Abstract
We continually draw on, and link, conceptual knowledge with perception as we process and interact with our surroundings. This chapter highlights issues at the intersection of perceptual and conceptual processing in human memory. First, it discusses the role of the brain’s perceptual systems and connected regions during conceptual processing. Next, a case study of real-world (or ‘canonical’) size is used to illustrate questions and issues that arise when seeking to understand phenomena that can require information from both perceptual input and semantic memory to be integrated. The influence of conceptual processing on perception is then described, before outlining some additional related factors: conceptual granularity, episodic memory, and individual differences. The chapter concludes by looking to the future of this research area – a field that requires a unique understanding of issues that lie at the heart of perception, memory, and more.
Switching on a lamp, searching for our keys, and interpreting a street sign, all draw on incoming perceptual input and our semantic memory of concepts. In fact, we continually link our conceptual knowledge and perceptions. It is perhaps not a surprise, then, that the field of conceptual processing is broad, with questions that touch upon perception, semantic memory, learning, language, and categorization, among others. This chapter is not intended to be exhaustive, but highlights a set of issues that lie at the intersection of perceptual and conceptual processing in memory. These terms – ‘perceptual’ and ‘conceptual’ – are easily defined at their extremes, but less so at their intersection. Coming back to a previous example, switching on a lamp requires the perceptual (i.e., sensory-grounded) process of detecting its shape – a basic function of the visual system. On the other hand, understanding a lamp’s function, or features that are necessary (connected to outlet, working bulb) versus optional (outside markings), are conceptual processes that are based in semantic memory. This distinction is less clear-cut when input from semantic memory and perception interact. For instance, noticing that a lamp is unusually small or large requires knowing the typical size of lamps, and identifying this lamp’s size from visual size cues (the domain of real-world size is discussed further below).

The link between conceptual and perceptual information in memory is an important topic: the organization of perceptual and conceptual systems is central to understanding relevant neurological deterioration (Fujimori et al., 1997). Patients with semantic dementia frequently experience failures in object memory (Fuld et al., 1990) and associated knowledge (Hodges & Patterson, 2007). In fact, semantic dementia patients can be particularly vulnerable to memory failures for those concepts that rely on combinations of rich perceptual features (Hoffman et al., 2012), reflecting just how intertwined perception and concept knowledge are.

*The role of perceptual systems during conceptual processing*

Perceptual systems are not simply involved in processing visual input, but are also crucial for conceptual processing. Accessing a concept is often accompanied by reactivation of the sensorimotor areas that underlie its features (Lambon Ralph et al., 2017, Martin, 2007; though see Bainbridge et al., 2021 for recent evidence that
reactivation might employ close but non-overlapping voxels). This has been explained through sensorimotor models in which the neural substrates of concepts are distributed across sensorimotor cortical areas (Kiefer & Pulvermüller, 2012; Martin, 2007), and through hub(s)-and-spokes models, in which hubs coordinate with sensorimotor areas to represent concepts (Lambon Ralph et al., 2017). Importantly, areas of the brain that are activated while viewing images of certain domains, such as places (Epstein & Kanwisher, 1998) and objects (Eger et al., 2008; Malach et al., 1995) are activated by both basic features of perceived items and conceptually associated stimuli (Bi, 2020). For instance, the parahippocampal place area responds to viewing certain basic visual properties such as specific shapes (Nasr et al., 2014), but also to landmark-related sounds (Adam & Noppeney, 2010). Similarly, lateral occipitotemporal cortex responds to viewing basic shapes (Chen et al., 2018), as well as to tool-associated sounds (Lewis et al., 2005) and names (Noppeney et al., 2006). Regions that are sensitive to more basic visual features, such as motion, similarly respond to novel objects that have been associated with this feature, even when learned through a verbal descriptor (e.g., "hops"; James & Gauthier, 2003). This may not be universal across all concepts, however. Certain types of concepts do not always show responses to both visual and non-visual stimuli (e.g., animate items; Adam & Noppeney, 2010; Lewis et al., 2005; Noppeney et al., 2006), possibly due to different response mappings across domains (Bi, 2020; Peelen & Downing, 2017). An alternative account for the above findings proposes that these activations reflect response selection processes rather than the representations of concepts per se (Mahon & Caramazza, 2008), though relatively early effects of semantic knowledge (such as the P100; Abdel Rahman & Sommer, 2012) argue against this (see Galetzka, 2017 for a review of this debate).

The sensory regions discussed above are connected to the anterior temporal lobe (ATL), which is thought to act as a ‘graded hub’ that integrates perceptual features into heteromodal concepts (Lambon Ralph et al., 2017), and processes abstract properties (Binney et al., 2016). Studies of the timecourse of ATL activity suggest that an initial rapid pass from perceptual systems to the ATL (Chan et al., 2011) might identify a concept’s superordinate category, which then feeds back to the ‘spokes’ (sensorimotor areas) to help identify concepts at subordinate levels (Lambon Ralph et
The organization of the ATL appears to reflect distinctions between more perceptual, and more conceptual, processing. Across hemispheres, verbal semantic access depends more on left ATL, while image-based tasks depend more on the right (Mion et al., 2010; Ralph et al., 2001). A parallel organization is a dorsal-to-medial and abstract-to-concrete gradient, where concrete features draw on the medial-ventral ATL, and abstract features rely on dorsal-lateral areas (Hoffman et al., 2015; Striem-Amit et al., 2018; Vignali et al., 2020; but see Wang et al., 2010), possibly due to differential connectivity with the ventral stream and language systems, respectively (Lambon Ralph et al., 2017). A key region for integrating perceptual and conceptual processing is the perirhinal cortex (Barense et al., 2012; Clarke & Tyler, 2014; Cowell et al., 2010; Martin et al., 2018). Lying at the intersection of the ventral stream and medial temporal cortex (Miyashita, 2019; Suzuki & Naya, 2014), this area is particularly important for distinguishing within-category exemplars, which is often impaired in patients with semantic dementia, despite their successful recognition of coarser categories (Wright et al., 2015).

Another key region of conceptual processing is the angular gyrus – a higher-order association area that integrates components of semantic concepts (Binder & Desai, 2011; Patterson et al., 2007), and is important in semantic retrieval (Humphreys & Lambon Ralph, 2015). Intriguingly, the angular gyrus responds more strongly when conceptual decisions draw on memory rather than on external stimuli (Murphy et al., 2018). The angular gyrus may control the balance of conceptual and perceptual processing through connectivity with visual cortex (Gonzalez Alam et al., 2019).

When perceptual and conceptual knowledge intersect

In practice, as we process and interact with the world, we frequently draw on both perceptual processing and conceptual knowledge, with each reciprocally feeding into each other. The processing of real-world or 'canonical' size (which occurs frequently during perception; Collegio et al., 2019; Wolfe, 2017) is an excellent test case for

---

1 The terms ‘concrete’ and ‘abstract’ typically describe the nature of the referent (physical versus non-physical, respectively). This contrasts with the ‘perceptual’ and ‘conceptual’ dimensions, which often are used to refer to the source of the information (sensory versus semantic memory, respectively).
understanding this interaction. Real-world size draws on both perceptual and conceptual information, as reflected in some papers framing it as 'perceptual' (Long et al., 2018), while others consider it 'conceptual' (Harel et al., 2014). Unlike purely perceptual features (shape, color, etc.), real-world size is often not apparent from an object’s visual percept. Instead, additional information is often needed, such as explicit knowledge, comparisons to nearby familiar items, spatial context, and other cues.

A debate is ongoing on the relative contributions of perceptual versus conceptual knowledge to how the brain processes real-world size. On the one hand, mid-level perceptual features appear to co-vary with real-world size, suggesting that visual information provides cues to an object’s size (Long et al., 2016). Such mid-level perceptual features can account for a large amount of variance in ventral temporal (VT) activity patterns (Long et al., 2018), and the real-world size of man-made (but not animate) items is reflected in the spatial topography of (univariate) blood-oxygenation level dependent (BOLD) responses in mid-level VT cortex, but not in earlier visual cortex (Konkle & Caramazza, 2013; Konkle & Oliva, 2012).

On the other hand, recent studies have found real-world size information in multi-voxel activity patterns of early visual cortex, even without visual cues, including for concept labels (“camel” versus “mouse”; Borghesani et al., 2016) and geometric shapes that have been associated with different sizes (Gabay et al., 2016). Further, real-world size is detectable in early visual cortex, even when the common correlation between size and taxonomic category is removed (Coutanche & Koch, 2018). In a design that draws on a learning intervention, we have found that declaratively learning a novel item’s real-world size can lead to changes in activity patterns in early visual cortex (Coutanche & Thompson-Schill, 2019). While being scanned with functional magnetic resonance imaging (fMRI), these participants viewed images of previously unfamiliar (and familiar) species, before and after learning the unfamiliar animals’ real-world size through on-screen text. Learning an animal’s real-world size led early visual cortex (though not VT) activity patterns to become more similar between the new species and the size-matched familiar species (unlike in the non-size control conditions). How was this change in early visual cortex activity generated? The right angular gyrus showed the same shift after learning, and showed greater informational connectivity (a measure
of information synchrony; Coutanche & Thompson-Schill, 2013, 2014) with early visual cortex than would be expected by chance. The angular gyrus, which is central to both semantic memory and spatial systems (Seghier, 2013), might provide top-down feedback to these early visual regions.

The role of conceptual knowledge during perceptual processing

Real-world size is not the only domain where conceptual knowledge can influence perceptual systems (Collins & Olson, 2014). Human performance in a simple visual matching task improves once objects are associated with semantic knowledge, even when the semantic features are not necessary for the task (Gauthier et al., 2003). Semantic knowledge, but not verbal labels alone, can also reduce viewpoint-dependency during object recognition (Collins & Curby, 2013). Relatedly, cueing a person with the conceptual identity of an ambiguous image leads them to automatically experience the identity of this concept when they subsequently view the image (Figure 1; Gorlin et al., 2012). This conceptual cueing shifts the neural activity of early visual and lateral occipital cortices to become more similar to the activity patterns evoked by (unambiguous) images of the cued item (Gorlin et al., 2012; Hsieh et al., 2010), reflecting an automatic and rapid influence of conceptual knowledge on the brain's visual systems.

Several mechanisms could underlie the modulation of visual cortex. The rapidity of such changes suggest that neuronal connections are unlikely to be directly modified, but instead that other regions feed back to visual cortex. In some cases, this could reflect direct modulation based on the content of the represented information. For instance, Cohen and Tong (2015) developed a computational model suggesting that feedback from an average object template to V1-like units could improve behavioral detection and discrimination. More broadly, associated conceptual knowledge might help organize perceptual representations in a more efficient or useful manner, leading to reduced demands in subsequent perceptual tasks (Collins & Olson, 2014).
In other cases, modulation of visual cortex might reflect frontoparietal-driven shifts in visual attention (Gilbert & Li, 2013). For instance, the functional connectivity (correlated fluctuations of the BOLD response) of low and high level visual cortex is influenced by top-down attentional demands (Al-Aidroos et al., 2012). The influence of memory on attention is a somewhat understudied area, possibly due to the common but overly simplistic dichotomy of "bottom-up" versus "top-down" attention, which fails to easily include automatic and task-independent memory-driven influences (Chun, 2000; Chun & Jiang, 1999; Zhao et al., 2013). Indeed, memory-guided attention appears to operate independently from top-down functions, with minimal interactions in their behavioral effects (Schwark et al., 2013).

Perceptual and conceptual granularity matters for memory

Memory successes and failures frequently occur at specific semantic levels. For instance, we might misremember the name or breed of a friend’s dog, but are unlikely to misremember Fido as a rabbit, and certainly not as a screwdriver. Similarly, domains of expertise, which reflect an impressive degree of conceptual and perceptual knowledge, typically involve distinguishing within-category items, such as different people, rather than between people and objects (Bruett et al., 2018). The granularity of information in the human ventral stream is reflected in how and where it is represented (Coutanche et al., 2016). For instance, increased (coarse) univariate activity is associated with the presence of an object (Sergent et al., 1992), face (Kanwisher et al., 1997), and scene (Epstein & Kanwisher, 1998). On the other hand, more fine-grained multi-voxel patterns can represent more fine-grained distinctions, such as between different orientations (Kamitani & Tong, 2005), directions of motion (Kamitani & Tong, 2006), colors (Parkes et al., 2009), shapes (Drucker & Aguirre, 2009), and objects (Eger et al., 2008; Haxby et al., 2001). Fine-grained multi-voxel patterns are similarly associated with retrieving different objects from semantic memory (Coutanche & Thompson-Schill, 2015), recognizing different familiar items (LaRocque et al., 2013), and other memory processes (Rissman & Wagner, 2012).

We have previously demonstrated the ability and value in analyzing concepts at different granularities in a test of whether retrieving a known object draws on an
integration site, or “convergence zone” (Damasio, 1989; Meyer & Damasio, 2009). While fMRI participants were cued to look for a known object from within (pure) visual noise, we have been able to decode the shape (spherical or elongated) of the retrieved object in LOC (a key shape region), and its color (orange or green) from right V4 (a known color-processing center; Coutanche & Thompson-Schill, 2015). A “generalization test”, where a classifier is trained and tested on different items (such as training on tangerine versus lime, but testing on carrots versus celery), gives confidence that this reflects the hypothesized color distinction: tangerines and limes differ in many ways, but only their color difference should generalize to distinguishing carrots from celery. An exploratory “searchlight” found that only the left ATL encoded the retrieved object’s identity. Consistent with accounts of convergence in the ATL, the presence of object identity in ATL activity patterns was predicted by the joint presence of shape decoding in LOC and color decoding in V4 (Coutanche & Thompson-Schill, 2015).

Recent studies of memory reactivation have also examined concept memories across granularities, including at the level of item and semantic category (Kuhl & Chun, 2014; Lee et al., 2019; Mack & Preston, 2016), finding that each has a corresponding consequence for future memory success (Lee et al., 2019).

Episodic memories of percepts and concepts

Neural patterns collected during encoding suggest that visual and semantic representations each contribute to memory retrieval during perceptual and conceptual tests through a complex interaction of representation, test-type, and brain region (Davis et al., 2021). For instance, visual representations in visual cortex are predictive of perceptual memory performance, but then predict conceptual memory performance when in more anterior regions. Key memory regions, however, such as the hippocampus, respond across retrieval tests. When conceptual relationships are explicitly task-relevant, their complex organization (such as nested hierarchies of learned words) is found within hippocampal activity (Viganò & Piazza, 2021).

When we encode and retrieve an episodic association, such as between an image and word, the robustness (i.e., reliability) of the activity patterns at encoding, and the reactivation of patterns between encoding and retrieval, predict subsequent memory
performance (LaRocque et al., 2013; Ritchey et al., 2013). Nonetheless, concepts are more complex and variable than word–image associations. For instance, the concept of "dog" must incorporate visual depictions that vary based on differences across exemplars (Fido, Scooby-Doo) and viewpoints (facing forward, side-ways, occluded, etc.) We recently asked how activity pattern robustness and reactivation predict a person’s memory for associations between novel words and visually varying concepts (here, unfamiliar animals) approximately one month after initial encoding (Bruett et al., 2020). Pattern robustness during encoding significantly predicted subsequent memory in a variety of regions, including left ventrolateral prefrontal cortex, left medial parietal cortex, right perirhinal cortex, and early visual cortex. In contrast, pattern reactivation between encoding and retrieval was not predictive of subsequent memory, suggesting that in a more concept-motivated design that introduces visual variation and long encoding/retrieval delay, the similarity between encoding and retrieval activity patterns might not have the same relationship with memory performance as found for simple word–image pairs. The question of how encoding and retrieval patterns differ for concepts, and the consequences for subsequent memory, is a fascinating question for future research.

An object’s properties play a strong role in determining whether we remember encountering an item or not (Bainbridge, 2019). The features that influence memorability are consistent across individuals, with high consistency in which objects will, or will not, be remembered (Isola et al., 2011), though there is uncertainty about what these factors are. Perceiving images that are memorable (compared to those that are not) leads to greater neural activity across areas of the later ventral stream and in memory networks, including perirhinal cortex, parahippocampal cortex, and the medial temporal lobe (MTL; Bainbridge et al., 2017). Notably, such activation is observed even if these memorable images are not actually remembered, so relates to the stimuli rather than memory outcomes (Bainbridge et al., 2017).

We have recently shown that the memorability of images depends on levels of perceptual processing in different ways (Koch et al., 2020). By using a convolutional neural network (CNN) trained for object recognition to select images that participants view, we determined that high discriminability in low-level visual features, but greater
similarity at higher (category) levels, predict greater memorability. The organization of a CNN, in which internal "hidden" layers of nodes feed information forward to a final output layer, has parallels with the organization of the human visual system (Yamins & DiCarlo, 2016), in which representations shift from more basic visual features to higher-level object categorization as one moves from posterior to anterior regions (Coutanche et al., 2016). Recently, studies of memorability have been extended to semantic features, finding that the interrelatedness of item features positively predicts hit rates on visual and lexical memory tests (Hovhannisyan et al., 2021). Interestingly, hit rates are correlated across tests, suggesting some similarity in the basis for perceptual and conceptual memory traces. It has yet to be determined if the roles of semantic features also differ based on their level in a hierarchy.

**Individual differences**

A common practice in human neuroscience research is to collapse across subjects, or to transform analyzed brain data into a “standardized” form. This approach helps detect shared neural characteristics across a group by eliminating irrelevant variance. Although this continues to contribute important findings to the memory literature, by design it removes (potentially informative) between-subject neural differences (Dubois & Adolphs, 2016). Yet, these differences can represent a substantial portion of the data. For example, 54% of between-subject variance in subjects’ functional brain activity, over and above structural variance, remains unexplained (Miller et al., 2012).

Recent studies suggest that individuals reliably differ in the extent that they draw on different memory systems (Palombo et al., 2013). Some individuals are more likely to encode and later retrieve information that is vivid and context-bound (“episodic”) versus abstracted and removed from specific events (“semantic”; Palombo et al., 2013). Intrinsic (at-rest) brain networks of individuals differ based on whether they report a tendency to recall episodes very vividly versus abstractly (Sheldon et al., 2016). These networks show strong connectivity between memory structures (MTL) and perceptual regions in individuals who frequently retrieve vivid episodes, consistent with a system that replays events with vivid sensory detail. In contrast, people who are more likely to
retrieve information in an abstracted “semantic” manner have stronger connectivity between prefrontal, and MTL and temporal cortices – important for retrieving factual knowledge (Binder & Desai, 2011). Behaviorally, such individual differences appear to interact with the type of encoding and retrieval test in complex ways. For instance, individuals with high semantic trait scores show a greater tendency to lexically integrate new words through fast mapping (Coutanche & Koch, 2017), while those higher in the more visually vivid episodic and spatial dimensions show superior performance in free (but not cued) recall of naturalistic visual episodes (Coutanche, Koch, et al., 2020). This direction of research is still young, with enormous opportunities to account for currently unexplained between-participant variance within behavioral and neuroimaging data.

Looking forward

The question of how percepts and concepts interact is ripe for significant advances that will deepen our knowledge of both facets, and their interaction. To give one example, memorability is often framed as the likelihood that an image will be remembered, but this memory for whether an object was encountered is a small part of our memory for objects. The knowledge that we acquire about objects can include their typical motion (Schlack & Albright, 2007), size (Coutanche & Thompson-Schill, 2019), category (van der Linden et al., 2008), value (Murray & Richmond, 2001) and more. A full understanding of memorability requires greater understanding of an item's associated features and conceptual knowledge.

Another direction of future research is understanding how top-down signals that modulate visual cortices (discussed above) interact with the concurrent signals generated by visual input. Specifically, how is existing knowledge reflected in visual cortex while current external sensory input is being processed? This has been noted in some studies of visual cortex. For example, using an attentional task with two stimuli simultaneously presented (one attended; one unattended), Jehee and colleagues observed that “task-related activity was found even when the attended stimulus differed in orientation from the unattended stimulus, indicating that this top-down orientation-selective signal can operate independently of the bottom-up input” (Jehee et al., 2011). One possible answer is that neural activity patterns generated by top-down signals are
intrinsically different, allowing them to co-exist with sensory signals without negative consequences for either. Alternatively, a third neural system, spatially separated from perceptual and memory-related structures, may provide modulation that prevents mutual interference. Recent methods that combine multi-voxel pattern information with approaches to connectivity (Anzellotti & Coutanche, 2018) have potential to shed light on such questions.

Finally, a burgeoning subfield of ‘conceptual combination’ asks how we can flexibly combine concepts in new ways, allowing us to understand 'gingerbread house', 'picture book', and others, when we encounter them (Coutanche, Solomon, et al., 2020). The question of how we select and extract the key perceptual features of one concept, and integrate them into another, has been the focus of some fascinating recent research (e.g., Price et al., 2015) that in turn illuminates the workings of the semantic memory system.

**Summary**

In this chapter, I have sought to describe ways that perceptual and conceptual processing interact in memory. This topic is broad, necessitating that some issues and debates have been left out. Nonetheless, the issues discussed above alone reflect decades of fascinating and rigorous research that has brought us to this point, as well as the enormous future potential to deepen our understanding.


Bibliography


Sergent, J., Ohta, S., & MacDonald, B. (1992). Functional neuroanatomy of face and
object processing. A positron emission tomography study. *Brain: A Journal of

lobe connectivity relates to individual differences in episodic autobiographical

visual concepts in people born blind. *Nature Communications, 9*(1), 5250.
https://doi.org/10.1038/s41467-018-07574-3


together: Experience-driven formation of visual object categories in human ventral
temporal cortex. *PloS One, 3*(12), e3995. https://doi.org/10.1371/journal.pone.0003995

hierarchical relations between words during concept learning. *Hippocampus.*
https://doi.org/10.1002/hipo.23320

Timecourse and convergence of abstract and concrete knowledge in the anterior
https://doi.org/10.1101/2020.06.04.134163

representation of abstract and concrete concepts: A meta-analysis of
https://doi.org/10.1002/hbm.20950


Wright, P., Randall, B., Clarke, A., & Tyler, L. K. (2015). The perirhinal cortex and
conceptual processing: Effects of feature-based statistics following damage to the
https://doi.org/10.1016/j.neuropsychologia.2015.01.041

https://doi.org/10.1038/nn.4244

https://doi.org/10.1177/0956797612460407