Abstract Concepts: Sensory-motor Grounding, Metaphors, and Beyond

Diane Pecher, Inge Boot, & Saskia Van Dantzig

Author Note
Diane Pecher, Erasmus University Rotterdam, The Netherlands; Inge Boot, Erasmus University Rotterdam, The Netherlands; Saskia Van Dantzig, Leiden University, The Netherlands. Starting in September 2010 Inge Boot will be at the University of Amsterdam, The Netherlands.

This research was supported by grants from the Netherlands Organization for Scientific Research (NWO) to Diane Pecher. We are grateful to Larry Barsalou and Gerard Steen for stimulating discussions and Seana Coulson for her very helpful comments on an earlier version of this manuscript.

Correspondence concerning this article should be addressed to Diane Pecher, Psychology Department, T12-33, Erasmus University Rotterdam, Postbus 1738, 3000 DR Rotterdam, The Netherlands. E-mail: pecher@fsw.eur.nl
Abstract

In the last decade many researchers have obtained evidence for the idea that cognition shares processing mechanisms with perception and action. Most of the evidence supporting the grounded cognition framework focused on representations of concrete concepts, which leaves open the question how abstract concepts are grounded in sensory-motor processing. One promising idea is that people simulate concrete situations and introspective experiences to represent abstract concepts (Barsalou & Wiemer-Hastings, 2005), although this has not yet been investigated a lot. A second idea, which more researchers have investigated, is that people use metaphorical mappings from concrete to abstract concepts (Lakoff & Johnson, 1980). According to this conceptual metaphor theory, image schemas structure and provide sensory-motor grounding for abstract concepts. Although there is evidence that people automatically activate image schemas when they process abstract concepts, we argue that situations are also needed to fully represent meaning.
1. Grounded cognition

1.1. Introduction

How do people think? People may describe their thoughts as mental images, imagined movements through space, simulated sequences of actions, and so on. The idea that the elements of thought are not words or symbols, but visual and motor images is at the core of recent theories of cognition (e.g., Barsalou, 1999b; Glenberg, 1997; Grush, 2004; Zwaan, 1999), that are commonly referred to by the label grounded cognition. According to the grounded cognition view, thinking relies on the activation of the sensory-motor system. When a person thinks about an object such as a banana, the neural patterns in sensory-motor brain areas that have been formed during earlier experiences with bananas are reactivated. This reinstatement of neural activation results in a sensory-motor simulation. The visual system simulates seeing a banana, the motor system simulates the acts of grasping, peeling, and eating the banana, and the olfactory and gustatory system simulate the smell and the taste of the banana. In other words, when thinking about a banana, the brain acts partially as if the person is actually perceiving and interacting with a banana.

The grounded cognition approach solves two issues that have not been addressed very thoroughly by previous theories. First, most theories are quiet on the issue of how mental symbols get their meaning, other than by links to other symbols. In order for mental symbols to be meaningful, they need to be grounded in experience. Sensory-motor simulations might provide such grounding. Second, the grounded cognition approach provides an explanation for the flexibility of concepts (i.e., mental representations). Concepts are dynamic constructions that are highly variable, dependent on context and recent experience (Anderson et al., 1976; Barsalou, 1982, 1993; Pecher & Raaijmakers, 2004; Zeelenberg, Pecher, Shiffrin, & Raaijmakers, 2003). Simulations can involve different sensory modalities to different extents (Pecher, Zeelenberg, & Barsalou, 2004). For example, a concert pianist may form different simulations of piano in various contexts. When thinking about an upcoming performance, the simulation might include the piano’s sound, along with the fine hand movements involved in playing the
instrument. When planning to move the piano into a new apartment, however, the simulation might include the shape, size, and weight of the piano, along with the gross movements necessary for lifting and moving it. Furthermore, different instances of a category can be instantiated from one simulation to another. Thus, the simulation mechanism provides a logical explanation of concept variability. These and other reasons have led many researchers to put the grounded cognition framework to the test.

Most of the evidence supporting the grounded cognition framework focused on representations of concrete concepts (e.g., objects; Pecher, Zeelenberg, & Barsalou, 2003) or actions (Glenberg & Kaschak, 2002). A frequent criticism of these studies is that they fail to explain how more abstract concepts such as power or democracy are represented. Because people do not have direct sensory-motor experiences with abstract concepts, it is not immediately obvious how they can represent such concepts by sensory-motor simulations. In this chapter we will discuss various views on how abstract concepts might be represented through sensory-motor simulations. In particular, we will investigate the idea that metaphorically used image schemas may provide a mechanism to ground abstract concepts in embodied experience.

1.2. The grounded cognition framework

At first glance it may seem more logical to adopt a symbolic approach to explain abstract concepts. After all, democracy does not have a particular color, shape, smell, sound, or weight. People cannot kick, bite, or squeeze it. It is associated with other abstract concepts such as freedom, majority, or republic. Thus, one may wonder how the sensory-motor systems in the brain can ever represent abstract concepts. On the other hand, the cumulative evidence for the idea that cognition is grounded in sensory-motor systems is too strong to ignore. This theory solves the grounding problem (Harnad, 1990) for concrete concepts, involving objects and actions. If we now require symbolic representations to explain abstract concepts, the grounding problem returns. Therefore, it makes sense to investigate whether the same mechanisms that have proven so successful for concrete concepts can also be used for abstract concepts.
Probably the most influential view of grounded cognition is given by Barsalou’s Perceptual Symbol Theory (Barsalou, 1999b). Barsalou proposes that mental representations used in cognitive tasks are grounded in the sensory-motor system. The building blocks of mental representation are *perceptual symbols*, partial reinstatements of the neural patterns that were stored in perceptual and motor brain areas during actual experience and interaction with the environment. These perceptual symbols are organized into simulators. A simulator is a set of related perceptual symbols, representing a particular concept. A simulator should not be viewed as a re-enactment of an entire previous experience. Rather, the simulator uses a subset of components that were captured during different experiences. Because perceptual symbols can be combined dynamically to produce a simulation, concepts are specific instances of a category. For example, a person may represent a *banana* by simulating eating a ripe banana or by simulating seeing a green banana hanging on a tree. This ability to combine components of experience dynamically allows the system to form new concepts. Since simulations are composed of perceptual symbols, they are analogous to actual perception and action. According to Barsalou, a perceptual symbol system is a fully functional conceptual system, which “represents both types and tokens, produces categorical inferences, combines symbols productively to produce limitless conceptual structures, produces propositions by binding types to tokens, and represents abstract concepts.” (Barsalou, 1999b, p. 581).

1.3. Evidence for the grounded cognition view: Concrete concepts

1.3.1. Perceptual simulations

On this account, even when people are not actually perceiving or interacting with objects in the world, the sensory-motor systems are still actively involved in cognitive processes such as language comprehension, categorization, and memory retrieval. Because representations are formed in sensory-motor brain areas, they should retain perceptual qualities. This hypothesis has been confirmed by the results of several studies. For example, Solomon and Barsalou (2001) asked participants to verify sentences such as
a bee has wings. Earlier in the experiment the same property (e.g., wings) had been presented with a different entity. During this earlier presentation, the property either had a similar perceptual form (e.g., wasp-wings) or a different perceptual form (e.g., butterfly-wings). Performance on the property-verification task was better if the property had a similar form as in the earlier presentation than if it had a different form. Priming for shape similarity has also been observed in tasks that require less semantic processing, such as word naming and lexical decision (Pecher, Zeelenberg, & Raaijmakers, 1998), although the effect tends to be quite fragile in such tasks. Thus, overlap in perceptual features facilitates processing. This effect of shape similarity indicates that visual properties are part of the conceptual structure. This is corroborated by studies showing effects of visual perspective (Borghi, Glenberg, & Kaschak, 2004; Solomon & Barsalou, 2004; Wu and Barsalou, 2009). For example, Borghi et al. presented sentences that evoked an inside perspective (e.g., You are driving a car) or an outside perspective (e.g., You are washing a car). Following the sentence they presented the name of an object part (e.g., steering wheel, antenna). Borghi et al. found that participants were faster to verify properties that were congruent than incongruent with the evoked perspective. This result indicates that participants simulated perceiving the object (e.g., car) from the evoked perspective, and used this simulation to perform the subsequent property verification. Properties that were more likely to be included in the simulation were verified faster than those that were less likely to be included in the simulation. In addition, information from different modalities can be more or less relevant for a task, which affects the composition of a particular representation. A couple of studies showed that there is a processing cost associated with switching the relevance from one sensory modality to another (Marques, 2006; Pecher et al., 2003; Vermeulen, Niedenthal, & Luminet, 2007). For example, participants were faster to verify sentences such as a banana is yellow (visual property) if on the previous trial they had verified that a gemstone is glittering (visual property) than if they had verified that leaves are rustling (auditory property). Studies such as these provide evidence that perceptual features are retained in concepts and strongly suggest that cognition uses the same systems as perception and action.

In support of the behavioral findings described above, neuroscientific studies showed that modality-specific sensory brain areas are active during conceptual
processing (for a review, see Martin, 2007). Verifying properties from different sensory modalities is associated with the activation of specific sensory brain areas. For example, verifying that a banana is yellow is associated with activation of the visual area, whereas verifying that a banana is sweet corresponds with activity in the gustatory area (e.g., Goldberg, Perfetti, & Schneider, 2006; Kan, Barsalou, Solomon, Minor, & Thompson-Schill, 2003). The retrieval of perceptual knowledge thus appears to rely on the activation of modality-specific brain regions that are associated with experiencing instances of the represented concepts.

Because simulations involve processing in sensory-motor systems, the account also predicts that there should be direct interactions between representational processes and perceptual processes. An important line of evidence comes from studies that show effects of mental representation on processing of visual stimuli. In such studies, visual stimuli are typically better processed if their visual features overlap or match with those of the mental representation, For example, Stanfield and Zwaan (2001) presented sentences in which the horizontal or vertical orientation of an object was implied. For instance, the sentence *John put the pencil in the cup* implies a vertically oriented pencil, whereas the sentence *John put the pencil in the drawer* implies a horizontally oriented pencil. Immediately following the sentence a picture was presented, and the participant decided whether the object in the picture was mentioned in the preceding sentence. On the trials of interest, the depicted object was presented either in the implied orientation or in the perpendicular orientation. Participants were faster and more accurate to recognize the object when the orientation of the picture matched the orientation implied by the sentence than when it did not match. Similar results were obtained for overlap in shape (e.g., sliced vs. whole apple, Zwaan, Stanfield, & Yaxley, 2002) and color (brown vs. red steak, Connell, 2007). These compatibility effects even arise when representation and perception are separated by a delay (Pecher, van Dantzig, Zwaan & Zeelenberg, 2009) or when representation and perception only overlap in sensory modality and not in the specific concept that is simulated (Van Dantzig, Pecher, Zeelenberg, & Barsalou, 2008), which indicates that the effects are not due to strategic imagery.

1.3.2. The role of action
In addition to perception, action plays an important role in representing concepts (Borghi, 2005). Since the main function of cognition is to support our interactions with the environment (Glenberg, 1997) action should be central to representations. This is supported by multiple demonstrations of interaction between representation and action. For example, in various language comprehension studies, participants are instructed to perform a particular motor action while reading words or sentences. The typical outcome of these studies is that participants read and verify a sentence faster when the performed motor action corresponds to the action described by the sentence. This effect is referred to as the Action Compatibility Effect (ACE). For instance, participants are faster to verify the sentence *You handed Jane the ball* by making a hand movement away from themselves than by making a hand movement toward themselves. Oppositely, they are faster to verify the sentence *Jane handed you the ball* by moving the hand toward than away from themselves (Glenberg & Kaschak, 2002; see also Bub & Masson, 2010; Girardi, Lindemann, & Bekkering, 2010; Glenberg et al., 2008; Klatzky, Pellegrino, McCloskey, Doherty, 1989; Masson, Bub, & Warren, 2008; Pezzulo, Barca, Bocconi, & Borghi, 2010; Van Elk, Van Schie, & Bekkering, 2009; Zwaan & Taylor, 2006).

Furthermore, neuroscientific studies showed that motor areas are active while reading action words. Interestingly, a somatotopic organization is found, such that words referring to actions performed by different parts of the body (e.g., hand actions such as ‘grasp’, foot actions such as ‘kick’, mouth actions such as ‘kiss’) are associated with activations of different somatotopic areas of the motor cortex (Hauk, Johnsrude, & Pulvermüller, 2004).

To sum up, theories of grounded cognition propose that mental representations are formed by sensory-motor systems. On this account, concepts are simulations of interactive experiences in the real world. This view is supported by empirical studies that show that representations have similar features as perception and action. In addition, many studies show interactions between representation and sensory-motor tasks. Finally, neuroscientific studies indicate that cognition shares neural structures with perception and action.
1.4. Are all concepts grounded in sensory-motor processing? The scope problem

The studies presented in the previous sections demonstrate that the sensory-motor system is involved in many cognitive processes, such as conceptual processing and language comprehension. Based on this and related empirical evidence, we can conclude that cognition at least partially involves the systems of perception and action. Clearly, cognition is not completely amodal and symbolic. However, the question remains whether cognition can be completely grounded in perception and action. In other words, can all cognitive tasks be performed with analogue, sensory-motor representations, or do some cognitive tasks require more symbolic, abstract representations? This issue is called the scope problem by Machery (2007). He argues that findings of embodiment in a particular cognitive task cannot automatically be generalized to other tasks. According to Machery, “the interesting question is not, ‘Are concepts amodal or perceptual?’, but rather ‘To what extent do we use reenacted perceptual representations in cognition and to what extent do we use amodal representations?’”. (Machery, 2007, p. 42)

Some researchers take a wide-scope embodied standpoint on this issue, claiming that cognition does not need any amodal representations. For example, Barsalou used the following argument: “Perceptual symbol systems can implement all of the critical cognitive functions that amodal symbol systems have implemented traditionally. If so, then why did evolution add a redundant layer of amodal symbols?” (Barsalou, 1999a, pp. 75-76). However, others argue that amodal representations may be required for tasks that involve more abstract reasoning. For example, consider thinking about a rather abstract domain such as law or politics. According to Clark,

[I]t is not clear how rich sensorimotor simulation could possibly account for all the kinds of moral and abstract reasoning required – reasoning about rights, implications, responsibilities, economics, and so on. (…) [I]t is hard to see how sensorimotor simulation could in principle account for all the kinds of thought and reasoning that the problem demands. (…). It does seem that the more decoupled and abstract the target contents become, either the less applicable the sensorimotor simulation strategy
is, or the less clearly it can then be differentiated from the more traditional approaches it seeks to displace. (Clark, 1999, p. 348)

M. Wilson (2002) has argued that cognitive theories can be placed on a continuum, ranging from a purely embodied account of cognition to a purely disembodied account. On one end of the continuum, pure embodied accounts claim that cognition is completely grounded in the sensory-motor system. Oppositely, pure disembodied accounts claim that cognition is completely symbolic and amodal. These opposite accounts define the ends of the continuum, leaving space for other theories to occupy the middle ground between the two extremes. Several of such intermediate theories have been proposed, each suggesting that cognition is partially based on sensory-motor processing and partially based on amodal, symbolic processing. For example, Mahon and Caramazza (2008) proposed the grounding by interaction framework. They hypothesized that concepts partially consist of amodal symbols, and partially of sensory-motor information. While the core of a concept is formed by amodal, symbolic information, sensory-motor information “colors conceptual processing, enriches it, and provides it with a relational context” (Mahon & Caramazza, 2008, p. 10). A related idea was proposed by Barsalou, Santos, Simmons and Wilson (2008; see also Simmons, Hamann, Harenksi, Hu, & Barsalou, 2008), in their Language and Situated Simulation (LASS) theory. This theory assumes that concepts are represented in two distinct ways; by means of linguistic representations and by means of situated, sensory-motor simulations. During conceptual processing, both the linguistic system and the simulation system become active. The systems interactively contribute to the representation of concepts. The linguistic system is thought to underlie relatively superficial processing, whereas deeper conceptual processing requires engagement of the simulation system (see also Solomon & Barsalou, 2004).

The scope problem is related to two questions; the necessity question and the sufficiency question (Fischer & Zwaan, 2008). The necessity question asks whether sensory-motor representations are necessary for cognitive processing. The sufficiency question asks whether sensory-motor representations are sufficient for cognitive processing. Whereas the pure embodied view claims that sensory-motor representations
are necessary and sufficient for cognition, the disembodied view claims that sensory-motor representations are neither necessary nor sufficient. The intermediate theories suggest that sensory-motor representations are necessary for deep conceptual processing but not for shallow processing. Additionally, they argue that sensory-motor simulation may not be sufficient for cognition, but that it plays an important role by enriching mental representations and grounding them in experience.

2. Representing abstract concepts: Some evidence for grounding

Are abstract concepts grounded in sensory-motor processes? Answering this question would shed some light on the scope problem. The view that sensory-motor grounding is necessary for cognitive processing predicts that such grounding should be consistently found for abstract concepts. Moreover, such a view predicts that in the absence of grounding full understanding of abstract concepts is impossible. Although the latter claim is very hard to test (it might entail removal of all sensory-motor systems), we argue that consistent findings of sensory-motor effects for abstract concepts provides good indications that grounding is necessary. Because abstract entities by definition have no perceptual or motoric details, we should not expect sensory-motor effects to arise as a mere by-product of processing. Rather, such findings would indicate that sensory-motor processes are fundamental to the representation of abstract (and concrete) concepts.

2.1. Emotional valence

An important line of research that suggests that sensory-motor grounding may extend beyond representations of concrete objects and actions is the one investigating interactions between body and emotional valence. Unlike shape or color, the emotional valence of entities may not be perceived directly. Moreover, interaction effects are also obtained for words that refer to abstract concepts such as *peace* or *hostility*. Many studies demonstrated that positive and negative words automatically trigger approach or avoidance actions (e.g., Chen & Bargh, 1999; Lavender & Hommel, 2007; Rotteveel & Phaf, 2004; Solarz, 1960; Wentura, Rothermund, & Bak, 2000). This phenomenon, called
the approach/avoidance effect, has been brought forward as an example of grounded cognition (e.g., Niedenthal, Barsalou, Winkielman, Krauth-Gruber, & Ric, 2005). In studies of this phenomenon, participants typically respond to stimuli with positive and negative valence by making an arm movement toward or away from the stimulus. Approach can be defined as the tendency to reduce the distance between a stimulus and the self, whereas avoidance can be defined as the tendency to increase the distance between a stimulus and the self. Pleasant stimuli automatically elicit approach tendencies, while unpleasant stimuli elicit avoidance tendencies. Accordingly, people are faster to respond to positive stimuli by making an approach movement (resulting in reducing the distance toward that stimulus) than by making an avoidance movement (resulting in increasing the distance). Conversely, they are faster to respond to negative stimuli by making an avoidance movement than by making an approach movement. For example, in the study by Chen and Bargh (1999), participants categorized positive and negative words by pulling a joystick toward themselves or pushing it away. In response to positive words, participants were faster to pull the joystick than to push it away. In response to negative words, in contrast, they were faster to push the joystick away than to pull it toward themselves. The effect is not located in the specific arm movement that has to be made, but rather seems to depend on the expected effect of an action. For example, approach/avoidance effects have even been found when participants do not move their arms with respect to the stimulus, but instead make button presses that result in an apparent movement of the stimulus toward or away from the participant (Van Dantzig, Pecher, & Zwaan, 2008).

In order to investigate whether the approach/avoidance effect truly reflects embodied action, Markman and Brendl (2005) introduced a method in which the participant’s “self” was disembodied. They presented the participant’s name on the computer screen, and argued that this would lead to an abstracted representation of the “self” that was spatially dissociated from the participant’s real body. Markman and Brendl instructed participants to make valence judgments to words by moving a joystick in one of two directions, such that the word would move towards or away from the name. They obtained the usual congruency effect, but with respect to the participant’s name on the computer screen instead of their real body. According to Markman and Brendl, their
study demonstrates that approach/avoidance actions are not executed with respect to the body, but with respect to a symbolic, disembodied representation of the self. They conclude that their results constrain grounded theories of cognition, because they show that some cognitive tasks require higher order symbolic representations.

However, the results of a study by Van Dantzig, Zeelenberg, and Pecher (2009; see also Proctor & Zhang, 2010) suggest that the findings of Markman and Brendl are the result of an artifact of their task, rather than a reflection of disembodied, symbolic processing. Van Dantzig et al. showed that similar approach/avoidance effects were obtained with respect to a meaningless rectangle on the computer screen, indicating that the effect was not due to an abstract representation of the self. Thus, the study of Markman and Brendl does not necessarily undermine grounded theories of cognition.

These studies show that valence influences action. Effects in the opposite direction -motor actions influencing evaluations- have also been observed. Such effects are found even when participants are unaware of the relation between their action and the evaluated stimulus. In several studies participants received a misleading cover story instructing them to perform certain motor actions, which affected evaluations of unrelated stimuli. For example, valence judgments are more positive when participants are instructed to hold a pen in their mouth using only their teeth, such that their facial expression resembles a smile. Judgments are more negative when participants hold a pen only using their lips, such that their facial expression resembles a frown (Strack, Martin, & Stepper, 1988). Recognition of positive stimuli was better when participants were nodding their head, and better for negative stimuli when they were shaking their head (Forster & Strack, 1996). Extending the middle finger leads to higher hostility ratings, whereas pointing the thumb upwards leads to higher liking ratings (Chandler & Schwarz, 2009). Conversely, when participants are prevented from making spontaneous facial expressions, they are worse at recognizing emotion (Oberman, Winkielman, & Ramachandran, 2007; Havas, Glenberg, Gutowski, Lucarelli, & Davidson, in press). These results provide evidence that representations of emotional valence are grounded in embodied interactions.

2.2. Abstract transfer
Evidence for sensory-motor grounding has also been observed for non-emotional concepts. There appears to be an interesting parallel between transfer of objects and transfer of information. The action compatibility effect (ACE) described earlier was also found for sentences describing abstract transfer events (Glenberg & Kaschak, 2002; Glenberg et al., 2008). Similar to responses to sentences such as *Jane handed you the ball* in which the described event involves actual hand movements, sentences such as *Liz told you the story* also facilitated responses toward oneself, even though the described situation does not involve any specific hand movements. These results indicate that both concrete and abstract transfer events are represented by simulating concrete transfer, which provides sensory-motor grounding for abstract transfer.

2.3. Force dynamics

Another interesting parallel between the concrete and abstract domain is found in the representation of forces (Talmy, 1988). Talmy's theory of force dynamics claims that many events, as well as linguistic descriptions of events, can be understood in terms of patterns of forces, both physical and psychosocial. According to this theory, event descriptions are understood in terms of naïve physics and a perceived tendency of described entities to act or rest. In its simplest form, the theory supposes that each event consists of two force-exerting entities. The agonist is a focal entity with a tendency to act or to rest, while the antagonist is a secondary entity that exerts a force to oppose the agonist. The forces exerted by the agonist and antagonist arise through a drive (e.g., to eat, sleep, leave, say something) or a physical circumstance (gravitational force to fall or to stay still). Depending on the relative strengths of these two force-exerting entities, the focused agonist is either able to manifest its force tendency or is overcome by the antagonist. For example, the sentence *the gate prevented the boy from leaving the yard* can be understood by perceiving the boy as the focused entity (agonist) that has a tendency to leave the yard, and the gate as a stronger entity (antagonist) that instead forces the boy to stay (or rest, in Talmy’s terminology). Talmy shows that this basic idea of opposing agonists and antagonists yields various complex patterns that give rise to
such semantic constructs as causation, allowing, and helping (see Wolff, 2007 for a model of causality that relies on force dynamics primitives). According to Talmyn’s theory, concrete events (a player kicking the ball) are represented in the same manner as abstract events (a lawyer persuading the jury to convict the defendant). In the case of abstract sentences, the agonist may tend toward rest or action in less physical ways (jury tends towards rest in not convicting the defendant) while the antagonist (the lawyer) opposes this force. Again, whether or not the agonist results in carrying out its own force tendencies or those of the antagonist is determined by the relative strength of the agonist and antagonist. If the lawyer is strong enough, he will force the jury to act and convict the defendant.

To test Talmyn’s theory, Madden and Pecher (2010) presented sentences to participants for sensibility ratings. The sentences described one of four force patterns (force action, force rest, allow action, allow rest) in either a concrete situation (e.g., The bulldozer pushed the pile of dirt across the lot) or an abstract situation (e.g., Her friends persuaded the girl to come to the party). Before each sentence, an animation was shown in which two geometrical shapes interacted in a way that matched or mismatched the force pattern of the following sentence (e.g., a circle moving towards a rectangle, forcing it to topple). The results showed that sensibility ratings were faster and more accurate for sentences that matched with the animation than for sentences that mismatched with the animation, suggesting that prior activation of the correct force pattern made it easier to represent the meaning of the sentences. The advantage of match over mismatch was found for both abstract and concrete situations, indicating that representations of concrete and abstract forces are similar.

2.4. Summary

The three cases described above, regarding the representation of emotional valence, abstract transfer events, and force dynamics, show that sensory-motor processing plays a role for some abstract concepts. In these cases, the sensory-motor effects are similar for abstract concepts and for their more concrete counterparts. For example, the approach/avoidance effect extends from concrete things that people want to
physically avoid (e.g., a dangerous animal) to more abstract things that are physically harmless (e.g., the word *hostility*). Thus, these studies suggest that the representations underlying abstract and concrete concepts are at least partially similar. In the next section, we will compare theories that focus on fundamental distinctions between concrete and abstract concepts with theories that focus on similarities between concrete and abstract concepts.

3. Explanations of abstract concepts

3.1. Dual code theory

While the Perceptual Symbols Theory and related frameworks advocate an extreme view where ultimately all representations are grounded in sensory-motor systems, there are others who propose a more moderate view. Mahon and Caramazza (2008) argued that the cognitive system might be only partially grounded in sensory-motor systems. They suggest that sensory-motor information might be co-activated with other, more amodal information during representation, supporting representations through feedback rather than forming the entire basis for representations. Thus, representations themselves could still be largely amodal with only some grounding in sensory-motor systems (see also Dove, 2009). The idea that different representational systems exist is related to Paivio’s dual code theory (e.g., Paivio, 1991). This theory proposes that two independent systems underlie cognitive processing: a verbal and a nonverbal system. The nonverbal system uses sensory-motor analogues (called *imagens*) for representation, while the verbal system represents the visual, auditory, and perhaps motoric aspects of linguistic units such as words (called *logogens*). Connections between the systems exist such that non-verbal information can be named and conversely that names can evoke non-verbal experiences such as images. This theory explains why performance in memory tasks is often better for words that refer to concrete entities than for words that refer to abstract entities. According to the dual code theory, this effect occurs because concrete entities can be represented by both systems, whereas abstract concepts can only
be represented by the verbal system. As a result, more retrieval cues can be used for concrete than abstract entities.

It is important to note that the verbal system does not use abstract or amodal symbols such as propositions to represent verbal information. Rather, it represents the visual and auditory features of words and sentences. Therefore, it is unlikely that this system is suited to represent the meaning of abstract words (Barsalou, 2008). Rather, there may be associations between different verbal representations, without grounding these representations in meaningful experiences. Such associations might support correct performance in tasks that require only superficial activation of meaning, such as the free association task (Barsalou et al., 2008; Solomon & Barsalou, 2004), although Crutch, Conell, and Warrington (2009) showed that even in tasks requiring deeper meaning abstract concepts rely more on associative relations whereas concrete concepts rely more on similarity relations.

3.2. Linguistic co-occurrence

Related to this verbal association idea are models based on statistical analyses of language such as LSA (Latent Semantic Analysis; Landauer & Dumais, 1997) and HAL (Hyperspace Analogue to Language; Burgess, 1998). These models represent word meaning as points in a high dimensional space based on their patterns of co-occurrence with other words. The basic principle is the extraction of meaning representations from large corpora of texts. For each word a frequency count is made of its occurrence in many different contexts. Contexts can be the other words within a window (as is done by the HAL model) or paragraphs of texts (as is done by LSA). These frequency counts can be put in vectors such that each word has a single vector containing all its frequency counts. This vector of co-occurrence values for that word can be viewed as a reference to a point in multi-dimensional space that corresponds to semantic memory. Words with highly similar meanings are closer in this semantic space than words with less similar meanings. It is important to note that vector similarity is not a measure of direct co-occurrence. For example, the words kitten and puppy may have very similar co-occurrence vectors, yet they may hardly occur together. Instead, they frequently occur in similar linguistic
contexts. LSA can compute vectors for single words and for larger units such as sentences. The cosine between two vectors reflects the similarity of the vectors, and can be used as a measure of semantic similarity or text coherence. Landauer and Dumais (1997) suggest that language learning starts out by learning words in experiential contexts. By the time children have been in school for several years, however, most of their new words are acquired through text reading. According to these researchers, learning the meanings of these new words thus depends on knowing the meanings of other words.

These co-occurrence models are quite successful at predicting human performance (Burgess & Lund, 2000) and have been presented as evidence that meaning can be represented by purely symbolic systems (Louwerse & Jeuniaux, 2008). The advantage of these models is that they do not need to distinguish between different types of words such as concrete and abstract, and they have tremendous computational power. One should be cautious in interpreting the co-occurrence vectors as meaningful, however, because it is not clear what they represent. The corpora that provide the input to these models are texts produced by people. Thus, they are measures of how people produce language, which reflects how people think. The abstract nature of the vector representation and the corpus it is based on has no bearing on whether people’s representations are modal or amodal. Thus, the fact that there is systematicity in how words relate to each other, and that two different measures of how people use language (e.g., the co-occurrence vectors and performance in experimental tasks) are related does not by itself explain how meaning is represented. Furthermore, although the co-occurrence vectors are derived from linguistic input, they might reflect the actual distribution of the words’ referents across real-world situations. After all, language typically refers to objects and situations in the world. Words might co-occur in language because their referents co-occur in the real world. For example, co-occurrence models consider the words kitten and puppy as similar because they often occur in similar linguistic contexts (e.g., sentences). However, the reason that they occur in similar sentences might be because these sentences typically describe real-world situations in which kittens and puppies occur. In the real world, kittens and puppies are often found in similar contexts (e.g., in and around the house). As a result, the sentences describing
these contexts will be similar. Indeed, a recent study by Louwerse and Zwaan (2009) showed that the spatial distance between pairs of US cities can be derived from their distributions across texts. In the words of Louwerse and Zwaan (2009), “cities that are located together are debated together.”

3.3. Hybrid models

Some researchers suggested that meaning may be partially grounded in sensory-motor processing and partially in relations between words. Andrews, Vigliocco, and Vinson (2009) developed a model in which sensory-motor information and linguistic information are treated as a single combined data set (for similar work, see Steyvers, in press). They used feature production norms as a proxy for sensory-motor experiences, and word co-occurrences as linguistic input. The word co-occurrence data were based on an approach similar to LSA. Andrews et al. argued that a probabilistic model that was based on a combination of linguistic and sensory-motor data was a better predictor of human performance than models based on either source of data alone or based on the two sources independently.

Thus, instead of language and sensory-motor representations as being separate, parallel processes, Andrews et al. (2009) propose that they are part of the same network. Language binds different sensory-motor experiences to form coherent categories. During language processing, meaning is represented by a pattern of activation across sensory-motor systems. However, in tasks in which performance can be based on word associations, activation of other words might be sufficient. The advantage of such shortcuts is that the system does not need to wait for full activation of representations in the sensory-motor system. In order to fully represent meaning, however, sensory-motor representations are necessary. Such a mechanism was proposed by Barsalou et al. (2008) and Simmons et al. (2008). They argued that when linguistic materials are processed, linguistic associations are activated first. In turn, these linguistic representations very quickly activate sensory-motor simulations.

3.4. Abstract concepts are represented by situations
On such an account, the important issue still remains how sensory-motor systems can represent abstract concepts. Barsalou and Wiemer-Hastings (2005; see also Barsalou, 1999b) proposed that in the case of abstract concepts, the specific situations in which those concepts occur as well as the introspective experience in response to those concepts might be simulated. Representations might be formed by collections of concrete situations that share the abstract concept. Such a mechanism is central to exemplar models of categorization (Hintzman, 1986; Nosofsky, 1988). In exemplar models, each experience with an exemplar is stored. Abstraction across exemplars is achieved when a cue activates several different exemplars. The combined activation as a response to the cue results in a kind of summary representation. In a similar way might a cue for an abstract concept activate several concrete situations in which that abstract concept played an important role. Representation by concrete situations allows these abstract concepts to be grounded in sensory-motor simulations.

To investigate the content of abstract concepts, Barsalou and Wiemer-Hastings asked participants to produce properties for concrete and abstract concepts. They found that introspections and situations were important aspects for all types of concepts, but even more for abstract than concrete concepts. The introspective and situational experiences can be simulated by sensory-motor systems because they have the perceptual and motoric details that abstract concepts in isolation lack.

The role of situations is suggested by the context availability model proposed by Schwanenflugel and Shoben (1983; Schwanenflugel, Harnishfeger, & Stowe, 1988). They argue that in order to understand a word or sentence, people need to represent a context in which the linguistic material has meaning. They further argue that the main difference between concrete and abstract concepts is that it is more difficult to find an appropriate context for abstract than concrete materials. They presented sentences with and without contexts, and found that the difference in reading time between abstract and concrete materials that existed in the isolated condition disappeared when a context was provided. Barsalou and Wiemer-Hastings (2005) argued that abstract concepts are used in a wider variety of contexts, and that typical events for abstract concepts are much more complex than those for concrete concepts. Barsalou (1999b; 2003) further proposed that
abstract concepts might identify complex relations between physical and mental events. Such relations might be spatial, temporal, or causal. These relations are simulated by ‘filling in’ the regions that are connected by the relations with specific entities, events, or mental states.

To sum up, abstract concepts might be grounded in sensory-motor simulations in an indirect way. Words for abstract concepts might activate specific, concrete situations that are instances of the concept or that provide a context for the concept. There is at present still very little evidence for this view, so more research will be needed in order to draw stronger conclusions.

3.5. Conceptual metaphor theory

A very different proposal originated in cognitive linguistics, where researchers have suggested that abstract concepts are understood by way of analogy to representations of concrete, embodied experiences. For example, people may understand the process of solving a problem in terms of travelling from a starting point (the problem situation) to a destination (the solution) along a path (the method that is used to solve the problem), as is illustrated by expressions such as to get sidetracked or to have something get in one's way (Lakoff, 1987; Lakoff & Johnson, 1980, 1999; Gibbs, 1994, 2005; but see Barsalou & Wiemer-Hastings, 2005). In this view, the concrete concept is used as a metaphor (the ‘vehicle’) to represent the abstract concept (the ‘topic’). Thus, the concrete situation of travelling along a path provides a metaphor for the abstract situation of solving a problem. Lakoff and Johnson used the term conceptual mapping and conceptual metaphors because they argue that metaphorical mappings that are found in linguistic expressions (e.g., love is a journey) are conceptual and occur even beyond language comprehension. The basic claim of the conceptual metaphor theory is that the concrete vehicle partially structures the mental representation of the abstract topic and that the mental representation of the vehicle is necessary to fully understand the topic (Lakoff & Johnson, 1980, 1999; Johnson, 1987). Because the metaphor’s vehicle refers to a concrete physical experience, conceptual metaphor theory can explain how the sensory-motor system can be used to represent the abstract concept.
Lakoff and Johnson (1980) proposed that there is a set of basic concepts that are central to human experience. These basic concepts on which concrete concepts as well as embodied conceptual metaphors (or primary metaphors, Gibbs, 2006; Grady, 1997) are based have been called image schemas (Johnson, 1987; Lakoff, 1987). These image schemas are concepts such as source-path-goal, containment, or balance. In general, they are analogue representations of mostly spatial relations and movements, although they are not sensory-motor representations themselves (Gibbs, 2006). Rather, image schemas give structure to experiences across modalities, and this multi-modal nature makes them more abstract than actual sensory-motor representations. At the same time, they are viewed as grounded because they originate from sensory-motor experiences. Therefore, if such image schemas are grounded in sensory-motor experiences, and abstract concepts are partially represented by image schemas, one could argue that abstract concepts are ultimately grounded in sensory-motor experiences as well. It is important to note, however, that image schemas are abstractions from experience, so this logic relies on a two-step grounding operation.

Lakoff and Johnson (see also Gibbs, 1994, 2005) thus argue that image schemas are fundamental to human experience and are used to understand more abstract concepts. They argue that this metaphorical mapping is reflected in language. Therefore, most evidence for this idea has come from analyses of linguistic expressions. Only recently have researchers started to collect experimental evidence for the metaphorical representation of abstract concepts, although some earlier indirect evidence comes from studies on metaphorical language comprehension. These earlier studies investigated whether image schemas were activated after people had read metaphorical statements. For instance, Allbritton, McKoon, and Gerrig (1995) obtained priming effects between words in a recognition memory test if these words were related to the same metaphor in a previously read text passage. Gibbs, Bogdanovich, Sykes, and Barr (1997) investigated priming effects for metaphor related words in a lexical decision task. They used metaphorical mappings such as life is a journey. For example, after reading the metaphorical sentence She was over the hill there was more priming for the word journey than after the literal sentence She was getting old. There was no difference in priming between the literal sentence and an unrelated control sentence. Thus, these studies
support the idea that metaphorical sentences activate the concrete vehicle (but see Keysar, Shen, Glucksberg, and Horton, 2000; McGlone, 1996 for failures to find evidence for such activation).

3.5.1. Evidence for conceptual metaphors and methodological issues

Most researchers who investigated the role of image schemas were interested in how people process metaphorical language, and did not directly address the representation of abstract concepts beyond metaphorical language. If metaphorical mappings are used only to comprehend metaphorical language, effects of such mapping should be found exclusively during processing of metaphorical sentences (Gibbs, 1992, 1994). When people comprehend metaphorical language, image schemas may be activated simply by words that also literally refer to the concrete vehicle (e.g, *prices have risen*). On the other hand, the conceptual metaphor theory claims that metaphorical mappings are used more generally to represent abstract concepts. Therefore, the image schema should be part of the representation itself, and activation of the image schema is needed for full understanding of the abstract concept. In order to support the idea that metaphorical mappings are conceptual rather than linguistic we need evidence that image schemas are also activated when no metaphorical language is used. Because the studies described above did not test this directly, they do not allow the conclusion that conceptual metaphors are used outside metaphorical language.

Much of the evidence discussed in the previous section has used metaphorical language in some way. The conceptual metaphor theory, however, claims that metaphors are the expression of an underlying conceptual mapping. An interesting line of evidence comes from studies that show image schemas in people’s spontaneous gestures as they talk about abstract concepts (Cienki & Müller, 2008; Núñez & Sweetser, 2006). Future studies should test whether such gestures are nonverbal expressions of the linguistic metaphors (e.g., someone gesturing ‘straight’ to express the linguistic metaphor *straight is honest*) or reflect embodied image schemas (e.g., the concept *honesty* is represented by *straightness*).
Spatial image schemas can be mapped on various abstract domains, such as valence (happy is up and sad is down), power (powerful is up and powerless is down), and divinity (god is up and the devil is down). For example, Meier, Hauser, Robinson, Friesen, and Schjeldahl (2007) investigated whether the up-down image schema is activated by the abstract concept divinity. They used god-like, devil-like, and neutral pictures. First, they asked participants to study the pictures that were presented at different, random locations on the screen. Following the study phase, they tested participants’ memory for the location of each picture. Their results showed that participants remembered god-like pictures at a higher location than neutral pictures and devil-like pictures at a lower location than neutral pictures. Thus, even when no linguistic reference was made to the up-down image schema, it still affected spatial memory. Comparable results were found by Giessner and Schubert (2007), who investigated the relation between the up-down image schema and power. They demonstrated that the amount of power attributed to a manager was affected by spatial aspects, such as the vertical distance between the manager and his subordinates in an organizational chart. Participants read a brief description of a manager in a fictive organization and were shown a typical organizational tree-diagram in which the members of the organization were depicted as boxes connected by lines. The manager was judged as more powerful if the vertical line between the manager and his subordinates was longer than if it was shorter, all other factors (number of subordinates, leader’s position in the tree diagram) remaining equal.

Another study, by Casasanto and Boroditsky (2008), investigated the spatial representation of time. They presented moving dots and manipulated the spatial displacement and duration of the movement independently. When they asked participants to estimate the duration of the movement, time estimations were affected by the spatial distance that the dot had travelled. This result suggests that people use the mental representation of space in order to fully understand time.

An important question regarding these studies is whether the image schema plays a role during representation of the abstract concept or during selection of the response (Boot & Pecher, 2010). There is some evidence that irrelevant information can sometimes affect responses in situations with high uncertainty. For example, participants’ judgments
of justice in conditions of uncertainty (e.g., when there is no information about the others’ outcome) was influenced by irrelevant information (e.g., affect, Van den Bos, 2003; Van den Bos, Lind, Vermunt, & Wilke, 1997). In studies investigating activation of image schemas during processing of abstract concepts a similar type of uncertainty might play a role as well. When participants have to choose from many response options (e.g., duration of a stimulus, power of an unfamiliar person) and are uncertain about the accuracy of their choice, irrelevant information may affect responses. For example, in Giessner and Schubert’s study, participants had no idea how powerful the person in the chart really was. It is possible that the irrelevant information (position in the chart) might have influenced the response selection in congruence with the metaphor (e.g., people presented at the top are more powerful than people presented at the bottom). In sum, while it is obvious that the conceptual metaphor (e.g., power is up) was active during performance, it is unclear whether the image schema affected the representation of the person’s power or the response selection process.

There are some studies in which the image schema was unlikely to have affected response selection. Several researchers have found that identification of a stimulus (e.g., a p or q) at the top or bottom of the screen is affected by a prior event in which a word had to be judged on an abstract quality such as valence (Meier & Robinson, 2004) or power (Van Dantzig, 2009; Zanolie et al., 2010) or even when participants listened to sentences containing abstract verbs such as respect or argue (Richardson, Spivey, Barsalou, and McRae, 2003; Spivey, Richardson, & Gonzalez-Marquez, 2005, but see Bergen, Lindsay, Matlock, & Narayan, 2007). Van Dantzig asked participants to judge whether a word (e.g., president, slave) presented in the center of the screen represented a powerful or powerless entity. Immediately following the power judgment a letter was presented at the top or bottom of the screen, and participants identified this letter as quickly as possible. Because the target response (the letter identity) was unrelated to the image schema, it is highly unlikely that participants used the image schema in order to select their response. Rather, making the power judgment directed attention to the congruent spatial location (up for powerful and down for powerless). When the subsequent letter appeared in the attended location response was facilitated compared to when it was presented in the unattended location. Using the same paradigm in an EEG study, Zanolie et al. measured
ERPs (Event Related Responses, a measure of the brain’s electrical signal in response to an event such as presentation of a stimulus) to the letter targets. They found an early difference in brain activation, the N1, which is a negative deflection occurring at 160-200 msec after stimulus presentation. This component has been observed in many studies of spatial attention, and is assumed to reflect an enhancement of stimuli at the attended location (Luck & Hillyard, 1995; Mangun & Hillyard, 1991). This study thus provides further evidence that the image schema influenced spatial attention.

Studies that do not require participants to select a response provide additional evidence that the effects of image schemas are not due to uncertainty. Teuscher, McQuire, Collins, & Coulson (2008) obtained evidence for the automatic activation of a spatial image schema when participants mentally represented time. They presented sentences that described events involving space or time. On half of the trials, the sentence had an ego-moving perspective (She approached the buffalo, We got near the day of the exam) and the other half had an object-moving perspective (The buffalo approached her, The day of the exam was getting closer). Following the sentence they presented a cartoon of a smiley face and a geometric object that represented a congruent or incongruent spatial movement. Rather than asking participants to respond to the cartoons they only asked them to view the cartoons while their EEG was measured. The results showed effects of congruency on the ERPs that were time-locked to the cartoons, with effects of spatial congruency affecting an earlier portion of the ERP waveform for descriptions of events in space than for descriptions of events in time.

Finally, several studies have shown effects of image schemas on processing of abstract concepts using tasks that were so easy that it was very unlikely that participants would have used the task-irrelevant image schema in a strategic way. Boot and Pecher (2010) investigated the role of distance for the representation of similarity. They asked participants to make similarity judgments to colored squares that were presented at various distances from each other. The colors were either very similar or very different, making the task quite easy to perform. Still, an interaction effect was found between distance and similarity. Specifically, for similar colors, responses were faster when the spatial distance between the colored squares was small than when it was large, whereas the opposite was found for dissimilar colors (see Casasanto, 2008, for interactions
between distance and similarity in more difficult tasks). In another study, Boot and Pecher (in press) investigated the role of the container image schema for the representation of categories. The idea was that categories can be represented as containers (Lakoff, 1987). Things that belong to a category are represented as being inside the container, while things that do not belong to the category are represented as being outside the container. They asked participants to categorize pictures of animals and vehicles. Crucially, pictures could be presented inside or outside a visual frame. Category decisions were faster when pictures were presented congruent with the image schema than when they were presented incongruent with the schema. Again, the task was so easy that it was unlikely that participants used the task-irrelevant frame for response selection.

To sum up, many studies have obtained evidence that image schemas play a role in representations of abstract concepts. Effects have been found in a variety of tasks for several abstract concepts. Moreover, several studies have found such effects without using metaphorical language, thus providing evidence for the claim that metaphorical mappings are conceptual rather than linguistic.

### 3.5.2. Evaluation of conceptual metaphor theory

An important issue for theories on grounded cognition is whether abstract concepts automatically activate a concrete vehicle or image schema. The evidence presented in the previous section seems to support this view. However, there are also findings that may be more problematic. For example, McGlone (1996) failed to find evidence for activation of the underlying vehicle when subjects were asked to paraphrase metaphors or to come up with another metaphor for the same concept. Keysar et al. (2000) found that processing of sentences containing expressions of metaphorical mappings (e.g., *the journey should not last too long* for the metaphor *argument is a journey*) was not facilitated by earlier conventional expressions of the same metaphorical mapping (e.g., *point out your position*), even when the mapping was explicitly mentioned (e.g., *think of an argument as a journey*). They did find facilitation when novel expressions of the same metaphor (e.g., *parking at a compromise*) were used. These results question the assumption of conceptual metaphor theory that metaphorical
mappings are activated automatically when people represent the topic of the metaphorical expression. Rather, metaphors may be understood as categorizations (Glucksberg & Keysar, 1990) or may evolve from explicit analogies when they are novel to categorizations or polysemy (i.e., words that have multiple related meanings) when they have become conventional (Bowdle & Gentner, 2005). According to the category view, a metaphorical sentence (e.g., lawyers are sharks) can be understood when a post-hoc category is formed by selecting relevant features and deselecting irrelevant features (e.g., vicious, predatory but not fast swimmer). This selection process, and thus formation of the post-hoc category, can only occur when both the topic and vehicle are activated.

According to the analogical view, a metaphorical sentence (e.g., lawyers are sharks) can be understood when the relational structure (e.g., harming others) of the concrete vehicle (e.g., shark) and topic (e.g., lawyer) are aligned. Indeed, being able to see structural similarities might be at the very core of representing and understanding abstract concepts (Gentner, 2003). This alignment process can only occur when both the topic and vehicle are active.

Similarly, Coulson and Van Petten (2002; see also Coulson & Matlock, 2001) propose that comprehension of metaphorical mappings requires blending of the topic and the vehicle in a blended space. Such blending is a process by which a subset of features from the vehicle and topic are combined to form a new representation. In order for this to happen, vehicle and topic first need to be represented separately. Coulson and colleagues showed that this process is similar for metaphorical (her voice was sweet syrup) and literal blends (in the movie Psycho, the blood was really cherry syrup). The literal blends were slightly easier to process than the metaphorical blends, but both were harder than simple literal meanings (one of Canada’s major exports is maple syrup). Giora (2002) also noted that there is no fundamental difference between literal and metaphorical meanings. Rather, she posits that saliency (e.g., frequency, contextual fit) determines which meaning is activated first, the literal or metaphorical.

On the other hand, according to the conceptual metaphor theory, the concrete vehicle partly structures the abstract topic. Because representation of the topic depends on the vehicle, mapping of the vehicle on the topic domain should occur before the topic is fully represented. Such mapping can only work for conventional metaphors because
these are already part of the abstract concept. Therefore, novel metaphor comprehension should rely on a different mechanism such as analogy or blending. Another important difference between novel and conventional metaphors is that novel metaphors may be used deliberately by the speaker in order to force the listener to see a concept in a new perspective, whereas conventional metaphors do not have this communicative goal (Steen, 2008). In contrast, conventional metaphors might reflect conceptual mappings.

Some of the experiments described earlier indicate that metaphorical language can lead to the activation of image schemas, although the evidence was mixed. Lakoff and Johnson’s claim is much stronger, however, because it states that people use metaphors to understand concepts. To prove this claim one needs to show not only that abstract concepts activate image schemas, but rather that activation of an image schema precedes full understanding of the abstract concept. For example, processing of the sentence Our relationship was at a crossroads may activate the source-path-goal (or journey) schema. What must be shown, however, is that the source-path-goal schema needs to be activated in order to understand the concept love. In other words, the source-path-goal schema should be necessary for the representation of love. Thus, the image schema should be activated irrespective of whether the abstract concept is presented in a metaphorical or literal sentence.

Commenting on this issue, Murphy (1996) argued that the view of conceptual metaphor theory can make strong or weak claims as to how fundamental metaphorical mapping is to the representation of concepts. In the strong view, the abstract concept does not have its own structure but derives all of its structure from the concrete concept. A serious problem with the strong view is that if the abstract concept does not have a structure of its own, it is impossible to interpret the metaphorical relation with the concrete concept. For example, which features of journey overlap with love must be determined by the similarities between the two concepts. How else would someone be able to understand that the statement This marriage has been a long bumpy road means that the relationship has had difficulties but not that it is made of tarmac or has a line down the middle. In order to determine the similarities, however, both concepts need to already have a structured representation, and therefore, Murphy argues, conceptual mapping cannot fully account for the representation of abstract concepts.
The weaker version of the conceptual metaphor theory does not hold that representations receive their full structure from conceptual mappings. Rather, the abstract concept has its own structure, but this structure is influenced by the conceptual mapping. It is unclear whether such influence is the result of the metaphorical expression (e.g., *love* becomes more like a *journey* because of the way it is talked about) or whether the metaphorical expressions are the result of the structural similarity between the abstract topic and the concrete vehicle (e.g., because *love* and *journey* share structure people have developed metaphorical expressions that reveal such similarity) (Murphy, 1996).

Proponents of the conceptual metaphor theory, however, have claimed that the systematicity of image schema mappings for abstract concepts as expressed by metaphorical language provides evidence that the metaphorical mappings are conceptual rather than linguistic. They argue that language reflects the underlying representational structure of concepts, and the systematic use of metaphorical mappings indicates that the mapping must be part of the concept itself. Such a framework predicts that image schemas should be activated whenever the abstract concepts are activated, even when people are not using metaphoric language. In other words, thinking about abstract concepts should activate image schemas directly, and not via language. Language is merely the expression of such activation, not the cause. The studies using non-linguistic stimuli that we presented in the previous section provide some evidence that image schemas are activated when abstract concepts are processed non-linguistically.

4. Discussion

We started this chapter by asking how abstract concepts might be grounded in sensory-motor processing. Two grounded views have been proposed; representation by concrete situations and the conceptual metaphor view. Whereas only a few studies have investigated the idea that abstract concepts are represented by concrete situations, many more have investigated the role of metaphor or image schemas. As the evidence reviewed above shows, image schemas are relevant for abstract concepts. Image schemas are activated quickly and automatically even when people are performing simple unambiguous tasks that do not involve metaphoric language. Having established that
image schemas play a role, we should now turn to the question of what that role is. One of the big challenges for applying conceptual metaphor theory to mental representation is clarifying the process of mapping image schemas onto abstract concepts. This turns out to be non-trivial. One of the main problems is that the concept of image schemas is still shrouded in mystery, and the term has been used to mean quite different things. Hampe (2005; see also Gibbs, 2005), for example, noted that there is no set of clear-cut criteria that distinguishes image schemas from other types of mental representations. Gibbs (2006) described image schemas as embodied but at the same time abstracted from modality-specific sensory activation, whereas Grady (2005) defined image schemas as mental representations of concrete sensory experience. Dodge and Lakoff (2005) argued that image schemas are categories of experiences for which the same words are used (e.g., high is used for spatial position, quantity, power, divinity, etc.). Mandler (2005), however, viewed image schemas as the earliest concepts that infants acquire. These image schemas do not have perceptual properties, but at the same time they are preverbal. Gentner (2003) argued that image schemas are abstract rather than concrete. Although this short summary of views is not exhaustive, it illustrates that the idea of an image schema is not well-specified.

The one thing that researchers seem to agree on is that image schemas apply to embodied experiences in different sensory modalities. These image schemas can then be extended to more abstract concepts. Thus, image schemas reflect commonalities between distinct, recurrent experiences in different domains and modalities. The spatial component is an important aspect of image schemas. For example, the up-down, source-path-goal, container, and distance image schemas all refer to spatial experiences, whether they are visual, tactile, or motoric. Image schemas thus may be important for binding perception from different modalities and action together. As an example, in order to have coherent experiences with verticality, the image schema binds visual experiences, proprioceptive experiences, and motor actions so that the representation of visual vertical orientation can be similar to the representation of feeling vertical orientation and the representation of acting in vertical orientations (similar to the idea of event coding, Hommel, 2004). Image schemas thus may be skeletal, providing a basic structure, to
which the flesh (specific perceptual detail, valence, and so on) is added by particular meaning and context (e.g., Gibbs, 2006; Johnson, 2005).

On this account, image schemas are not identical to concrete concepts, because they lack the perceptual detail of the latter. Rather, they might be viewed as features that provide an element of structure to both concrete and abstract concepts. Experiencing verticality in a particular modality may reactivate previous experiences with verticality in other modalities. We can even extend this mechanism to language. Words that refer to verticality might get bound to physical experiences of verticality because they are used in the same situation. Image schemas may thus reflect associations between words and experiences in different modalities.

An important question that remains, however, regards the origin of the connection between image schemas and abstract concepts. This may be a chicken and egg problem. The relation between an image schema and an abstract concept might be formed during early experiences in which the two domains are conflated (Johnson, 1997). For example, early experiences with power typically involve a perception of verticality, such as when a child looks up to a more powerful adult or taller child. These frequent experiences might underlie the relation between the up-down image schema and power. Alternatively, metaphorical language may have activated representations of physical experiences, which then created a link between concrete and abstract domains. For example, expressions such as high power or low status may have activated concrete vertical simulations, thus leading to representations in which power and verticality were combined. Although this view does not explain the historic origin of metaphorical expressions, for individuals it might work because they are exposed to metaphorical language during development. A third possibility is that the image schema and abstract concept have completely separate representations, and people may construct the metaphorical mapping on the fly as an analogy or conceptual blend.

In order to fully represent abstract concepts, however, image schemas are not sufficient. The fact that the up-down schema can be mapped on several different concepts such as power, valence, divinity, and quantity shows that more features are needed to distinguish these concepts, just like the concepts blood, apple, traffic light, and carpet all share the feature red, but this feature is not sufficient to represent these different
concepts. The *up-down* image schema, for example, indicates that concepts all have some scale on which things can be compared (e.g., powerful vs. powerless), but for full understanding more is needed.

Situations or events might provide such additional representational power. Representations of concrete situations and events may provide a full spectrum of perceptual, motoric, and affective experiences (Barsalou & Wiemer-Hastings, 2005). Except for psychological experiments using single words, abstract concepts are hardly ever processed in isolation. Rather, a concept such as power is represented because there is a specific entity (e.g., mother, boss, country) that has power over another specific entity (e.g., child, employee, other country) making it do specific things (e.g., clean up, write a report, make budget cuts). Image schemas provide similar structure to such diverse situations, but the actual details provide the meaning. On this account, image schemas are not the final solution to the grounding problem for abstract concepts. The pervasiveness of image schemas in processing of abstract concepts shows that similarities between concrete and abstract domains are central to understanding. However, the full grounding of abstract concepts in sensory-motor experience needs an additional step. Most likely, the rich perceptual, motoric, and evaluative details of specific situations provide such grounding.
References


Proctor, R. W., & Zhang, Y. (2010). "Mother nature doesn't have a bullet with your name on it": Coding with reference to one's name or object location? *Journal of Experimental Social Psychology, 46*, 336-343.


