrelated words, and found contradictory results. Some studies show that priming effects for two words whose referents have the same color or the same shape were evident only when the experimental task directed participants' attention to these features, implying that the activation of shape or color information is context- and task-dependent (Pecher et al., 1998; Yee et al., 2012). Other studies, however, that examined responses to similar types of relationships between words, found a similarity effect even without an explicit task that directed participants' attention to the shared features, undermining the proposal that activation of shape or color is necessarily context- or task-dependent (Rommers et al., 2013; 2015; Schreuder, Flores d'Arcais, & Glazenborg, 1984). These contradictory results underscore the need to have a clear a priori assignment to context-independent and context-dependent features, in order to make appropriate predictions to test this hypothesis.

Finally, at the other end of the spectrum of the degree of flexibility sits *contextualism*, which claims that representations are context-dependent, unstable entities, that have no cores that are necessarily activated (Casasanto & Lupyan, 2015; Lebois et al., 2015; Löhr, 2017; Ludlow, 2014; Yee & Thompson-Schill, 2016). The Ad Hoc Cognition view, for example, champions coreless conceptual representations, claiming that concepts, categories, and meanings are not fully formed entities sitting in storage ready for immediate use when called upon but rather are constructed ad hoc upon each instantiation. On this view, words serve as retrieval cues, and concepts have no stable core that is obligatorily and automatically activated regardless of context, given that no concept has any feature that is both necessary and sufficient for its definition. For instance, although most typical living tigers are large and furry, a small sculpture of a tiger may be neither, and yet it still can rightfully be referred to as a "tiger." Lebois et al. (2015) likewise maintain that word representations do not have conceptual cores and that even the most salient features of a word's meaning are not obligatorily activated. Instead, all conceptual features are presumed to be flexible and context-dependent. Importantly, the ad hoc view maintains that the "core serves only as a security blanket to make linguists and philosophers feel better about their concepts" (Casasanto & Lupyan, 2015; citing Greg Murphy, p. 552). The main premise of this view is that all concepts are created ad hoc each time we use them; conceptual cores do not exist as they are neither stable nor context-independent, contrasting with Barsalou's proposal that at least some features are context-independent.

Many of the findings discussed earlier, among others, seem to be consistent with the coreless *Ad Hoc Cognition* view of contextually dependent conceptual representations, including conceptual features that would seem to be salient or central to their meaning. For example, N400 amplitude for the animacy violation at the sentence end in "the peanut was **in love**," with no prior context, is large, compared to the sentence end in "the peanut was **salted**." However, when these same sentences are embedded in a discourse context describing an animated dancing peanut, which is happily in love, the pattern of N400 amplitudes is reversed: "was **salted**" elicits a larger N400 than "**in love**" (Nieuwland & Van Berkum, 2006). This finding indicates that discourse context can override highly associated information (such as the feature *salty* for *peanut*), counter to the proposal that there are context-independent core features that are equally activated regardless of the context.

However, as compelling as it may be, a view based on "coreless" conceptual representations cannot explain Barsalou's (1982) and others' (Whitney et al., 1985) findings in which context had no effect on how quickly knowledge about context-independent properties is retrieved. Proponents of a stable core often raise this lack of context effect on CI features as challenging for contextualism (see Machery, 2015; Mazzone & Lalumera, 2010), on which "retrieval from long-term memory is always context dependent" (Machery, 2015, p. 576).

Moreover, also contrary to the "coreless" view, conceptual representations do seem to possess some highly associated information that is often activated across and within participants, across different contexts. For instance, there is a wide agreement in participants' lists of features for a large number of concepts (McRae, Cree, Seidenberg, & McNorgan, 2005), as well as in lists of words related to given concepts (Nelson, McEvoy, & Schreiber, 2004). Even the agreement that a sculpture of a tiger—though neither furry nor large—is referred to by almost everyone as "tiger" which challenges the idea that concepts are "coreless," amorphous entities, constructed anew on the fly at each use. It seems that there are some shared and perhaps definitional conceptual features that allow humans to communicate. The Ad hoc view regards the emergent statistical pattern that points to a core as "a figment of the cognitive scientist's imagination" or an "illusion" (Casasanto & Lupyan, 2015, p. 552). According to them, it is from concepts (aggregate pattern), not conceptual representations (instantiations), that an illusory stability and sometimes an illusory core of context-independent features emerges.

They take this stability to be an abstract, statistical pattern that is epiphenomenal and does not reflect any isomorphic structure in the brain that houses a conceptual representation with features, core or otherwise, that can be accessed, activated, or retrieved. Future work, therefore, according to Casasanto and Lupyan (2015), should explain how apparent stability emerges from pervasive variability.

In summary, a strictly context-dependent view of conceptual activation has yet to successfully account for the lack of context effect on context-independent features of concepts. Likewise, the classification of some features as belonging to a stable, context-independent core also is insufficient to account for the results in all cases when it has been tested directly. It seems then that a debate about the existence versus non-existence of automatically activated core features is fruitless, as neither stance can account for all the extant findings. In Part 3 of the review (Sections 4.1-4.3), we consider a possible account for these seemingly contradictory findings.

3.2. Do flexible conceptual representations have flexible relations?

One of the questions that we would like to raise is the nature of the relations between flexible representations. Since there is no clear consensus, as detailed in the previous section, which parts or features of a concept are stable and context-independent, and which are flexible and context-dependent, it may likewise be unclear whether the relations between conceptual representations are stable or flexible. Although implied, most of the neurocognitive theories of conceptual flexibility do not directly address the nature of the relationships among flexible conceptual representations. Are these relationships stable or flexible? Does the degree of similarity (i.e., overlap in features) or relatedness change as the representations of the concepts themselves change? Does the activation of a conceptual representation prime or facilitate the activation of a fixed set of related or associated words, or are different words primed on each occasion, at least to some extent, depending on the specific conceptual representations activated?

In semantic priming experiments, participants are presented with word pairs that are followed by a task, usually a lexical decision. Typically, participants respond faster when the prime target are related and/or associated (e.g., *nurse–doctor*), compared to when they are unrelated (e.g., *nurse–bread*; Neely, 1977, 1991). This phenomenon has been classically assumed to reflect the organization of semantic memory and the relations between concepts (Fischler, 1977; Lucas, 2000). It is important to note that these classical models treat concepts as stable units and their relations as fixed and have not addressed the possibility of flexibility either within conceptual representations or their interrelations, including the idea of changing relations among flexible representations.

Some findings, however, do support the flexibility of the relationships among words. In some studies, p.riming effects for two words whose referents are the same color (e.g., *cucumber–emerald;* Yee & Thompson-Schill, 2016) or the same shape (e.g., *coin–button*) (Pecher et al., 1998) were observed only if the experimental task directed participants' attention to these conceptual features. This shows that priming effects for the same two words can change in size or scope by activating different representations of the same concept, which affords the suggestion that flexible conceptual representations may have flexible interrelations: that is, different sets of words that share some similarity or associative relationship with the activated conceptual representations can be primed or facilitated under various circumstances.

Nevertheless, the majority of the neurocognitive models of flexible conceptual representations do not make explicit predictions about different types of relationships among flexible conceptual representations. For instance, it is not clear what context-dependent and context-independent features a la Barsalou (1982) would predict about the relationships between flexible representations. Would concepts that share a context-independent feature exhibit a similar degree of priming regardless of task or context, while concepts that share a context-dependent feature exhibit a difference in relatedness and/or priming between their flexible conceptual representations? For example, given that "smelly" is a salient feature of SKUNK and perhaps also of GARLIC, is it the case that priming effects between these two concepts will not be affected by whether or not the recent context highlighted the "smelly" feature? And alternatively, will ROOF and CURB show priming effects (compared to unrelated word pairs) only if the recent context highlighted their non-salient shared feature "can be walked upon?" Likewise, in controlled semantic cognition, a static representational system is assumed. Does this mean that concepts can be operationalized flexibly while their interrelationships with other concepts remain stable? As for the Ad hoc cognition view that posits that all information is context-dependent, is it the case that SKUNK–GARLIC, though sharing a salient feature, will nevertheless show varied facilitation effects according to whether or not the context highlights their shared features? Or, perhaps, relationships among concepts are determined completely separately and independently of conceptual representation construction.

This discussion can be distinct from that about the nature of concepts and conceptual representations as entities, as it is not clear what might affect the functional and/or anatomical organization of semantic space, and whether relations between representations are also flexible to any extent, in any way. It may be that the predictions of upcoming words and priming effects depend on the changing relationships between flexible representations; or, alternatively, it may be that although representations are flexible, their relative locations in the semantic space are stable. Presumably, focused research will tell.

3.3. Is feature activation all or none?

Another issue with the extant neurocognitive models that we would like to raise is that, by and large, they do not discuss the degree of activation of a concept's properties, features, or attributes: Activation of a feature is typically described dichotomously, either activated or not (Barsalou, 1982; Yee, 2017) or automatic versus context-dependent (Lebois et al., 2015). Most current research examines the activation of a single feature, such as color (Connell, 2007; Yee et al., 2012) or shape (Pecher et al., 1998; Rommers, Meyer, & Huettig, 2015; Rommers et al., 2013; Zwaan et al., 2002). Even in the few studies that examined more than one feature, there are typically just two placed in opposition to each other, such as context-dependent versus context-independent features (Barsalou, 1982) or as features relevant versus irrelevant for constituting a conceptual representation (Hoenig, Sim, Bochev, Herrnberger, & Kiefer, 2008).

The underlying assumption of these experiments seems to be all-or-none feature activation, since there is no comparison of many features' activation: the features are not organized along any scale, and their activation is not discussed in any terms other than happening or not. However, it seems unlikely to us that this is a valid assumption, as studies report one of two patterns of opposing results. As discussed above, some studies show that a shared property can lead to perceptual similarity effects only when the shared property of the concepts is attended (Pecher et al., 1998; Yee et al., 2012), whereas, in others, this similarity effect for a shared property is observed regardless of attention (Huettig & Altmann, 2007; Rommers, Meyer, & Huettig, 2015; Rommers et al., 2013).

In addition, feature verification of a concept seems to be graded, regardless of context effects. In Vivas, Kogan, Yerro, Romanelli, and Vivas (2021), participants performed a concept-feature verification task. The features were organized into three categories according to their frequency and order of production in a separate feature norming study: (1) core features—features that were produced by most people to describe the specific concept, (2) partially shared features—features that were produced by only some participants, and (3) idiosyncratic features—which were produced by very few participants. The fastest verification was observed for core features and the slowest RTs for idiosyncratic features. It seems that, even without the modifying effects of context, different features are verified at different rates, presumably based on the importance of their contribution to the concept's meaning (more on the topic of features' varying contribution to a concept's meaning in Section 4.1). Meaning, even outside the framework of task-driven or context-driven flexibility, feature activation is not all or none.

Taken together, these types of activation patterns challenge the assumption of a dichotomous state of feature activation, in which some properties are context-independent and invariably activated, and others are activated only when the context requires it and then to the same degree. Instead, flexibility in the conceptual system may be expressed as variability in the degree of activation of a concept's features. We elaborate on this idea in Section 4.2*: Graded activation*.

In Part 2, we have considered how flexible representations are constructed by examining some of the open questions within the conceptual flexibility framework. How can flexibility be reconciled with the apparent stability, seemingly required (by most albeit not all researchers, see Casasanto & Lupyan, 2015) in order to anchor meaning and reach mutual understanding? What is the nature of the relations among flexible representations? Are features activated in an all-or-none manner? In Part 3, we offer possible solutions to some of these questions.

4. PART 3

In this final part of this review, we discuss how the challenge of reconciling the evident flexibility and the apparent need for stable definitions of concepts raised in Part 2 might be resolved.

4.1. Context is not the only factor driving flexible representations, featural saliency also matters

In Part 2, we introduced the notion of "coreless," "ad hoc" concepts in which no feature contributing to the construction of a concept is context-independent. We discussed some findings that contradicted this view (e.g., *smells bad* for *skunk* seems to be context-independent; Barsalou, 1982, 1993), which had led to the suggestion that conceptual representations could have context-independent cores. However, even this view of context-independent, inflexible "cores" is at odds with data of contextual override even for conceptual features considered central or highly associated (e.g., *salty* for *peanut*; Nieuwland & Van Berkum, 2006), which, in turn, suggests that there are no context-independent features.

If all feature activations are context-dependent, how can we explain the observed lack of context effects on the activation of context-independent features in a principled manner?—that is, without assuming that there are context-independent features or that all the challenging findings reflect experimental design decisions. Moreover, if all concepts are "coreless" and all conceptual features are similarly context-sensitive and "unstable," how can we explain the emergence of the "illusion of a core?" Here, and in the next sections, we suggest a framework that can reconcile these seemingly contradictory results. We rely on previous work to provide evidence that some features may indeed make a relatively stronger contribution or have a greater weight in constructing a concept's meaning (this Section), and although there may be no entirely context-independent stable "core" definitional features, the differential import of features for constructing a concept matters: Features that are central to the meaning of a concept may be activated more strongly (Section 4.2), and/or may be more likely to be activated (Section 4.3) during the construction of conceptual representations.

Some neuropsychological observations may provide a clue regarding the greater value of some features over others in the conceptual representation, suggesting that representations are not completely "coreless" but rather have some features that are more central to constituting their meaning. As described earlier, some individuals with brain-compromising lesions exhibit what initially looks like a category-specific semantic loss, exhibiting significant difficulty naming animals but not man-made tools (Farah & McClelland, 2013; Hart & Gordon, 1992). Closer examination, however, revealed that these individuals did not have a category-specific semantic loss: Although they did struggle considerably in verifying animals' visual features, they were able to verify animals' non-visual features, suggesting that they had not lost all that they knew about the category of animals. Farah and McClelland (2013), among others, proposed that visual features are more salient in defining natural objects than in defining tools or other inanimate objects. They argued that the apparent specific category impairment was artifactual, resulting from the degree of association between the salient attribute and the semantic category under investigation rather than the category per se. Warrington and McCarthy (1987, p. 1273) offer a similar explanation: "different weighting values from multiple sensory channels will be important in the acquisition of different categories of knowledge and […] such differential weightings could be the basis of the categorical organization of systems in the brain subserving semantic knowledge." This explanation echoes Warrington and Shallice's (1984) "differential weighting hypothesis" of the relationship between the categorical deficits observed and the impairment in the featural knowledge most salient for defining different categories of objects. The feature types of greatest relevance were those known in the parlance of the Martin sensory/motor model as “core properties” or “semantic primitives” (Antonucci & Alt, 2011). On this account, some features are relatively more important or salient in constructing a concept’s meaning than others.

A similar notion has gained traction in the neuroimaging literature with healthy participants. Hoenig et al. (2008), for example, compared patterns of neural activation to highly relevant (dominant) versus less relevant (non-dominant) attributes for constituting an object’s conceptual representation. Across two experiments, they compared fMRI or electroencephalogram (EEG) responses to congruent and incongruent attribute-noun pairs. The attribute was either perceptual (“elongated,” “round”) or action-related (“cut,” “peel”), and the target nouns were either man-made tools (“knife”) or natural objects (“orange”). On a separate materials validation task, Hoenig et al. found that perceptual attributes were rated as highly relevant (dominant) for constituting the conceptual representation of natural objects but not of tools, whereas the reverse held for the action attributes. Greater fMRI activation was observed when the non-dominant attributes of each category had to be verified, that is, visual attributes for tools and action-related attributes for natural objects. EEG effects began as early as 166 ms post-stimulus onset. Over the left fronto-central scalp region, sensory N1 amplitudes were largest when action attributes had to be verified for a natural category, smallest for visual verifications of natural category items, and intermediate for visual verifications of artifactual categories. The authors suggested that lower activation in brain responses while verifying highly dominant attributes may reflect easier access to these features in these concepts.

Interestingly, when verifying a less dominant feature (perceptual for tools, and action-related for natural objects), fMRI still showed some activation in the brain regions associated with the dominant feature, that is, some activation in pre-motor (action-related) area when verifying visual features for tools and some activation in the temporal gyri (visual-perceptual related) when verifying action-related features for natural objects. The reverse was not true—there was no activation in action-related areas when verifying visual-perceptual features of natural objects and no activation in visual-perceptual areas when verifying action-related features of tools. This pattern would seem to suggest that highly associated feature information can still be activated to some degree, even when it is task-irrelevant (Machery, 2015). Hoenig et al. (2008) concluded that conceptual features contribute to conceptual representations to varying degrees in a flexible, context-dependent manner, depending on the task but also, critically, on the importance or salience of these features for a particular concept or category.

Other studies suggest that interference with highly dominant features of concepts can impede the activation of the concept itself (reviewed in Yee & Thompson-Schill, 2016). For example, in one study, participants named pictured objects while concurrently performing hand motions that were incompatible with the depicted objects being named (Yee, Chrysikou, Hoffman, & Thompson-Schill, 2013). Naming times were longer, and error rates were higher for the depicted objects typically handled with the hand (e.g., *pencils*) versus those that were not (e.g., *tigers*). Since the ability to name objects was disrupted by concurrent incompatible hand motions, manual interaction was viewed as a central feature of the mental representation of frequently manipulated objects. Another study compared sentence verification of objects' properties while participants held three tones or a picture in memory (Vermeulen, Corneille, & Niedenthal, 2008). Participants were slower and less accurate in property verification when the modality of the property matched that of the items being held in working memory and imposing a cognitive load. For example, RTs were longer and accuracy was lower when holding three tones in memory when verifying *Blenders can be loud* versus when verifying *Lemons can be yellow*. If some features of a concept are more central for constituting its meaning, then interfering with these features will impede the construction of the conceptual representation.

If interference with its dominant features impedes a concept's activation, then greater attention to its dominant features might facilitate its activation. And indeed, Bermeitinger, Wentura, and Frings (2011) found that when an independent task directed participants' attention to shape, priming for words referring to natural kinds was greater than those referring to tools, presumably because shape is a more dominant feature of natural kinds. In contrast, when the interspersed task directed attention to action, priming was greater for tools, for which action is known to be a particularly important feature.

If concepts are "coreless" (Lebois et al., 2015), and all feature activation is context-dependent, how can we explain the activation of task-irrelevant dominant features (Hoenig et al., 2008) or the lack of context effect on the retrieval timing of context-independent features (Barsalou, 1982; Whitney et al., 1985)? Moreover, even though a small sculpture of a tiger lacks some key characteristics of living tigers, most speakers will nevertheless use the same label, "tiger," to refer to it and will be understood by other speakers in doing so. Therefore, it seems to us that context is not the only driving factor determining what conceptual representation is activated. The differential importance of features in constructing a concept needs to be seriously considered: Featural salience also matters and determines what conceptual representation is activated. Some features are more central to constructing a concept's meaning, and it may be this variable importance that creates the "illusion of a core."

However, it may be that there are no features that are *necessarily* activated in every context, regardless of their importance to communicative or learning aims. A better model to describe feature activation during conceptual representation construction may be one that considers the *probability* and the *degree* of activation of a feature in any given context. Perhaps features deemed as more central to a concept's meaning are more likely to be activated, and/or to be more strongly activated, than less central features. We elaborate on these two potential outcomes in Sections 4.2 and 4.3.

4.2. Conceptual features may be activated in a graded manner

We noted in Section 3.3 that the underlying assumption of most neurocognitive models is that feature activation is all or none, and stated that we think that this view may not find much support, as is already evident in contradictory data (Pecher, Zeelenberg, & Raaijmakers, 1998; Yee, Ahmed, & Thompson-Schill, 2012; vs. Rommers, Meyer, & Huettig, 2015; Rommers et al., 2013; Huettig & Altmann, 2007) as well as in the graded pattern of feature activation observed even in the absence of contextual modulations (Vivas et al., 2021). We suggest, instead, that feature activation may be graded, and perhaps affected by the feature's contribution to a concept's meaning, while still subject to the specific demands of the context.

One alternative to all-or-none activation can be found in Barclay et al. (1974), who entertain two types of flexible selective activation. *Partial activation* refers to activation limited to the contextually relevant features of a concept. G*raded activation* allows for the activation of many conceptual features but to varying degrees as a function of some weighting mechanism. Each of these types of selective activation can support a flexible conceptual system, but while partial activation suggests all-or-none activation of only attended conceptual features, graded activation allows for activation of non-attended features as well, albeit to a lesser degree. Graded activation also provides a means for reconciling the contradictory results in the literature: Varying degrees of activation of task-irrelevant features can produce inconsistent results across different studies, depending perhaps on the specific features' importance for the specific concepts that were included in each study.

Perhaps graded activation can also account for relationships among the conceptual representations. In computational language models, the relation between two words, referred to as edge, can be represented by a number on a continuous scale as the strength of association. The strength of the association is termed the weight of the edge. It is well established that the strength of the association between two words, such as the degree of overlapping features or forward association strength, affects priming (see review: Hutchison, 2003). Several studies in computational models, for instance, have found that adding edge weights to association networks, rather than just signifying whether two words are associated or not in a dichotomous manner, improves computational model performance in mimicking human experimental data (De Deyne & Storms, 2008; De Deyne, Navarro, & Storms, 2013). This suggests that relations among words may also be represented in a graded fashion in the human mind/brain.

As mentioned in the previous section, different features have various degrees of importance for a concept, and this importance is manifest in the representations activated. This importance may be expressed in the *degree* of activation of each feature. Note that this proposal shifts the debate from the existence versus non-existence of core features to a question about the degree of activation of each feature. On this view, constituent features of a concept are more likely to be more strongly activated, while still subject to specific dynamic contextual demands.

This possibility has not yet been extensively tested. A strong test of this proposal calls for research with more conceptual features on a graded scale and including the contextual effects of feature activation. We elaborate on this proposal in our Summary and Discussion section.

4.3. Bayesian updating of prior information according to current demands allows for the construction of flexible representations

We have discussed one of the biggest challenges faced by proponents of flexible conceptual representations, namely, how flexible representations can be constructed according to dynamic changes in momentary demands while still preserving the definitional features central to a concept's meaning. Perhaps a better way to discuss features' activation, specifically "core" features, is not in deterministic terms—whether salient or dominant features are necessarily activated every time a concept is encountered—but rather in probabilistic terms. As it happens, proponents on either side of the representational flexibility/stability divide mention probability as an organizing factor. Even though the ad hoc view rejects the idea of an internal representation with "core" features, Casasanto and Lupyan do suggest that "the more frequently a piece of information is activated in response to a cue [...] the more core-like that information may appear" (Casasanto & Lupyan, 2015, p. 551). Conversely, although rejecting a view of ad hoc representations constructed on the fly and supporting a view of stable cores, Mazzone and Lalumera claim that "stored information has a probabilistic structure" (Mazzone & Lalumera, 2010, p. 59). In this section, we would like to highlight this suggestion as the possible solution that shifts the debate from the existence of a definitional stable core that is obligatorily activated to a discussion about the *degree* and *probability* of activation of the information constituting a conceptual representation.

One model that has been proposed in this context is the Bayesian update of prior conceptual knowledge (such as featural saliency) according to task or context demands (Barsalou, 2016; Lebois et al., 2015; Michel, 2020). Empirical work in the field of language processing, and cognitive processing in general, suggest that "the brain is Bayesian," claiming that although cognitive processing is not necessarily optimal, it is able to deal with uncertainty by considering both the information it has acquired up to the moment and the current task demands and their integration in a probabilistic manner (Chater, Oaksford, Hahn, & Heit, 2010; Knill & Pouget, 2004). Bayesian modulation of cognition in real time has been implicated in word learning (Chater & Manning, 2006; Xu & Tenenbaum, 2007) and semantic memory (Steyvers, Griffiths, & Dennis, 2006), as well as visual perception (Mamassian, Landy, & Maloney, 2002; Orbán, Fiser, Aslin, & Lengyel, 2008), sensory-motor control (Körding & Wolpert, 2006), among other cognitive tasks.

When applied to conceptual representations, Lebois et al. (2015) suggest that the flexible activation of conceptual information reflects Bayesian sampling: The probability that a given feature of a concept is activated reflects (1) the overall frequency with which that feature has been processed (its prior) and (2) its relevance in the current context (its likelihood) (see also Barsalou, 2016). To better understand how context and features interact, we can revisit Nieuwland and Van Berkum (2006) who showed that "*the peanut was* ***in love***" and "*the peanut was* ***salted***" were processed differently as a function of the discourse context in which they appeared. Without any immediate discourse context, "the peanut was **salted**" showed attenuated N400 amplitudes time-locked to **salted**, suggesting greater activation of the feature relatively more strongly associated with peanuts compared to "**in love**"—an animacy violation not typically associated with peanuts. Within a larger discourse context, revolving around a dancing peanut who has happily fallen in love, the pattern of N400 amplitudes to these two sentences' final words are reversed.

To determine which factors lead to the inverted predictive word probabilities and the associated N400 amplitude inversion, Werning, Unterhuber, and Wiedemann (2019) compared three computational (quantitative) models of predictive probability of a word given prior discourse: (a) The *Semantic Similarity Model* based on distribution semantics, determined by statistical regularities of co-occurrences in large corpora; (b) the *Relevance Model* on which context relevance is the only factor guiding lexical prediction (proxy for processing of target utterances); and (c) the *Bayesian Pragmatic Model* on which both

Table 1
Examples of the four conditions compared in Werning et al. (2019)

	+Similarity	−Similarity
Std Ctx	*Maria prepares a cake buffet to impress her friends and makes ready everything necessary for it. She is already about* ***cream*** *to* ***whip***	*Maria prepares a cake buffet to impress her friends and makes ready everything necessary for it. She is already about* ***cream*** *to* ***draw***
NewCtx	*Maria practices for a picture of a cake buffet and uses her notebook for her preliminary study. She is already about* ***cream*** *to* ***whip***	*Maria practices for a picture of a cake buffet and uses her notebook for her preliminary study. She is already about* ***cream*** *to* ***draw***

Note. The sentences in the experiment are in German, this translation to English preserves the original word order of the final sentence.

prior information about between-word similarity (through co-occurrence) and context relevant information (e.g., discourse relevance), which can update the prior with its likelihood, affect the processing of a target utterance. Werning et al. (2019) compared the predictions of these three models to human data in a multiple-choice cloze task in which participants chose the word they thought would appear next in a sentence and to their N400 amplitudes for the four experimental conditions. The 2 × 2 condition set interlaced semantic similarity (based on co-occurrence) between words—for example, "whip" and "cream" are more "related" than "draw" and "cream"—with context relevance—for example, a context about painting a picture of a cake and a context about baking a cake (see Table 1). Multi-linear regression analyses revealed that while each model explained some variance in the human data from each experiment, the Bayesian Pragmatic Model explained the largest proportion and outperformed the other two models in accounting for both the human cloze probability and the N400 amplitude ERP findings. These results indicate that both factors—prior semantic knowledge *and* ongoing processing context and task demands—contribute to context-guided modulation of lexical meaning and comprehension and propose a mechanism: context-driven updating of priors that allows for the construction of flexible representations.

Hence, a Bayesian model assumes relevance-guided modulation of a word's lexical meaning through "Bayesian updating of learned statistical regularities stored in semantic memory" (Werning et al., 2019). This Bayesian update of prior information can explain how featural activation is modified as a function of context: Frequent features have a high chance of becoming activated, yet context can override this, thereby advantaging contextually relevant features. As context becomes increasingly salient and specified, contextually relevant features may become increasingly salient (Barsalou, 2016; Lebois et al., 2015). A possible mechanism for achieving flexible representations starts with a differential assignment of weights to various features of a concept as suggested by Barclay et al. (1974) and Warrington and Shallice (1984) who proposed some feature weighting mechanism in which higher weights can be assigned to stronger priors, that is, to more salient features that contribute more strongly to a concept's meaning. Then, these weights are updated according to context: The activation of prior information is adjusted according to contextual (recent or concurrent) demands. By adjusting the weights assigned to different pieces of information we have about a concept according to contextual demands, our minds/brains can construct context-sensitive and task-sensitive flexible representations while still preserving some of the constituent features of the concept.

One possible way to quantify features' saliency or importance in constructing a concept's meaning is by looking at the production frequency of different features for a concept. McRae et al. (2005) Feature Norms include feature lists for concepts, organized by production frequency, that is, the proportion of participants who listed that specific feature for a given concept. For example, the feature "red" for CHERRIES was produced by 27 out of 30 participants, while the feature "round" was produced by 15 out of 30 participants. If we apply the Bayesian model to our CHERRIES example, the prior information about features' saliency assigns a higher weight to "red" than to "round" since the former may be seen as a more salient feature. These prior weights are then updated according to context: In the sentence *I could tell from her lips that she had been eating cherries*, more weight is assigned to the context-relevant feature "red," whereas in the sentence *Several rolled to the floor as I grabbed a handful of cherries*, more weight is assigned to the context-relevant feature "round." As a result, we can predict what the representation of CHERRIES is likely to be in each of these sentences: In the first sentence, it is more likely that the representation is highly focused on CHERRIES' color rather than their shape, or the typical place where cherries can be found, or their taste, and so forth. In the second sentence, we would expect a representation that assigns more weight to the shape of CHERRIES but also some, perhaps smaller, weight to their color since the color feature has a strong prior weight. In such a manner, our minds are able to incorporate the momentary task or context demands into the current representation while maintaining some of the more salient features of the concept.

It is important to note that we are not claiming that feature dominance or salience is the only type of prior information activated by contextual demands. Priors include all previously learned information related to a concept: its features, the environments in which it occurs (linguistic and physical), and pragmatic information, among other information. The major claim is that the probability that a feature will be activated results from updating the prior saliency of previously learned information with its likelihood given contextual relevance. These two interacting factors—prior knowledge and relevant context—update the priors and determine what the current representation of a concept becomes.

The Bayesian account affords a compromise between a "coreless" view that states that "congruency effects should only occur when task conditions make relevant features salient" (Lebois et al., 2015, p. 1790) and contradictory evidence for activation of task-irrelevant and/or context-irrelevant features (e.g., *smells bad* for *skunk*, Barsalou, 1982, 1993; Whitney et al., 1985; or task-irrelevant features in Hoenig et al., 2008). It provides a single framework to explain how all features can be context-sensitive on the one hand and still preserve the concept's constituting meaning on the other—by assigning higher prior weight to central features that constitute the core meaning of a concept and updating all weights according to task and context demands. This view can explain both the 34% variability and 66% overlap in concepts' definitions found in Barsalou (1993) that were provided by the same participants two weeks apart. Bayesian update of prior information shifts the debate from the existence versus non-existence of a definitional conceptual core that is always activated to a discussion about the *probability* and *degree* of activation of the information constituting a concept, thereby reconciling representational flexibility with the need to anchor meaning in shareable definitions that form the core meaning of a concept.

5. Summary and discussion

How do our cognitive faculties make use of a set of fixed labels attached to concepts to provide a veridical representation of a constantly changing reality? In this review, we have examined findings from a whole host of cognitive experiments conducted with healthy and clinically compromised individuals, employing a variety of methods, to better understand how flexible conceptual representations are formed and used by our minds in real time.

We have come to the view that our minds represent the ever-changing world by constructing ever-changing flexible representations: The fixed label attached to a concept seems to activate a different representation at each use. We began our review by listing key characteristics of flexible representations. Flexible conceptual representations are context-dependent and task-dependent. They are componential—constituted by a variety of different components that are systematically distributed in different brain areas (including sensorimotor, among others). These characteristics of the concepts and their functional organization allow for partial selective activation of the knowledge one possesses about any given concept. Flexible conceptual representations also linger beyond their initial activation and can thereby exact measurable consequences on the same concept on a subsequent encounter for some limited time period. This sort of flexibility offers a potential explanation for why activation in sensorimotor brain areas is not always evident during word reading: When embodied features of conceptual representations are selectively activated from among the entire set of possible featural components by context or task demands, they are likely to result in a flexible embodied representation. Otherwise, if the modifying factors do not selectively trigger activation of the sensorimotor components, the conceptual representation will not necessarily carry evidence of its embodied aspects.

In Part 2, we examined several open questions regarding the flexibility of conceptual representations. How is it possible to understand what words mean if they are not anchored to any immutable or fixed definition? Some stability in the conceptual representation system would seem to be essential. We discussed several views on how conceptual representations, which differ in the degree of stability they entertain, come to be: These range from the idea that the (1) same "core information" is activated each time a concept is encountered (Ralph et al., 2010), (2) there are context-independent conceptual features that are stable and context-dependent conceptual features that are not (Barsalou, 1982), to the proposal that (3) representationals are "coreless" and all information activation is context-dependent (Casasanto & Lupyan, 2015; Lebois et al., 2015). We find that none of these accounts seem to explain all the empirical results in the literature. We also examined the implicit assumption that conceptual features or components are activated in an all-or-none manner and suggested that this seems unlikely given the graded activation of conceptual features by their saliency, even without the modifying effects of context. Last, we raised the question of how flexible representations affect or interact with each other, namely, whether the relations among flexible representations are also flexible, and if so, to what extent, noting that the nature of the interrelations among conceptual representations may be independent of the nature of the representations themselves.

In Part 3, we reconcile the evident flexibility of conceptual representations with the communicative need for stability. We maintain that (1) representations are not "coreless"—featural salience matters, (2) feature activation can be graded, and (3) Bayesian updating of prior information as a function of current demands can explain how flexible representations are constructed. We maintain that a complete reconciliation can follow if we shift the debate from whether there exists a definitional core for each concept that is always activated to questions about the *degree* and *probability* of the activation of the information that constitutes a concept. Concepts have meanings shared across speakers, usually consisting of central features that are salient to a concept's meaning. Literature on context-dependent and flexible representations sometimes seems to have neglected the significance of the varying degrees of salience of specific features. We think that constituent features are *more likely* to be activated and that interference with the constituent features is more likely to impede the activation of the conceptual representation, compared to interference with non-constitutive features. Critically, these central constituent features need not be stable nor context-independent. Indeed, salient contexts can override even central constitutive conceptual features.

Additionally, we think that the graded activation of features as an organizing principle of the internal structure of conceptual representations should be given serious consideration. To explain, we would like to draw a parallel between this idea and Barsalou's (1987) response to the proposal that concepts within categories are organized by typicality. Graded structure is evident in the order and the frequency with which category members are produced, even for novel ad hoc categories created to achieve a goal (e.g., the category of *things to pack in a suitcase*). This graded structure, according to Barsalou, "is a highly flexible and unstable phenomenon" (Barsalou, 1987, p. 113) since these categories are constructed anew and subject to changing conditions.

Perhaps a parallel may be drawn between categories and their concepts and our discussion of concepts and their features. Some features of a concept can be seen as more "typical" than others. The graded nature of features is evident in feature production studies. For example, McRae et al.'s Feature Norms (2005), which include a list of nouns and their features organized by the frequency of their production, represent wide agreement on the concepts' features' "typicality." For example, almost all participants produced the feature "yellow" for "banana," but only some said it is "eaten by monkeys." We propose that the importance of a feature can be represented not just by the order or frequency of feature production but also by its *degree* of activation during comprehension: We predict that constitutive features will be activated more strongly. If graded categories create instability or flexibility (Barsalou, 1987), perhaps flexibility in conceptual representations affords graded activation. It may be that flexibility and graded structure are coupled since graded structure allows for more variability than a dichotomous structure: A larger number of unique representations can be created when each feature can be activated to varying degrees, compared to the number of possible representations when each feature can assume only one of two states.

A model that includes an update of the *degree* and *probability* of specific feature(s) activation as a function of context allows us to describe how flexible representations that preserve some of their constituent meaning may be created by our minds in real time. A Bayesian probabilistic model can assign different weights to different features of a concept, based on the features' importance in constructing the concept's meaning (prior), and update their relative importance in a specific context accordingly in real time (likelihood; Barsalou, 2016; Lebois et al., 2015; Werning et al., 2019). We propose that the updating of different weights assigned to different features can theoretically describe not just the *probability* of activating a feature but also the *degree* of a feature's activation where relatively more salient features or context-relevant features will be more strongly activated. This proposal, thereby, provides a computational mechanism for constructing flexible, graded representations, in which the prior importance of features matters, such that a portion of the constituent information about a concept's meaning is preserved. One possible advantage of creating partial graded representation as a function of context, instead of activating all the information one possesses about a concept, may be that these "good-enough" representations are more cost-effective as introduced in Section 2.3.

It is important to note that other computational models, like the connectionist models, also have been able to capture many important characteristics of how language is learned, represented, and processed (Joanisse & McClelland, 2015). One advantage of Bayesian probability models, however, is that they allow for a greater degree of representational and structural diversity (Griffiths, Chater, Kemp, Perfors, & Tenenbaum, 2010). These models do *not* make a strong statement about the mental representations a person uses while performing a task (Jacobs & Kruschke, 2011). For example, in several experiments, participants performed a property induction task: They learned that three members of a category have a certain property and were asked whether other items in the category share this property (Atran, 1998; Rips, 1975). The results were compared to those obtained with different types of representations of the possible underlying structure of the category (Kemp & Tenenbaum, 2008). When the items referred to animals, inferences about novel properties were better supported by tree-structured representations of the category. Inferences about relationships between cities, however, were better captured by a low-dimensional space than a tree (see review: Griffiths et al., 2010). Griffiths et al. (2010, p. 359) claim that "unlike the connectionist approach, the probabilistic approach is open to the idea that qualitatively different representations are used for different types of inferences." This finding may provide an advantage to Bayesian probabilistic models over connectionist models. At the same time, however, we acknowledge that "people are not truly Bayesian, but only approximately Bayesian" (Jacobs & Kruschke, 2011, p. 9), and that other computational models also may offer a good description of the cognitive processes that underlie the construction of flexible conceptual representations that are stored in the brain or created ad hoc. A large body of work describes in detail the various ways that neural networks and other computational language models have been used to study human conceptual representations (see reviews: Castro & Siew, 2020; Kumar, Steyvers, & Balota, 2022); however, much less has been said about psycholinguistic and neurolinguistic findings to explain flexibility in representations, which is the focus of this review. Future work by computational modeling experts that directly compares different computational language models to explain and characterize flexibility would be most welcomed.

Further research is needed to answer the open questions and test the viability of our proposed solutions in this review. Yet it is our belief that the conceptual system as a whole should be treated as a flexible and (for)ever-learning system, in which meaning is constructed, as opposed to merely retrieved from stored memory. It is this ability that allows us to use the same label for a specific man and the same label for a specific river, while still acknowledging—by activating flexible idiosyncratic conceptual representations—that the same river cannot be stepped into twice, for it is neither the same man nor the same river.

References

Allport, D. A. (1985). Distributed memory, modular subsystems and dysphasia. In S. K. Newman & R. Epstein (Eds.), *Current perspectives in dysphasia* (pp. 32–60). New York: Churchill Livingstone.

Amsel, B. D., DeLong, K. A., & Kutas, M. (2015). Close, but no garlic: Perceptuomotor and event knowledge activation during language comprehension. *Journal of Memory and Language*, *82*, 118–132.

Antonucci, S. M., & Alt, M. (2011). A lifespan perspective on semantic processing of concrete concepts: Does a sensory/motor model have the potential to bridge the gap? *Cognitive, Affective, & Behavioral Neuroscience*, *11*(4), 551–572.

Atran, S. (1998). Folk biology and the anthropology of science: Cognitive universals and cultural particulars. *Behavioral and Brain Sciences*, *21*(4), 547–569.

Barclay, J. R., Bransford, J. D., Franks, J. J., McCarrell, N. S., & Nitsch, K. (1974). Comprehension and semantic flexibility. *Journal of Verbal Learning and Verbal Behavior*, *13*(4), 471–481.

Barsalou, L. W. (1982). Context-independent and context-dependent information in concepts. *Memory & Cognition*, *10*(1), 82–93.

Barsalou, L. W. (1987). The instability of graded structure: Implications for the nature of concepts. In U. Neisser (Ed.), *Concepts and conceptual development: Ecological and intellectual factors in categorization* (pp. 101–140). Cambridge: Cambridge University Press.

Barsalou, L. W. (1989). Intraconcept similarity and its implications for interconcept similarity. In S. Vosniadou & A. Ortony (Eds.), *Similarity and analogical reasoning* (pp. 76–121). Cambridge: Cambridge University Press.

Barsalou, L. W. (1993). Flexibility, structure, and linguistic vagary in concepts: Manifestations of a compositional system of perceptual symbols. In A. F. Collins, S. E. Gathercole, M. A. Conway, & P. E. Morris (Eds.), *Theories of memory* (pp. 29–101). Hove: Lawrence Erlbaum Associates.

Barsalou, L. W. (1999). Perceptual symbol systems. *Behavioral and Brain Sciences*, *22*(4), 577–660.

Barsalou, L. (2003). Situated simulation in the human conceptual system. *Language and Cognitive Processes*, *18*(5-6), 513–562.

Barsalou, L. W. (2008). Grounded cognition. *Annual Review of Psychology*, *59*, 617–645.

Barsalou, L. W. (2010). Grounded cognition: Past, present, and future. *Topics in Cognitive Science*, *2*(4), 716–724.

Barsalou, L. W. (2016). On staying grounded and avoiding quixotic dead ends. *Psychonomic Bulletin & Review*, *23*(4), 1122–1142.

Basso, A., Capitani, E., & Laiacona, M. (1988). Progressive language impairment without dementia: A case with isolated category specific semantic defect. *Journal of Neurology, Neurosurgery & Psychiatry*, *51*(9), 1201–1207.

Bermeitinger, C., Wentura, D., & Frings, C. (2011). How to switch on and switch off semantic priming effects for natural and artifactual categories: Activation processes in category memory depend on focusing specific feature dimensions. *Psychonomic Bulletin & Review*, *18*(3), 579–585

Borghi, A. M., Glenberg, A. M., & Kaschak, M. P. (2004). Putting words in perspective. *Memory & Cognition*, *32*, 863–873.

Brown, C., & Hagoort, P. (1993). The processing nature of the N400: Evidence from masked priming. *Journal of Cognitive Neuroscience*, *5*(1), 34–44.

Casasanto, D., & Lupyan, G. (2015). All concepts are ad hoc concepts. In E. Margolis & S. Laurence (Eds.), *The conceptual mind: New directions in the study of concepts* (pp. 543–566). Cambridge, MA: MIT Press.

Castro, N., & Siew, C. S. (2020). Contributions of modern network science to the cognitive sciences: Revisiting research spirals of representation and process. *Proceedings of the Royal Society A*, *476*(2238), 20190825.

Chao, L. L., & Martin, A. (2000). Representation of manipulable man-made objects in the dorsal stream. *Neuroimage*, *12*(4), 478–484.

Chater, N., & Manning, C. D. (2006). Probabilistic models of language processing and acquisition. *Trends in Cognitive Sciences*, *10*(7), 335–344.

Chater, N., Oaksford, M., Hahn, U., & Heit, E. (2010). Bayesian models of cognition. *Wiley Interdisciplinary Reviews: Cognitive Science*, *1*(6), 811–823.

Chiou, R., Humphreys, G. F., Jung, J., & Ralph, M. A. L. (2018). Controlled semantic cognition relies upon dynamic and flexible interactions between the executive 'semantic control' and hub-and-spoke 'semantic representation' systems. *Cortex*, *103*, 100–116.

Coltheart, M., Inglls, L., Cupples, L., Michie, P., Bates, A., & Budd, B. (1998). A semantic subsystem of visual attributes. *Neurocase*, *4*(4-5), 353–370.

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), 204–256.

Connell, L. (2007). Representing object colour in language comprehension. *Cognition*, *102*(3), 476–485.

Connell, L., & Lynott, D. (2014). Principles of representation: Why you can't represent the same concept twice. *Topics in Cognitive Science*, *6*(3), 390–406.

Damasio, A. R. (1989). Time-locked multiregional retroactivation: A systems-level proposal for the neural substrates of recall and recognition. *Cognition*, *33*(1-2), 25–62.

De Deyne, S., & Storms, G. (2008). Word associations: Network and semantic properties. *Behavior Research Methods*, *40*(1), 213–231.

De Deyne, S., Navarro, D. J., & Storms, G. (2013). Better explanations of lexical and semantic cognition using networks derived from continued rather than single-word associations. *Behavior Research Methods*, *45*(2), 480–498.

De Renzi, E., & Lucchelli, F. (1994). Are semantic systems separately represented in the brain? The case of living category impairment. *Cortex*, *30*(1), 3–25.

Debruille, J. B. (2007). The N400 potential could index a semantic inhibition. *Brain Research Reviews*, *56*(2), 472–477.

Dieciuc, M. A., & Folstein, J. R. (2019). Typicality: Stable structures and flexible functions. *Psychonomic Bulletin & Review*, *26*, 491–505.

Dry, M. J., & Storms, G. (2010). Features of graded category structure: Generalizing the family resemblance and polymorphous concept models. *Acta Psychologica*, *133*(3), 244–255.

Ellis, R., & Tucker, M. (2000). Micro-affordance: The potentiation of components of action by seen objects. *British Journal of Psychology*, *91*(4), 451–471.

Farah, M. J., & McClelland, J. L. (2013). A computational model of semantic memory impairment: Modality specificity and emergent category specificity (Journal of Experimental Psychology: General, 120(4), 339–357). In G. Cohen, R. A. Johnstone, & K. Plunkett (Eds.), *Exploring cognition: Damaged brains and neural networks* (pp. 79–110). Hoboken: Taylor and Francis.

Federmeier, K. D., & Laszlo, S. (2009). Time for meaning: Electrophysiology provides insights into the dynamics of representation and processing in semantic memory. In B. H. Ross (Ed.), *Psychology of Learning and Motivation* (Vol. 51, pp. 1–44). Waltham, MA: Elsevier.

Fernandino, L., Binder, J. R., Desai, R. H., Pendl, S. L., Humphries, C. J., Gross, W. L., Conant, L. L., & Seidenberg, M. S. (2016). Concept representation reflects multimodal abstraction: A framework for embodied semantics. *Cerebral Cortex*, *26*(5), 2018–2034.

Ferreira, F., & Patson, N. D. (2007). The ‘good enough’ approach to language comprehension. *Language and Linguistics Compass*, *1*(1-2), 71–83.

Ferreira, F., Bailey, K. G., & Ferraro, V. (2002). Good-enough representations in language comprehension. *Current Directions in Psychological Science*, *11*(1), 11–15.

Fischler, I. (1977). Semantic facilitation without association in a lexical decision task. *Memory & Cognition*, *5*, 335–339.

Gibson, J. J. (1979). *The ecological approach to visual perception*. Boston: Houghton Mifflin.

Glenberg, A. M., & Robertson, D. A. (2000). Symbol grounding and meaning: A comparison of high-dimensional and embodied theories of meaning. *Journal of Memory and Language*, *43*(3), 379–401.

Griffiths, T. L., Chater, N., Kemp, C., Perfors, A., & Tenenbaum, J. B. (2010). Probabilistic models of cognition: Exploring representations and inductive biases. *Trends in Cognitive Sciences*, *14*(8), 357–364.

Hardt, O., Einarsson, E. Ö., & Nader, K. (2010). A bridge over troubled water: Reconsolidation as a link between cognitive and neuroscientific memory research traditions. *Annual Review of Psychology*, *61*, 141–167.

Hart, J., & Gordon, B. (1992). Neural subsystems for object knowledge. *Nature*, *359*(6390), 60–64.

Hoenig, K., Sim, E. J., Bochev, V., Herrnberger, B., & Kiefer, M. (2008). Conceptual flexibility in the human brain: Dynamic recruitment of semantic maps from visual, motor, and motion-related areas. *Journal of Cognitive Neuroscience*, *20*(10), 1799–1814.

Hoffman, P., McClelland, J. L., & Lambon Ralph, M. A. (2018). Concepts, control, and context: A connectionist account of normal and disordered semantic cognition. *Psychological Review*, *125*(3), 293–328.

Huettig, F., & Altmann, G. T. (2007). Visual-shape competition during language-mediated attention is based on lexical input and not modulated by contextual appropriateness. *Visual Cognition*, *15*(8), 985–1018.

Huettig, F., Guerra, E., & Helo, A. (2020). Towards understanding the task dependency of embodied language processing: The influence of colour during language-vision interactions. *Journal of Cognition*, *3*(1), 41, 1–19.

Hutchison, K. A. (2003). Is semantic priming due to association strength or feature overlap? A microanalytic review. *Psychonomic Bulletin & Review*, *10*(4), 785–813.

Jacobs, R. A., & Kruschke, J. K. (2011). Bayesian learning theory applied to human cognition. *Wiley Interdisciplinary Reviews: Cognitive Science*, *2*(1), 8–21.

Joanisse, M. F., & McClelland, J. L. (2015). Connectionist perspectives on language learning, representation and processing. *Wiley Interdisciplinary Reviews: Cognitive Science*, *6*(3), 235–247.

Kan, I. P., Barsalou, L. W., Olseth Solomon, K., Minor, J. K., & Thompson-Schill, S. L. (2003). Role of mental imagery in a property verification task: FMRI evidence for perceptual representations of conceptual knowledge. *Cognitive Neuropsychology*, *20*(3-6), 525–540.

Kellenbach, M. L., Brett, M., & Patterson, K. (2003). Actions speak louder than functions: The importance of manipulability and action in tool representation. *Journal of Cognitive Neuroscience*, *15*(1), 30–46.

Kellenbach, M. L., Wijers, A. A., & Mulder, G. (2000). Visual semantic features are activated during the processing of concrete words: Event-related potential evidence for perceptual semantic priming. *Cognitive Brain Research*, *10*(1-2), 67–75.

Kemmerer, D. (2015). Are the motor features of verb meanings represented in the precentral motor cortices? Yes, but within the context of a flexible, multilevel architecture for conceptual knowledge. *Psychonomic Bulletin & Review*, *22*(4), 1068–1075.

Kemp, C., & Tenenbaum, J. B. (2008). The discovery of structural form. *Proceedings of the National Academy of Sciences*, *105*(31), 10687–10692.

Kiefer, M. (2005). Repetition-priming modulates category-related effects on event-related potentials: Further evidence for multiple cortical semantic systems. *Journal of Cognitive Neuroscience*, *17*(2), 199–211.

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. *Journal of Neuroscience*, *28*(47), 12224–12230.

Kiefer, M., & Pulvermüller, F. (2012). Conceptual representations in mind and brain: Theoretical developments, current evidence and future directions. *Cortex*, *48*(7), 805–825.

Körding, K. P., & Wolpert, D. M. (2006). Bayesian decision theory in sensorimotor control. *Trends in Cognitive Sciences*, *10*(7), 319–326.

Knill, D. C., & Pouget, A. (2004). The Bayesian brain: The role of uncertainty in neural coding and computation. *Trends in Neurosciences*, *27*(12), 712–719.

Kriegeskorte, N., & Kievit, R. A. (2013). Representational geometry: Integrating cognition, computation, and the brain. *Trends in Cognitive Sciences*, *17*(8), 401–412.

Kumar, A. A., Steyvers, M., & Balota, D. A. (2022). A critical review of network-based and distributional approaches to semantic memory structure and processes. *Topics in Cognitive Science*, *14*(1), 54–77.

Kutas, M., & Federmeier, K. D. (2011). Thirty years and counting: Finding meaning in the N400 component of the event-related brain potential (ERP). *Annual Review of Psychology*, *62*, 621–647.

Kutas, M., & Hillyard, S. A. (1980). Reading senseless sentences: Brain potentials reflect semantic incongruity. *Science*, *207*(4427), 203–205.

Lambon Ralph, M. A., Jefferies, E., Patterson, K., & Rogers, T. T. (2017). The neural and computational bases of semantic cognition. *Nature Reviews Neuroscience*, *18*, 42–55.

Lau, E. F., Phillips, C., & Poeppel, D. (2008). A cortical network for semantics: (De) constructing the N400. *Nature Reviews Neuroscience*, *9*(12), 920–933.

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. *Cognitive Science*, *39*(8), 1764–1801.

Löhr, G. (2017). Abstract concepts, compositionality, and the contextualism-invariantism debate. *Philosophical Psychology*, *30*(6), 689–710.

Lucas, M. (2000). Semantic priming without association: A meta-analytic review. *Psychonomic Bulletin & Review*, *7*, 618–630.

Ludlow, P. (2014). *Living words: Meaning underdetermination and the dynamic lexicon*. Oxford, England: Oxford University Press.

Lupyan, G., & Lewis, M. (2019). From words-as-mappings to words-as-cues: The role of language in semantic knowledge. *Language, Cognition and Neuroscience*, *34*(10), 1319–1337.

Machery, E. (2007). Concept empiricism: A methodological critique. *Cognition*, *104*(1), 19–46.

Machery, E. (2015). By default: Concepts are accessed in a context-independent manner. In E. Margolis & S. Laurence (Eds.), *The conceptual mind: New directions in the study of concepts* (pp. 567–588). Cambridge, MA: MIT Press.

Mahon, B. Z., & Caramazza, A. (2008). A critical look at the embodied cognition hypothesis and a new proposal for grounding conceptual content. *Journal of Physiology-Paris*, *102*(1-3), 59–70.

Mamassian, P., Landy, M., & Maloney, L. T. (2002). Bayesian modelling of visual perception. In R. P. N. Rao, B. A. Olshausen, & M. S. Lewicki (Eds.), *Probabilistic models of the brain: Perception and neural function* (pp. 13–36). Cambridge, MA: MIT Press.

Marocco, D., Cangelosi, A., Fischer, K., & Belpaeme, T. (2010). Grounding action words in the sensorimotor interaction with the world: Experiments with a simulated iCub humanoid robot. *Frontiers in Neurorobotics*, *4*, 7.

Martin, A., Haxby, J. V., Lalonde, F. M., Wiggs, C. L., & Ungerleider, L. G. (1995). Discrete cortical regions associated with knowledge of color and knowledge of action. *Science*, *270*(5233), 102–105.

Mazzone, M., & Lalumera, E. (2010). Concepts: Stored or created? *Minds and Machines*, *20*, 47–68.

McClelland, J. L., & Rogers, T. T. (2003). The parallel distributed processing approach to semantic cognition. *Nature Reviews Neuroscience*, *4*(4), 310–322.

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.

McRae, K., De Sa, V. R., & Seidenberg, M. S. (1997). On the nature and scope of featural representations of word meaning. *Journal of Experimental Psychology: General*, *126*(2), 99–130.

Michel, C. (2020). Concept contextualism through the lens of Predictive Processing. *Philosophical Psychology*, *33*(4), 624–647.

Nastase, S. A., Connolly, A. C., Oosterhof, N. N., Halchenko, Y. O., Guntupalli, J. S., Visconti di Oleggio Castello, M., Gors, J., Gobbini, M. I., & Haxby, J. V. (2017). Attention selectively reshapes the geometry of distributed semantic representation. *Cerebral Cortex*, *27*(8), 4277–4291.

Neely, J. H. (1977). Semantic priming and retrieval from lexical memory: Roles of inhibitionless spreading activation and limited-capacity attention. *Journal of Experimental Psychology: General*, *106*(3), 226–254.

Neely, J. H. (1991). Semantic priming effects in visual word recognition: A selective review of current findings and theory. In D. Besner & G. W. Humphreys (Eds.), *Basic processes in reading: Visual word recognition* (pp. 264–336). Hillsdale, NJ: Lawrence Erlbaum Associates.

Negri, G. A., Lunardelli, A., Reverberi, C., Gigli, G. L., & Rumiati, R. I. (2007). Degraded semantic knowledge and accurate object use. *Cortex*, *43*(3), 376–388.

Nelson, D. L., McEvoy, C. L., & Schreiber, T. A. (2004). The University of South Florida free association, rhyme, and word fragment norms. *Behavior Research Methods, Instruments, & Computers*, *36*(3), 402–407.

Nieuwland, M. S., & Van Berkum, J. J. (2006). When peanuts fall in love: N400 evidence for the power of discourse. *Journal of Cognitive Neuroscience*, *18*(7), 1098–1111.

Orbán, G., Fiser, J., Aslin, R. N., & Lengyel, M. (2008). Bayesian learning of visual chunks by human observers. *Proceedings of the National Academy of Sciences*, *105*(7), 2745–2750.

Paczynski, M., & Kuperberg, G. R. (2012). Multiple influences of semantic memory on sentence processing: Distinct effects of semantic relatedness on violations of real-world event/state knowledge and animacy selection restrictions. *Journal of Memory and Language*, *67*(4), 426–448.

Pecher, D., Zanolie, K., & Zeelenberg, R. (2007). Verifying visual properties in sentence verification facilitates picture recognition memory. *Experimental Psychology*, *54*(3), 173–179.

Pecher, D., Zeelenberg, R., & Barsalou, L. W. (2004). Sensorimotor simulations underlie conceptual representations: Modality-specific effects of prior activation. *Psychonomic Bulletin & Review*, *11*(1), 164–167.

Pecher, D., Zeelenberg, R., & Raaijmakers, J. G. (1998). Does pizza prime coin? Perceptual priming in lexical decision and pronunciation. *Journal of Memory and Language*, *38*(4), 401–418.

Plaut, D. C. (2002). Graded modality-specific specialisation in semantics: A computational account of optic aphasia. *Cognitive Neuropsychology*, *19*(7), 603–639.

Popp, M., Trumpp, N. M., Sim, E. J., & Kiefer, M. (2019). Brain activation during conceptual processing of action and sound verbs. *Advances in Cognitive Psychology*, *15*(4), 236–255.

Rabovsky, M., & McRae, K. (2014). Simulating the N400 ERP component as semantic network error: Insights from a feature-based connectionist attractor model of word meaning. *Cognition*, *132*(1), 68–89.

Ralph, M. A. L., Sage, K., Jones, R. W., & Mayberry, E. J. (2010). Coherent concepts are computed in the anterior temporal lobes. *Proceedings of the National Academy of Sciences*, *107*(6), 2717–2722.

Rips, L. J. (1975). Inductive judgments about natural categories. *Journal of Verbal Learning and Verbal Behavior*, *14*(6), 665–681.

Rogers, T. T., & McClelland, J. L. (2004). *Semantic cognition: A parallel distributed processing approach*. Cambridge, MA: MIT Press.

Rommers, J., Meyer, A. S., & Huettig, F. (2015). Verbal and nonverbal predictors of language-mediated anticipatory eye movements. *Attention, Perception, & Psychophysics*, *77*(3), 720–730.

Rommers, J., Meyer, A. S., Praamstra, P., & Huettig, F. (2013). The contents of predictions in sentence comprehension: Activation of the shape of objects before they are referred to. *Neuropsychologia*, *51*(3), 437–447.

Rosch, E. H. (2011). Slow lettuce: Categories, concepts, fuzzy sets, and logical deduction. *Concepts and Fuzzy Logic*, *8*, 89–120.

Sartori, G., & Job, R. (1988). The oyster with four legs: A neuropsychological study on the interaction of visual and semantic information. *Cognitive Neuropsychology*, *5*(1), 105–132.

Schreuder, R., Flores d'Arcais, G. B., & Glazenborg, G. (1984). Effects of perceptual and conceptual similarity in semantic priming. *Psychological Research*, *45*(4), 339–354.

Silveri, M. C., & Gainotti, G. (1988). Interaction between vision and language in category-specific semantic impairment. *Cognitive Neuropsychology*, *5*(6), 677–709.

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.

Steyvers, M., Griffiths, T. L., & Dennis, S. (2006). Probabilistic inference in human semantic memory. *Trends in Cognitive Sciences*, *10*(7), 327–334.

Van Berkum, J. J. (2009). The neuropragmatics of 'simple' utterance comprehension: An ERP review. In U. Sauerand & K. Yatsushiro (Eds.), *Semantics and pragmatics: From experiment to theory* (pp. 276–316). Basingstoke: Palgrave Macmillan.

Van Dantzig, S., Pecher, D., Zeelenberg, R., & Barsalou, L. W. (2008). Perceptual processing affects conceptual processing. *Cognitive Science*, *32*(3), 579–590.

Vermeulen, N., Corneille, O., & Niedenthal, P. M. (2008). Sensory load incurs conceptual processing costs. *Cognition*, *109*(2), 287–294.

Vivas, J., Kogan, B., Yerro, M., Romanelli, S., & Vivas, L. (2021). Describing the structure of concepts through different feature levels. *Journal of Cognitive Psychology*, *33*(1), 49–63.

Warrington, E. K., & McCarthy, R. A. (1987). Categories of knowledge: Further fractionations and an attempted integration. *Brain*, *110*(5), 1273–1296.

Warrington, E. K., & Shallice, T. (1984). Category specific semantic impairments. *Brain*, *107*(3), 829–853.

Werning, M., Unterhuber, M., & Wiedemann, G. (2019). Bayesian Pragmatics provides the best quantitative model of context effects on word meaning in EEG and cloze data. *CogSci 2019*, Montreal, Canada (pp. 3085–3091).

Whitney, P., McKay, T., Kellas, G., & Emerson, W. A. (1985). Semantic activation of noun concepts in context. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *11*(1), 126–135.

Winkielman, P., Coulson, S., & Niedenthal, P. (2018). Dynamic grounding of emotion concepts. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *373*(1752), 20170127.

Winkielman, P., Davis, J. D., & Coulson, S. (2023). Moving thoughts: Emotion concepts from the perspective of context dependent embodied simulation. *Language, Cognition and Neuroscience*, *38*(10), 1531–1553.

Xu, F., & Tenenbaum, J. B. (2007). Word learning as Bayesian inference. *Psychological Review*, *114*(2), 245–272.

Yee, E. (2017). Fluid semantics: Semantic knowledge is experience-based and dynamic. In S. Kotzor & A. Lahiri (Eds.), *The speech processing lexicon* (pp. 236–255). Berlin: De Gruyter Mouton.

Yee, E., & Thompson-Schill, S. L. (2016). Putting concepts into context. *Psychonomic Bulletin & Review*, *23*, 1015–1027.

Yee, E., Ahmed, S. Z., & Thompson-Schill, S. L. (2012). Colorless green ideas (can) prime furiously. *Psychological Science*, *23*(4), 364–369.

Yee, E., Chrysikou, E. G., Hoffman, E., & Thompson-Schill, S. L. (2013). Manual experience shapes object representations. *Psychological Science*, *24*, (6), 909–919.

Zwaan, R. A., Stanfield, R. A., & Yaxley, R. H. (2002). Language comprehenders mentally represent the shapes of objects. *Psychological Science*, *13*(2), 168–171.