COGNITIVE SCIENCE
A Multidisciplinary Journal

Cognitive Science 39 (2015) 1764–1801

ISSN: 0364-0213 print / 1551-6709 online
DOI: 10.1111/cogs.12174

Are Automatic Conceptual Cores the Gold Standard of Semantic Processing? The Context-Dependence of Spatial Meaning in Grounded Congruency Effects

Lauren A. M. Lebois,[a] Christine D. Wilson-Mendenhall,[b] Lawrence W. Barsalou[a]

[a]*Department of Psychology, Emory University*
[b]*Department of Psychology, Northeastern University*

Received 13 June 2013; received in revised form 7 February 2014; accepted 24 February 2014

Abstract

According to grounded cognition, words whose semantics contain sensory-motor features activate sensory-motor simulations, which, in turn, interact with spatial responses to produce grounded congruency effects (e.g., processing the spatial feature of *up* for sky should be faster for up vs. down responses). Growing evidence shows these congruency effects do not always occur, suggesting instead that the grounded features in a word's meaning do not become active automatically across contexts. Researchers sometimes use this as evidence that concepts are not grounded, further concluding that grounded information is peripheral to the amodal cores of concepts. We first review broad evidence that words do not have conceptual cores, and that even the most salient features in a word's meaning are not activated automatically. Then, in three experiments, we provide further evidence that grounded congruency effects rely dynamically on context, with the central grounded features in a concept becoming active only when the current context makes them salient. Even when grounded features are central to a word's meaning, their activation depends on task conditions.

Keywords: Conceptual processing; Lexical semantics; Representation; Grounded cognition; Congruency effects; Context effects; Automaticity

Correspondence should be send to Lawrence W. Barsalou, Department of Psychology, Emory University, Atlanta, GA 30322. E-mail: barsalou@emory.edu

"*To know an object is to lead to it through a context which the world supplies.*"
William James (1909/2008)

1. Introduction

1.1. Accounts of core and peripheral information in semantic access

Research on semantic access often assumes that word meanings have conceptual cores, namely, definitional, typical, and otherwise central information that is activated automatically whenever a word's meaning is processed (Barsalou, 1982, 1989; Conrad, 1978; Greenspan, 1986; Whitney, McKay, Kellas, & Emerson, 1985). By "automatic," we mean that a word's conceptual core becomes active quickly to the same level across all contexts (cf. Moors & De Houwer, 2006). Any other information that becomes active more slowly in a context-dependent manner is considered peripheral. Even when a word is semantically ambiguous (e.g., tick), its multiple core meanings are assumed to become active automatically in parallel initially (Swinney, 1979; Tanenhaus, Leiman, & Seidenberg, 1979). Any potential influence of context occurs after these conceptual cores become active.

A more modern perspective echoes these claims, again assuming that conceptual cores become active automatically before context-dependent information (Dove, 2009; Machery, 2007; Mahon & Caramazza, 2008). In this framing, conceptual cores are viewed as amodal and context-independent, whereas sensory-motor activations that may optionally become active later arise in a more peripheral and context-dependent manner. Indeed, modal activations may not occur on occasion for some concepts (Dove, 2009). The argument that amodal representations of concepts reside in the anterior temporal lobes often similarly assumes that concepts have amodal cores that are more central to semantic processing than modality-specific information (e.g., Chan et al., 2011; Ralph, Cipolotti, Manes, & Patterson, 2010).

Thus, one important class of theory proposes that conceptual cores become active automatically across all contexts before peripheral context-dependent information (what we will call "core first accounts"). From this perspective, context-independent automaticity is the gold standard of semantic processing. Automatically activated semantic features that are context-independent constitute the most important semantic information for a word, with dynamic context-dependent features being relatively irrelevant.

Alternatively, many other accounts propose the opposite, namely, superficial information tends to become automatically active initially, with conceptual cores only becoming active later when needed ("core last accounts"). Classical theories of concepts, for example, often propose this pattern of activation. According to Smith, Shoben, and Rips (1974), both peripheral (characteristic) features and core (defining) features become active initially during conceptual processing, with only the core features remaining active later, if necessary; as a result, peripheral features play an especially large role in early processing. When processing the concept of water, for example, peripheral features like wet and clear may dominate early processing, whereas core features like H_2O may only dominate later (cf. Malt, 1994). In the same spirit, Keil and Batterman (1984) argued that peripheral features are acquired during development before core features, such that it would not be surprising if, in a speeded deadline task, characteristic features dominated processing before defining features. Many other researchers have similarly argued that superficial information dominates conceptual processing early, whereas deep information dominates later only when further processing activates it (e.g., Blanchette & Dunbar, 2000; Forbus, Gentner, & Law, 1995; Gentner & Forbus, 1991). From this perspective, the deepest most defining information is not activated automatically in a context-independent manner, but only becomes active when necessary.[1]

Still other researchers have proposed that the distinction between core context-independent and peripheral context-dependent features is unnecessary and even misleading ("no-core accounts;" McCloskey & Glucksberg, 1979; also see Evans, 2009). Because concepts are fuzzy, they have no core, defining, context-independent features that are activated automatically. Instead, all conceptual features are dynamic and context-dependent to some extent. Rather than some core features being automatic, all features vary in activation stochastically as a function of context.

Even when a category, by definition, has no core features, people may still create the illusion that core features exist (e.g., Brooks & Hannah, 2006). Similarly, people often create the fiction that a category has an essence (e.g., Gelman, 2003). As findings like these and others illustrate, the construct of conceptual cores is intuitively appealing (Barsalou, Wilson, & Hasenkamp, 2010). For this reason, it is essential to assess the presence of conceptual cores empirically, rather than to simply assume theoretically that they exist.

As we review in the next section, much research argues against conceptual cores, demonstrating that semantic access for central conceptual features is dynamic and context-dependent. Even the most accessible information during semantic access generally appears to be context-dependent, not automatic.

1.2. Accumulating evidence against the automatic processing of symbolic stimuli

Many classic symbolic phenomena believed to exhibit automaticity—such as Stroop, SNARC, and Simon effects—are actually dynamic and context-dependent. In each of these phenomena, the critical stimuli are actually not processed automatically in a rigid ballistic manner across contexts (e.g., color words in the Stroop task). Instead, their processing varies dynamically across task conditions (e.g., Stroop interference varies from context to context). If symbolic stimuli do not exhibit automaticity in these classic "automaticity" phenomena, then the semantic access of features in concepts is unlikely to exhibit automaticity either. In a recent review of different literatures, Gawronski and Cesario (2013) similarly concluded that many additional phenomena generally assumed to be automatic are actually not. Kiefer, Adams, and Zovko (2012) reach a similar conclusion.

We believe that it is important for research on conceptual processing to take stock of findings that bear on automaticity in related literatures. To the extent that automaticity does not occur across a diverse array of other research areas, why should it occur for conceptual processing? Why should conceptual cores that are activated automatically exist? Why should only conceptual processing exhibit automaticity and not other symbolic phenomena?

1.2.1. Stroop effects

In the Stoop paradigm (Stroop, 1935; MacLeod, 1991), participants read color words (e.g., red, green) printed in both congruent and incongruent font colors, and indicate the font color of the word. When the color word and font color are incongruent (e.g., the word "purple" printed in yellow font), responses are slower and more erroneous compared to congruent pairings (e.g., the word purple printed in purple font).

Although the Stroop effect is thought to reflect the automatic access of color information for color words, growing evidence demonstrates that the effect can be modulated significantly. Researchers have reduced and eliminated the Stroop effect by varying the proportion of congruent to incongruent trials (Jacoby, Lindsay, & Hessels, 2003), by varying the frequency of congruent and incongruent trials for specific word-color pairings (Jacoby et al., 2003), by coloring a single letter in the color word instead of the whole word (Besner, Stolz, & Boutilier, 1997), by cuing attention to a single letter in the color word instead of to the whole word (Besner & Stolz, 1999), and by priming the concept of dyslexia (Goldfarb, Aisenberg, & Henik, 2011). Thus, semantic access of color information from color words in the Stroop paradigm is dynamic and context-dependent, not automatic, and context-independent.

1.2.2. SNARC effects

In the Spatial Numerical Associations of Response Codes (SNARC) paradigm, participants make a judgment about a number stimulus (e.g., parity, orientation, font color). Across judgments, response times for small numbers tend to be faster when responses are made with the left hand, whereas response times for larger numbers are faster with the right hand (Dehaene, Bossini, & Giraux, 1993; Keus & Schwarz, 2005).

The standard explanation of this congruency effect between magnitude and space is that participants simulate a mental number line increasing in magnitude from left to right, and respond faster when the spatial location of their response is congruent with their spatial simulation of magnitude. Thus, participants are faster when responding to large numbers with their right hand than when responding with their left hand (and vice versa for small numbers). The standard explanation often further assumes that magnitude information for a given number is accessed automatically and independently of context.

Evidence indicates, however, that the SNARC effect is context dependent, being modulated by the salience of number magnitude in the task. The SNARC effect disappears, for example, when participants simply judge the font color of the number instead of a more magnitude-related property like parity (Fias, Lauwereyns, & Lammertyn, 2001). Rather than magnitude information being accessed automatically for a number, its accessibility varies dynamically with task conditions.

1.2.3. Perceptual similarity effects in semantic access

Across tasks that assess pronunciation, lexical access, and semantic evaluation, responses to target words tend to be faster and more accurate when a related word prime precedes them (e.g., Meyer & Schvaneveldt, 1971; Schreuder, Flores D'Arcais, & Glazenborg, 1984). When, for example, participants indicate whether *coin* is a word or nonword, they are faster to respond when the preceding prime is semantically related (*money*) than when it is unrelated (*rabbit*).

In one version of word priming tasks, perceptually similar semantic primes speed responses to subsequent targets (e.g., Pecher, Zeelenberg, & Raaijmakers, 1998). Verifying that *coin* is a word, for example, is facilitated by first processing a perceptually similar object prime (e.g., *pizza*). When the feature of roundness for *pizza* becomes active, it activates roundness for *coin*, thereby speeding the access of semantic information for *coin* and identifying it as a word. In this, and many other lexical processing phenomena, the priming of semantic features is often assumed to occur automatically, in primes and/or targets.

Growing evidence demonstrates, however, that task conditions significantly modulate perceptual priming. Perceptual priming typically occurs when an orienting task draws attention to the perceptual features of a word's referents (e.g., judge whether word's referents are oblong or flat before the critical lexical decision task; Pecher et al., 1998; Yee, Ahmed, & Thompson-Schill, 2012). Perceptual priming does not occur, however, when the salience of perceptual features is diminished by not including a perceptually focused orienting task before the critical task (Pecher et al., 1998; Yee et al., 2012), or by varying salience across multiple dimensions (e.g., including both highly semantically related and perceptually related words; Pecher et al., 1998).

1.2.4. Affective priming effects

Across diverse affective priming paradigms, a general finding is that an initial stimulus can prime task-irrelevant affect implicitly, thereby influencing processing of a subsequent stimulus (Herring et al., 2013). For example, participants might first see a picture having a positive or negative valence (e.g., a burned house vs. a sunset) before judging the valence of a target word (e.g., wedding). When the primed affect is congruent with the affect of the target stimulus or task response, response times are faster compared to incongruent trials. Affect can be primed implicitly in diverse manners via primes (e.g., Storbeck & Robinson, 2004), target stimuli (e.g., Rotteveel & Phaf, 2004), and tasks (e.g., Spruyt, De Houwer, Hermans, & Eelen, 2007; Storbeck & Robinson, 2004).

Although affective priming is often assumed to occur automatically (e.g., Bargh, 1997), recent research demonstrates that manipulating the salience of valence modulates it. Affective priming is diminished, for example, when participants must only make affective judgments for 25% of the stimuli (Spruyt et al., 2007). Affective priming can be removed altogether if the task requires participants only to make nonaffective judgments about affective stimuli (Rotteveel & Phaf, 2004), or if the stimuli vary in both valence and semantic category (Storbeck & Robinson, 2004). As findings like these illustrate, accessing affective information from words and pictures is dynamic and context-dependent, not automatic.

1.2.5. Attentional cuing effects

Spatial cuing paradigms demonstrate that attention to a symbolic task cue is not generated automatically in a context-independent manner. These paradigms are of particular interest because attention researchers have used them to assess whether automaticity exists as an attentional process. Should attention researchers fail to find automaticity in attention paradigms, it seems highly unlikely that automaticity would occur in other paradigms that utilize attention as a component process.

In one of these attentional paradigms (e.g., Folk, Remington, & Johnston, 1992; Remington, Johnston, & Yantis, 1992), participants simply make a binary response to whether a target stimulus displayed in one of four possible screen locations is one of two characters (e.g., X vs. =). On critical trials, the target is preceded by a spatial cue that can be valid or invalid. When the cue is valid, it occurs in the same location as the subsequent target, draws attention to the location where the target will be, and facilitates detection. Conversely, when the cue is invalid, it occurs in a different location from the target, draws attention away from the location where the target will be, and inhibits detection.

Researchers often argue that such cuing effects result from the automatic, involuntary, and rigid orienting of attention that results from processing the spatial cue (Jonides, 1981). Much evidence indicates increasingly, however, that these cuing effects are context-dependent, with task-specific demands modulating their presence or absence. For example, Folk et al. (1992) found that when the features of an invalid cue matched features of the target (e.g., both flashed rapidly), the cue inhibited performance. When, however, features of the invalid cue did not match features of the target (e.g., the cue was static and the target flashed rapidly), the cue did not inhibit task performance. Such results demonstrate that an inconsistent spatial cue does not always produce inhibition when paired frequently with a target. Depending on feature overlap, the same cue can inhibit the target on some trials but not on others, thereby not functioning automatically.

1.2.6. Simon effects

Simon tasks are similar to spatial cuing tasks, except that the congruency between perception and action becomes important, with this congruency being central to the grounded congruency effects discussed next, and to our experiments later. In visual versions of the Simon task (e.g., Kornblum, Hasbroucq, & Osman, 1990; Lu & Proctor, 1995; Simon, 1969, 1990), an "X" appears either to the left or right of a fixation point, with participants pressing a response button on their left or right side to simply indicate when the stimulus appears (regardless of its location). Although the X's location is irrelevant to the task, participants' responses are facilitated when the X's location and the response button's position are congruent (e.g., X on the left, response button on the left), whereas responses are inhibited when these dimensions are incongruent (e.g., X on the left, response button on the right).

Accumulating evidence demonstrates that the processing of symbolic cues in the Simon effect is highly context-dependent. Similar to the Stroop effect, processing can be modulated by varying the proportion of congruent to incongruent trials (Borgmann, Risko, Stolz, & Besner, 2007), by changing the required task judgment (Lammertyn, Notebaert, Gevers, & Fias, 2007), and by manipulating the salience of a relevant task dimension (Duschere, Holender, & Molenaar, 2008). Conditions with a greater proportion of congruent trials, for example, exhibit an augmented Simon effect, whereas conditions with more incongruent trials cause the Simon effect to invert and become unreliable (Borgmann et al., 2007). A much larger Simon effect occurs when participants are required to evaluate the orientation of the stimulus compared to when they judge stimulus color (Lammertyn et al., 2007). As these results illustrate, extracting the meaning of a symbolic cue is dynamic and context-dependent, not automatic and context-independent.

1.2.7. Grounded congruency effects

In the accumulating literatures on grounded cognition (e.g., Barsalou, 2008), a wide variety of congruency effects have been reported as evidence that simulations underlie conceptual processing. A grounded congruency effect occurs when an irrelevant modality-specific feature of a task biases a critical modality-specific response (for a recent review, see Santiago, Roman, & Ouellet, 2011). The irrelevant feature can facilitate or inhibit the response depending on task materials and requirements (e.g., Estes, Verges, & Barsalou, 2008). Because these experiments typically assume that both the biasing features and the critical responses are processed in the brain's modal systems for perception, action, and/or internal states (Barsalou, 1999, 2008), the congruency effects that result are viewed as grounded.[2]

Imagine, for example, having to indicate whether the word *bird* appears in red vs. blue letters on the computer screen by making an up or down response, respectively, on a keyboard. According to grounded theories, reading *bird* activates a simulation of what it would be like to experience a bird (e.g., Zwaan & Madden, 2005). Besides simulating what a bird might look like, such simulations might also include an implicit attentional orienting response upward, given that birds are often seen in the sky. Once active, these simulations can then interact with a response action to be performed, such as an upward vs. downward response to indicate font color. When the simulated content contains elements that match the response action, response facilitation may occur (congruent condition); when the simulated content mismatches, response inhibition may occur (incongruent condition). When participants, for example, make an upward task response for a concept associated with a high verticality (e.g., *bird*), responses may be faster relative to conditions when a downward response is required (e.g., Brookshire, Casasanto, & Ivry, 2010; Casasanto, 2008; Glenberg & Kaschak, 2002; Zwaan & Yaxley, 2003).

Although research often assumes that grounded congruency effects result from automatic and context-independent activation of modality-specific features (e.g., high vertical position for birds), increasing evidence suggests that these effects are dynamic and context-dependent. As reviewed next, context modulates congruency effects in ways that should not be possible if core conceptual properties are activated automatically independent of context.

In one well-studied task (Meier & Robinson, 2004), words whose meanings are associated with a salient vertical position are presented at the top vs. the bottom of the computer screen. When performing judgments on these words, participants are faster when the vertical position associated with the word's meaning is congruent with the word's screen location (high–high and low–low) than when they differ (high–low and low–high). In subsequent research, however, Santiago, Ouellet, Roman, and Valenzuela (2012) demonstrated that these congruency effects do not reflect automatic processing of the target words, but instead vary substantially with task conditions. By explicitly or implicitly drawing participants' attention to different task dimensions, Santiago et al. were able to make grounded congruency effects come and go. For a congruency effect to occur, participants had to be actively attending to both the relevant task dimension (e.g., stimulus valence) and to the irrelevant biasing dimension (e.g., stimulus location). The critical target words did not automatically produce attentional orienting responses in the vertical dimension. Instead, the presence of congruency effects depended on the overall allocation of attention across relevant cues.

Much research further shows that task conditions modulate grounded congruency effects, indicating that they are dynamic and context-dependent, not automatic. In behavioral research, congruency effects can be eliminated if participants' responses do not require certain critical movements that interact with semantic processing. In Borghi, Glenberg, and Kaschak (2004), for example, congruency effects did not occur when participants' responses required pressing a vertically positioned button on which their hand rested, without having to move their hand up or down to press the button. Similarly, congruency effects do not occur if the task fails to sufficiently orient participants' attention to stimulus-related actions (Bub, Masson, & Cree, 2008; Sato, Mengarelli, Riggio, Gallese, & Buccinoet, 2008; van Elk & Blanke, 2011). Another set of studies only observed congruency effects for motor responses but not for spatial attention (Thornton, Loetscher, Yates, & Nicholis, 2013). Finally, affective congruency effects associated with handedness only occur when attention is drawn to affective valence (de la Vega, de Filippis, Lachmair, Dudschig, & Kaup, 2012).

Related evidence demonstrates that task conditions dynamically modulate the neural activity associated with action-related stimuli. Clearly, action simulations in the motor system *can* become active very early during the semantic access of action words (as fast as 80 ms; Moseley, Pulvermüller, & Shtyrov, 2013; Shtyrov, Butorina, Nikolaeva, & Stroganova, 2014). Nevertheless, much other research demonstrates that action simulations in the brain do not become active automatically. Across multiple neuroscience methods, researchers have reported that action simulations and associated congruency effects are highly context-dependent, including findings from functional magnetic resonance imaging (Papeo, Rumiati, Cecchetto, & Tomasino, 2012; van Dam, van Dijk, Bekkering, & Rueschemeyer, 2011), transcranial magnetic stimulation (Labruna, Fernandez-del-Olmo, Landau, Duque, & Ivry, 2011; Papeo, Vallesi, Isaja, & Rumiati, 2009), and neuropsychological case studies (Papeo, Negri, Zadini, & Rumiati, 2010; Papeo & Rumiati, 2012). Tomasino and Rumiati (2013) provide a recent review.

Another line of work suggests that grounded congruency effects may not occur when linguistic forms represent word meanings instead of modality-specific simulations (Barsalou, Santos, Simmons, & Wilson, 2008). Since Paivio (1971), many researchers have assumed that linguistic forms become active during semantic access (also see Glaser, 1992; Landauer & Dumais, 1997). From this perspective, multiple systems contribute to the dynamic construction of word meaning, including the linguistic system and the simulation system. Under some task conditions, linguistic forms dominate conceptual processing, with simulations not becoming active, whereas under other task conditions, simulations become active and dominate processing. Changes in context may affect the influence and time course of each system's involvement, modulating the accessibility of linguistic forms and simulations. As a result, congruency effects may come and go, depending on whether simulations or linguistic forms dominate processing.[3]

As all of these findings illustrate, grounded congruency effects do not result from the automatic activation of modality-specific simulations. In a recent review, Santiago et al. (2011) observed a general lack of automaticity across a wide variety of grounded congruency effects, arguing instead that they only emerge under certain task conditions. Santiago et al. review many factors that modulate the presence of grounded congruency affects, including (but not limited to): (a) task conditions that influence a concept's activation (e.g., stimulus type, attention orientation, salience of the relevant and biasing task dimensions), (b) factors that influence the interaction between multiple concepts (e.g., simultaneity of processing), and (c) linguistic patterns of the particular culture. All these factors allow individuals to sample dynamically—depending on context—from the large, distributed amount of modality-specific information for a concept in long-term memory.

1.3. Conclusion

We began with three accounts of the time course that underlies semantic access: core first, core last, and no core. The evidence just reviewed supports the no-core account for two general reasons. First, across diverse tasks (including Stroop, SNARC, Simon, etc.), the processing of symbolic stimuli is not automatic, depending instead on task conditions. Second, in tasks that explicitly address semantic access (e.g., perceptual similarity effects in semantic access, spatial congruency paradigms, etc.), the activation of salient semantic features generally depends on task conditions. Rather than concepts containing cores that are activated automatically independent of context, concepts only appear to contain information that is dynamic and context-dependent.

On observing dynamic context-dependent information in a concept, researchers sometimes argue that some other information must constitute its conceptual core, for the sake of conceptual stability. Furthermore, when contextually varying information is modality-specific (e.g., visual and motor simulations), researchers sometimes further conclude that the conceptual cores providing stability must be constructed of amodal symbols (cf. Dove, 2009; Machery, 2007; Mahon & Caramazza, 2008). Evidence for dynamic context-dependent simulations, however, is not evidence for amodal conceptual cores. Even if concepts contain amodal information, it does not follow that this information is automatic and context-independent. Based on the literature reviewed, all information in concepts is likely to be dynamic and context-dependent, including amodal information. Should researchers wish to claim that concepts contain amodal cores, they will need to explain the literature that we have reviewed, and to provide direct empirical evidence that such cores exist.

More generally, the evidence we have reviewed suggests that concepts have no cores at all, amodal or modal. As we have seen, even the most accessible information for concepts varies in accessibility with context. We have not seen evidence that any information is activated automatically across all contexts.

Although none of a concept's features may be automatically activated in a context-independent way, some features may be more statistically likely to be active than others, thereby providing the concept with some stability (e.g., Barsalou, 1987, 1989, 1993). Clearly, some highly salient features often become active quickly across many contexts, with relatively little attention (e.g., Moseley et al., 2013) or intention (e.g., Heyes, 2011), reflecting well-established associative relations between concepts and features. In the Stroop task, color features tend to often become active quickly with relatively little processing, even when irrelevant. In all the other tasks we reviewed, other features can similarly become active quickly with little processing (e.g., spatial features in the Simon task, numerical features in the SNARC effect, affective features in affective priming). As we have also seen, however, context modulates the accessibility of these features extensively, sometimes decreasing their activation, sometimes preventing their activation. Rather than being automatic in a context-independent manner, these salient features of concepts are dynamic and context-dependent.

Together these robust patterns of stability and context-dependent accessibility implicate a general Bayesian perspective on the activation of word meaning: On the one hand, well-entrenched features tend to be active across many contexts, although not necessarily. On the other hand, features relevant in the current context become salient, thereby increasing in accessibility. Thus, on a given occasion, the specific features dynamically active for a word reflect both their overall entrenchment in memory and their current context relevance.

1.4. Overview of experiments

Two recent articles reported evidence that the modality-specific simulations responsible for grounded congruency effects become active automatically (Dudschig, Souman, Lachmair, de la Vega, & Kaup, 2013; Lachmair, Dudschig, De Filippis, de la Vega, & Kaup, 2011). Across varying task conditions, these researchers observed grounded congruency effects consistently. As described in the General Discussion later, specific aspects of these experiments may have produced consistent congruency effects. Importantly, however, it does not follow that these effects will appear under all other task conditions. To the extent that congruency effects fail to appear consistently, it follows that they do not reflect the automatic activation of grounded information. We report a series of experiments next designed to systematically assess whether grounded congruency effects become active under other conditions. As will be seen, we failed to find congruency effects on multiple occasions, suggesting that they reflect the dynamic context-dependent activation of grounded information.

Across three experiments, we explored two manipulations designed to assess the automaticity of grounded information. In Experiment 1, we manipulated whether task instructions drew participants' attention to spatial features or not. In Experiment 2, we manipulated whether trial-by-trial orienting tasks drew participants' attention to spatial features. We implemented these two manipulations in a paradigm shown recently to produce grounded congruency effects (Brookshire et al., 2010; Casasanto, 2008). On each trial, the word for a concept appeared in red or blue font, with participants simply indicating the font color, using an upward or downward response motion (the word's meaning was irrelevant). When a word appeared in red font, for example, participants made an upward response on a vertically positioned keyboard, but when it was blue, participants made a downward response. All critical target words had high vs. low verticality as a central conceptual feature (e.g., *sky* vs. *basement*). If these height features constitute core information accessed automatically, then congruency effects should always occur when the respective words are processed. According to grounded accounts of conceptual processing, attentional orienting to height should be simulated and produce a congruency effect. When this simulation matches the response direction for indicating word color (e.g., *sky* and an upward response), response times should be faster relative to trials when simulation and response direction conflict (e.g., *sky* and a downward response).

In Experiment 1, initial instructions either informed participants explicitly that verticality would vary across the words (aware condition) or stated nothing about this manipulation (unaware condition). If core spatial features are activated automatically in a context-independent manner, then verticality awareness should not modulate their accessibility. When responding, participants should exhibit congruency effects in both the aware and unaware conditions. If important spatial features are context-dependent, however, then participants should only exhibit congruency effects in the aware condition.

In Experiment 2, different orienting tasks were used to manipulate whether verticality was salient or not by drawing attention on a trial-by-trial basis to various semantic properties of the spatial target words. In three orienting task conditions (manipulated between groups), participants either judged a word's verticality, concreteness, or abstractness just before responding to its font color. Of interest was whether drawing attention to semantic properties modulated congruency effects, in comparison to two control conditions for which no orienting task was included. If the semantic access of central spatial features is automatic, then congruency effects should occur for all five groups, regardless of whether an orienting task draws attention to particular semantic properties. If accessing spatial features is context-dependent, however, then congruency effects should only occur when orienting tasks draw attention to spatial information.

Experiment 3 was designed to address the finding in Experiment 2 that congruency effects failed to occur across all conditions. Perhaps congruency effects did not occur in some conditions because participants only attended to font color and did not read the words, thereby precluding the automatic activation of spatial features. To address this possibility, participants performed the font color judgment task together with a surprise forced-choice recognition test for the spatial target words. Of interest was whether participants had memory for the target words even under conditions when they did not produce congruency effects, thereby showing that participants had read and processed the words. If participants do not display congruency effects under these conditions, but remember the words presented, then this offers further evidence that the activation of word meaning is context-dependent, with salient spatial features not always becoming active.

2. Experiment 1

2.1. Overview

Using a grounded congruency paradigm, Experiment 1 assessed whether spatial features are activated automatically in a context-independent manner. Experiment 1 also aimed to replicate previous studies that elicited spatial congruency effects using a color judgment task (e.g., Brookshire et al., 2010; Casasanto, 2008). It is unclear whether, in previous studies, participants were aware of the verticality manipulation, which is why

Fig. 1. Apparatus setup. The keyboard was strapped onto a wooden board apparatus with a bungee cord. The end of the keyboard with the “Q” key was positioned at the top, and the end with the “P” key at the bottom. In this setup, the computer screen was raised or lowered to match the participant’s eye level. The height of the keyboard was also adjusted for each participant to ensure that the white home key (“H” key) was aligned horizontally with stimulus presentation on the screen.

their awareness of verticality features was explicitly manipulated in the different experimental groups here. On each trial, participants simply had to indicate whether the font color of a word presented centrally on the computer screen was red or blue. To indicate font color, participants moved their right arm up or down to press a high or low key on a vertically aligned keyboard (see Fig. 1). On congruent trials, participants either pressed the high key for a word referring to a high concept (e.g., *sky*) or pressed the low key for a low concept (e.g., *basement*). On incongruent trials, participants either pressed the high key for a low concept (e.g., *basement*) or pressed the low key for a high concept (e.g., *sky*).

If the accessibility of verticality is context-independent for concepts saliently associated with a spatial position (e.g., *sky, basement*), then awareness of verticality should not affect its accessibility, and in turn, should not modulate congruency effects. To manipulate awareness, we told participants in the aware condition that they would be responding to words associated with salient vertical positions, but we said nothing about this to participants in the unaware condition. If vertical position is activated automatically in a context-independent manner for words associated with a vertical position, then congruency effects should emerge in both the aware and unaware conditions. If congruency effects are context-dependent, however, then they should only emerge in the aware condition, when verticality is made salient. Congruency effects should only occur when the instructional set activates relevant spatial information, not when it does not.

2.2. Methods

2.2.1. Participants

Participants were 48 (24 per group) students from Emory University, 12 males and 36 females, ages 17–34 years (M = 19). When asked at the end of the experiment, three participants in the unaware group had noticed that the words varied in verticality features, and so were replaced to maintain 24 participants in the group. The sample was 59% Caucasian, 27% Asian, 8% African American, and 6% Hispanic. Participants were native English speakers with normal or corrected vision. We obtained informed consent and treated the participants in accordance with the ethical standards of the American Psychological Association. All participants received course credit as compensation.

2.2.2. Design

The verticality awareness manipulation was implemented through random assignment to the two groups (aware vs. unaware). Orthogonally within groups, two variables were manipulated: word verticality (high vs. low) and response direction (up vs. down). On half of the 96 critical trials, word height and response direction were congruent, but on the other half they were incongruent. Half of the high vertical words were randomly assigned to the congruent condition, with the other half assigned to the incongruent condition. The low vertical words were analogously randomly assigned to congruent and incongruent conditions. Congruency (congruent vs. incongruent), font color (red vs. blue), and response direction (up vs. down) were counterbalanced between participants. Across four versions of the materials, each word appeared in four possible conditions: red font and up response; red and down; blue and up; blue and down. Equal proportions of high and low words occurred in each condition. Response time (RT) and errors were recorded for each trial.

2.2.3. Materials

The critical trials contained 96 randomly ordered concrete words whose literal meanings were associated with a spatial position (half high, half low). All critical words had been scaled previously for verticality (Casasanto, 2008). Ratings for high spatial words (5.27) and ratings for low spatial words (−3.61) were statistically different on a scale that ranged from 9 to −9, $t(47) = 79.61$, $SE = 0.111$, $p < .001$, $d_z = 11.49$ (Casasanto, 2008). The lists of high and low spatial words were controlled for number of letters, syllables, concreteness, and Kucera–Francis written frequency, not differing on these variables (Casasanto, 2008; based on data in the MRC Psycholinguistic Database: Machine Usable Dictionary, version 2.00[4]; Coltheart, 1981). See the Appendix for the words, including their average verticality ratings and standard errors.

2.2.4. Procedure

Participants indicated the font color of words presented in the center of a computer screen by pressing the corresponding color-key on a vertical keyboard positioned directly to the right of the screen (Fig. 1). The end of the keyboard with the “Q” key was positioned at the top, and the end with the “P” key at the bottom. As Fig. 1 illustrates, the keyboard was stabilized by strapping it to a wooden board apparatus.

Before participants in the aware group began the task, they were told about the verticality of the words: “The words you will see all have spatial qualities. For example, some will imply a low spatial location like ‘tunnel, culvert, dunk, grub, short’, while others will imply a higher location like ‘falcon, above, balcony, soar, top’. All of the words you will see come from a large online database of millions of sentences.” Participants in the unaware group were simply told, “All of the words you will see come from a large online database of millions of sentences.”

Each trial was self-paced and began with a centered black fixation cross. To initiate the trial, participants pressed and held down a white home key (“H” key), aligned with their horizontally-level eye fixation (adjusted for each participant). Once participants pressed the home key, a 500 ms pause occurred, followed by the target word in bold 28-point Arial red or blue font. When the word appeared, participants released the home key and pressed a color key to indicate the font color, counterbalanced as described earlier (e.g., the high key for red and the low key for blue). The “D” key was used for high responses, and the “L” key for low responses. Immediately after participants responded, the screen returned to fixation, participants returned to holding down the home key, and another trial began. Participants completed 40 practice trials before completing the 96 critical trials. The critical trial period contained two 30-s breaks.

There was one important exception to the color judgment task. At three random points, the word “emory” appeared, and participants had to release and re-press the home key, indicating that "emory" had occurred, regardless of its font color. These three catch trials, mixed in with the 96 critical trials, ensured that participants processed all target words semantically and did not just respond superficially to their font color. At the end of the experiment, participants in the unaware group were asked if they noticed anything about the words. If they mentioned that the words varied in spatial location or verticality, their data were discarded and replaced.

2.3. Results

2.3.1. Statistical method: Mixed effects analysis

Mixed effects regression modeling was used to analyze the data and to assess the a priori hypotheses (e.g., Baayen, Davidson, & Bates, 2008; Snijders & Bosker, 2012). Of primary interest was whether each group demonstrated a significant congruency effect. To test this, a priori contrasts were conducted within the mixed effects analysis using the following congruency effect calculation:

$$[(\text{high words \& down response} - \text{high words \& up response}) + (\text{low words \& up response} - \text{low words \& down response})]/2$$

or specified with respect to congruency:

$$[(\text{incongruent high words} - \text{congruent high words}) + (\text{incongruent low words} - \text{congruent low words})]/2$$

In addition to these a priori contrasts (which most directly tested our hypotheses), omnibus main effects and interactions were also examined. Specifically, the mixed effect model assessed fixed omnibus effects for group (aware vs. unaware), congruency (congruent vs. incongruent), verticality (high vs. low), group × congruency, group × verticality, congruency × verticality, and group × congruency × verticality. The mixed model also assessed random effects for participants and target words as intercepts to assess whether they, like the fixed effects, explained significant variance in response times. Assessing random effects for both participants and words further allowed us to generalize across both random factors, and to remove associated variance from the unexplained error. SPSS Mixed procedure (version 21; see Heck, Thomas, & Tabata, 2014) was used to perform these analyses, with restricted maximum likelihood estimation and an orthogonal variance component structure for the random effects. Significance of the random effects was assessed using the normal-deviate Wald test.

2.3.2. Response times

Reaction times that were three standard deviations above or below a participant's mean were removed. Initially we examined three different RT segments for congruency effects: complete RT from the word onset to the final color-key response press (RT1), word onset to the initial home key release (RT2), and home key release to the final color-key press (RT3). Our analyses focus only on the RT2 segment for the following reasons. For both the complete RT segment (RT1) and for the final color-key press segment (RT3), upward responses were faster on average compared to downward responses, regardless of congruency, indicating a motor bias for upward responses. This bias has been reported previously and may reflect biomechanical constraints that arise from human anatomy, independent of congruency effects (Brookshire et al., 2010; Santana & de Vega, 2011). Alternatively, upward bias may reflect salience asymmetry (Lakens, 2012; Proctor & Cho, 2006). Because "up" reflects the unmarked pole of the vertical axis, upward processing and movement are facilitated. Regardless of its source, this bias carries over into congruency effect calculations, biasing high words toward a congruency effect and low words toward a reverse congruency (or interference) effect. The RT2 segment when participants were engaged in response selection was the least affected by this bias in all three experiments. For this reason, we focused on RT2 in all analyses to follow. Table 1 presents the descriptive statistics for the RT2 response times in Experiment 1.

We also focused on RT2 because previous research on object processing and response selection has found that it tends to produce the strongest congruency effects (e.g., Jax &

Table 1
Estimated average and standard error of stimulus onset to initial release times (RT2) by experiment, group, and trial type

Experiment	Groups											
Expt 1	Aware				Unaware							
	U-L	U-H	D-H	D-L	U-L	U-H	D-H	D-L				
Mean	341	335	334	330	348	350	343	342				
SE	23	23	23	23	23	23	23	23				
Expt 2	Verticality				Concrete				Abstract			
	U-L	U-H	D-H	D-L	U-L	U-H	D-H	D-L	U-L	U-H	D-H	D-L
Mean	497	485	498	487	471	467	499	487	488	480	504	496
SE	12	12	12	12	12	12	12	12	12	12	12	12
	No-judgment control				Brookshire et al. replication							
	U-L	U-H	D-H	D-L	U-L	U-H	D-H	D-L				
Mean	454	452	454	454	347	349	350	345				
SE	12	12	12	12	12	12	12	12				
Expt 3	Brookshire et al. replication											
	U-L	U-H	D-H	D-L								
Mean	339	336	331	336								
SE	13	13	13	13								

Note. These are the estimated marginal means and standard errors from the mixed effects analysis for the four trial types (U-L, U-H, D-H, D-L) in each group (e.g., Aware, Unaware, etc.) for each experiment (1, 2, and 3). Expt is experiment. U-L is up responses to low verticality words. U-H is up responses to high verticality words. D-H is down responses to high verticality words. D-L is down responses to low verticality words. *SE* is standard error.

Buxbaum, 2010). This general finding suggests that the modality-specific activations underlying congruency effects have their largest impact on response selection, not on response execution.

Finally, we examined three specific congruency effects: the overall congruency effect just described in the congruency effect calculation, and two constituent congruency effects for high vs. low words, respectively. Critical a priori contrasts focused only on the overall congruency effect, because if motor bias influenced RT2, it would be averaged out. Each mixed model analysis, however, assessed whether the two constituent congruency effects were significant as well.

2.3.2.1. A priori directional contrasts: Based on previous research closely related to our paradigm, we predicted a priori that, if congruency effects occur, they would reflect facilitation, not interference (e.g., Brookshire et al., 2010; Casasanto, 2008; Dudschig et al., 2013; Lachmair et al., 2011; Thornton et al., 2013). Because we predicted that congruent trials would be faster than incongruent trials, we used one-tailed tests in critical contrasts within the mixed analysis, rather than two-tailed tests, which do not appropriately test directional predictions.

As Fig. 2 illustrates, the results supported our predictions. In the unaware group, the directional contrast for the overall congruency effect was not significant, with congruent trials being 0.35 ms slower than incongruent trials, $t(4175) = -0.12$, $SE = 2.87$, $p = .45$, $d_z = 0.003$. In a formal test of this null effect (Kass & Raftery, 1995; Rouder, Speckman, Sun, Morey, & Iverson, 2009), the JZS Bayes factor was 29.91 ($N = 2{,}304$, r scale = .5),

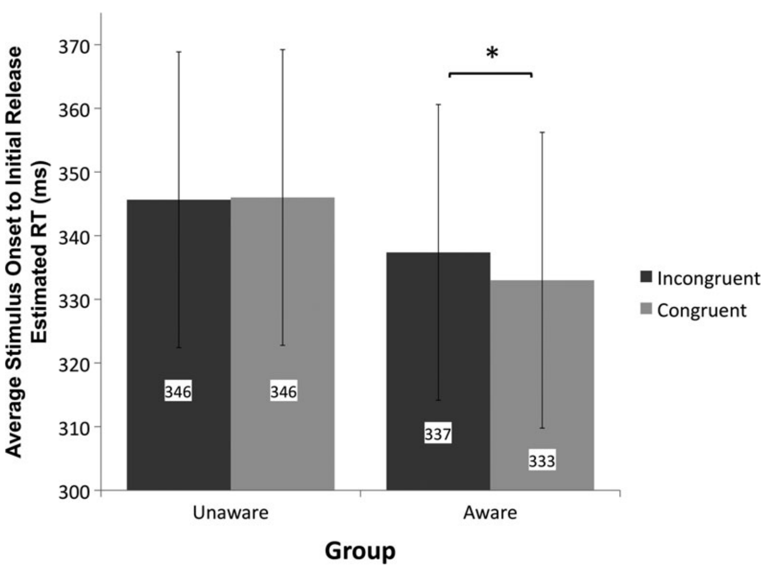

Fig. 2. Congruency effects in Experiment 1 for the unaware and aware groups (estimated marginal means from the mixed effects analysis on RT2). Error bars represent ± 1 *SE*. *Indicates $p = .05$.

providing substantial evidence for the null hypothesis (Jeffreys, 1961). In the aware group, however, the congruency effect fell right on the boundary of significance, with incongruent trials being 4.66 ms slower than congruent trials as predicted, $t(4162) = 1.64$, $SE = 2.84$, $p = .05$, $d_z = 0.05$).[5]

2.3.2.2. Fixed and random effects: Most omnibus fixed effects in the mixed model were not significant predictors of RT2, including group (aware vs. unaware), $F(1,46) = 0.11$, $p = .74$; congruency (incongruent vs. congruent), $F(1,4169) = 1.14$, $p = .29$; and verticality (high vs. low), $F(1,94) = 0.01$, $p = .94$. Although the omnibus interaction between group and congruency was not significant, $F(1,4169) = 1.54$, $p = .22$, the directional a priori contrasts just reported demonstrate that the difference between incongruent and congruent RTs bordered on significance for the aware group, but not for the unaware group. A significant interaction between congruency and verticality was observed, $F(1,4170) = 8.84$, $p = .003$, with low words exhibiting congruency and high words trending toward interference (see Table 1 for the condition means; low words, M difference = 8.16 ms, $t(4175) = 2.87$, $SE = 2.84$, $p = .002$, $d_z = 0.08$; high words, M difference = −3.85 ms, $t(4164) = -1.35$, $SE = 2.86$, p = ns [Correction added after initial online publication on September 22, 2014: p value was changed from ".09" to "ns."], $d_z = 0.04$). We address this pattern later in Experiment 2, along with implications for the polarity account of congruency effects, after observing a different set of patterns there. The group × congruency × verticality interaction was not significant, $F(1,4169) = 0.02$, $p = .88$.

Random effects for participants and words each explained significant variance (for participants, estimated variance = 12,838.42, $SE = 2687.28$, Wald $Z = 4.78$, $p < .001$; for words, estimated variance = 56.49, $SE = 22.59$, Wald $Z = 2.50$, $p = .01$).

2.3.3. Error rates

The average error rate was 4% for the aware group, and 2% for the unaware group (*SD*s = 4.37 and 2.65, respectively). These error rates did not differ significantly, $t(37.92) = 1.65$, $SE = 1.04$, $p = .11$, $d = 0.48$. All erroneous trials were removed when calculating congruency effects on RTs.

We also calculated "true errors," namely, errors only in response direction. Specifically, responses in the correct direction that landed on a key adjacent to the target key were coded as "correct" (e.g., pressing the K key when aiming for the L key). Computing errors in this manner decreased error rates (aware: $M = 1\%$, $SD = 1.82$; unaware: $M = 1\%$, $SD = 0.86$).[6]

2.3.3.1. "emory" error rate calculations: On "emory" trials, participants were instructed to always press the home key a second time, regardless of the font color. We examined the error rates for these trials to see if participants were actually reading the words. Both groups had error rates that were less than 5% (aware group, 4%; unaware group 1%). The error rates did not differ significantly between groups, $t(46) = 0.64$, $SE = 0.04$, $p = .53$, $d = 0.18$. These low error rates clearly indicate that participants were reading the words and not simply responding to font color.

2.4. Discussion

We assessed whether explicitly drawing attention to verticality modulates the accessibility of spatial information central to the meanings of spatial target words. If spatial information is accessed automatically in a context-independent manner, then congruency effects should have occurred in both the aware and unaware groups. If, however, spatial information is dynamic and context-dependent, then only the aware group should have exhibited a congruency effect.

The lack of a congruency effect in the unaware condition provides evidence against the hypothesis that spatial information is activated automatically. Although, drawing conclusions from a null result is always questionable, the associated Bayes factor indicates that this finding provides strong evidence for the null hypothesis. As we shall see in later experiments, we replicated this null result three more times, under conditions where we also predicted that null results would occur. Thus, this first predicted null result of four suggests that spatial features are indeed not activated automatically in grounded congruency paradigms.

Participants in the unaware condition were 99% accurate at detecting when the word "emory" was presented, indicating that they were clearly reading the words. Thus, the lack of a congruency effect for the unaware group did not occur because participants only processed font color. Instead, participants processed the words, just not in a manner that activated spatial features. On reading the words, spatial features did not become active automatically. Discrepancies with previous research (Brookshire et al., 2010; Casasanto, 2008) will be addressed in the General Discussion.

The congruency effect that bordered on significance at $p = .05$ in the aware condition is consistent with our general hypothesis that the activation of spatial features is dynamic and context-dependent. When task instructions made spatial features salient in the experimental context, a congruency effect appeared on the threshold of significance.

Perhaps the instructional set used in the aware condition did not make a particularly strong or lasting impression. As a result, the accessibility of spatial features across trials may have only increased slightly. If so, then a manipulation that makes spatial features salient on each trial may make them more accessible, such that congruency effects become more robust. Experiment 2 assessed this possibility, explicitly drawing attention to verticality on each trial, rather than only in the initial instructions.

3. Experiment 2

3.1. Overview

In this next experiment, we attempted to bring attention to verticality on each trial by asking participants to judge a semantic property of the target word before making a color judgment. On each trial, a word first appeared in black font before appearing in color.

While the word was in black, participants in three different groups performed one of three semantic judgments that drew their attention to the word's meaning. The word then changed to red or blue font, and participants made an up or down response, as in Experiment 1.

If congruency effects emerge from context-dependent processing that activates semantic properties (especially verticality), then these initial orienting tasks could potentially activate this information and produce these effects. We used a decreasingly explicit series of semantic orienting tasks to see how explicit these tasks must be to activate verticality. In the verticality group, participants initially judged whether a word's meaning was associated with a salient vertical position or not. In the concrete group, participants judged whether a word's meaning was concrete or not. In the abstract group, participants judged whether a word's meaning was abstract or not.

We predicted that the verticality judgment task would activate vertical features in concepts such as *sky* and *basement*, thereby producing a congruency effect. Also of interest was whether the more indirect judgments of concreteness and abstractness would also activate vertical features. If vertical features are central in the meanings of concepts like *sky* and *basement*, then any orienting task that produces deep semantic access should be likely to activate them. Thus, we predicted that judging the concreteness or abstractness of these concepts would be sufficient to activate their central height features, thereby producing congruency effects.

Including the concreteness and abstractness judgment groups allowed us to further assess whether initial responses to the orienting task interfered with subsequent color responses. As described in the Methods, the abstractness and concrete tasks were essentially inverses of each other in terms of response execution. Whereas the concreteness task required an initial key press to the critical targets before the color response, the abstractness task did not require an initial response. Manipulating the polarity of the same general task allowed us to assess whether the initial response affected congruency effects.

In the no-judgment control group, participants did not make a semantic orienting judgment while the word was black, but instead simply waited for the word to become red or blue, before making a color judgment. Similar to Experiment 1, this group provided an opportunity to assess whether congruency effects are automatic in a context-independent manner, given that no orienting judgments occurred.

Finally, a fifth group replicated a previously tested trial format (Brookshire et al., 2010). The words never appeared in black for this control group, and participants did not make a semantic orienting judgment. Instead, the words immediately appeared in red or blue, and participants only made a color judgment. Because Brookshire et al. observed congruency effects that appeared to occur automatically (i.e., with no assistance from instructions or orienting tasks), we wanted to assess whether replicating their trial format would again produce congruency effects. Although Brookshire et al. only used abstract words associated with metaphorical spatial meanings, we again used concrete words, given our theoretical assumption that verticality features should be most accessible for concrete words (i.e., verticality is most basically a concrete feature).

3.2. Methods

3.2.1. Participants

Excluding the verticality group, who knew about word verticality, seven participants in other groups revealed at the end of the experiment that they noticed the words varied in verticality (two in the abstract group, two in the concrete group, and three in the Brookshire et al. replication group). These participants were replaced, yielding a total of 105 participants (21 per group). All participants were students from Emory University, 28 males and 77 females, ages 18–24, (M = 19). The sample was 56% Caucasian, 23% Asian, 13% African American, 4% Hispanic, and 4% other. We used the same selection criteria as Experiment 1, and again obtained informed consent and treated participants in accordance with the ethical standards of the American Psychological Association. All participants received course credit as compensation.

3.2.2. Design

We used the same basic design as Experiment 1, but with five groups of participants instead of two: (a) verticality judgment, (b) concreteness judgment, (c) abstractness judgment, (d) no-judgment control, and (e) Brookshire et al. replication.

3.2.3. Materials

We used the same 96 spatial words as Experiment 1 but excluded the "emory" trials, given that they were no longer relevant in the context of orienting tasks. To implement the orienting tasks, the word set also included eight closed-class non-spatial words each repeated 3 times (24 presentations total), randomly intermixed with the 96 spatial words. The Appendix includes the eight closed-class words. These additional trials provided negative response items for the verticality and concreteness judgments, and positive response items for the abstractness judgments. To simplify counter-balancing constraints, the eight closed-class words were repeated three times instead of using 24 unique words. In the Brookshire et al. replication group, only the 96 spatial words were included. When computing congruency effects later, the closed-class words were excluded.

3.2.4. Procedure

The procedure was identical to Experiment 1, except: (a) the up and down response keys were located on "A" and apostrophe keys, in accordance with Brookshire et al. (2010); (b) participants received warning feedback if they released the home key faster than 200 ms or slower than 1,000 ms once the word changed from black to blue or red; (c) the instructions and trial sequences differed, as described next.

In the verticality, concreteness, and abstractness groups, participants pressed and held down the home key to initiate a trial, with a black fixation cross appearing in the screen center for 1,000–1,500 ms (the randomized duration prevented precise anticipation of the stimulus, thereby increasing attention to it). Following fixation, the word appeared for 2,000 ms, first in black font, cueing participants to make the initial word property judgment (verticality, concreteness, or abstractness, depending on the group). If the word had the designated property, participants briefly released the home key and then pressed it back down; if it did not, they continued to hold the home key down. In the verticality and concreteness conditions, participants briefly released the home key for all high and low words but not for the closed-class non-spatial words (80% of the trials). In the abstractness condition, participants briefly released the key for the closed-class words (20% of the trials). Following the initial 2000 ms judgment period, the word changed to either red or blue font, cuing participants to release the home key and press the corresponding color key.

In the no-judgment control, the procedure was identical except that participants did not make any additional judgment while the word appeared in black for 2,000 ms; they simply responded to the word color when it appeared. In the Brookshire et al. replication, the word never appeared in black, with the remainder of the procedure being identical.

3.3. Results

3.3.1. Response times

Reaction times that were three standard deviations above or below a participant's mean were removed. The same mixed effects analysis and a priori directional contrasts from Experiment 1 were conducted on the RT2 segments, except that the contrasts and fixed main effect of group included five groups, not two. Table 1 presents the descriptive statistics for RT2 in each group.

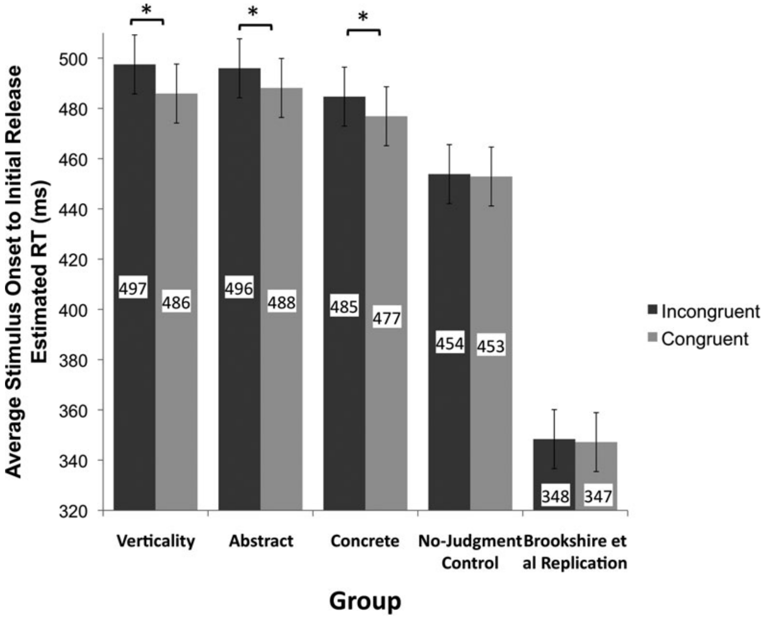

Fig. 3. Congruency effects in Experiment 2 for the verticality judgment, abstractness judgment, concreteness judgment, no-judgment control, and Brookshire et al. replication groups (estimated marginal means from the mixed effects analysis on RT2). Error bars represent ± 1 *SE*. *Indicates $p < .05$.

3.3.1.1. A priori directional contrasts: We hypothesized originally that if congruency effects in this paradigm require context-dependent activation of spatial features, then we would only see congruency effects when semantic-orienting tasks activate verticality in a context-dependent manner. Specifically, significant congruency effects should occur, at the very least, in the verticality group and perhaps also in the concreteness and abstractness groups, but not in the no-judgment control group, nor in the Brookshire et al. replication.

As Fig. 3 illustrates, the results supported our predictions. Participants produced significant overall congruency effects in the verticality, concreteness, and abstractness groups, but not in the no-judgment control or Brookshire et al. replication groups (verticality, *M* difference = 11.59 ms, $t(9188) = 2.85$, $SE = 4.06$, $p = .002$, $d_z = 0.09$; concreteness, *M* difference = 7.80 ms, $t(9151) = 1.93$, $SE = 4.05$, $p = .03$, $d_z = 0.06$; abstractness, *M* difference = 7.83 ms, $t(9200) = 1.92$, $SE = 4.07$, $p = .03$, $d_z = 0.06$; no-judgment control, *M* difference = 0.95 ms, $t(9184) = 0.21$, $SE = 4.53$, $p = .42$, $d_z = 0.01$; Brookshire et al. replication, *M* difference = 1.19 ms, $t(9216) = 0.29$, $SE = 4.14$, $p = .39$, $d_z = 0.01$). As in Experimental 1, all contrasts used one-tailed tests, given that response facilitation was predicted. In further assessments of the null results, the JZS Bayes factors in the no-judgment control and Brookshire et al. replication conditions were 27.58 ($N = 2{,}016$, r scale = .5) and 27.04 ($N = 2{,}016$, r scale = .5), respectively, offering strong odds in favor of the null hypothesis.

3.3.1.2. Fixed and random effects: When the verticality of word meaning was taken into account, the mixed effects model revealed different patterns of congruency across the different orienting tasks, as reflected in a significant three-way omnibus interaction between group × congruency × verticality, $F(4,9191) = 7.27$, $p < .001$. Table 1 presents the relevant condition means. For the verticality judgment group, both high and low words exhibited congruency (high words, *M* difference = 12.98 ms, $t(9215) = 2.26$, $SE = 5.73$, $p = .01$, $d_z = 0.07$; low words, *M* difference = 10.19 ms, $t(9153) = 1.78$, $SE = 5.74$, $p = .04$, $d_z = 0.06$). For the concreteness group, high words exhibited congruency, whereas low words exhibited interference (high words, *M* difference = 32.12 ms, $t(9152) = 5.60$, $SE = 5.73$, $p < .001$, $d_z = 0.18$; low words, *M* difference = −16.51 ms, $t(9151) = -2.88$, $SE = 5.73$, $p <$ ns [Correction added after initial online publication on September 22, 2014: p value was changed from ".001" to "ns."], $d_z = 0.09$). For the abstractness group, only high words exhibited congruency (high words, *M* difference = 24.08 ms, $t(9180) = 4.20$, $SE = 5.74$, $p < .001$, $d_z = 0.13$; low words, *M* difference = −8.41 ms, $t(9217) = -1.46$, $SE = 5.77$, $p =$ ns [Correction added after initial online publication on September 22, 2014: p value was changed from ".15" to "ns."], $d_z = 0.05$). Neither high or low words exhibited congruency for the control group, nor for the Brookshire et al. replication group (control group high words, *M* difference = 1.57 ms, $t(9175) = 0.24$, $SE = 6.46$, $p = .40$, $d_z = 0.01$; low words, *M* difference = 0.33 ms, $t(9209) = 0.05$, $SE = 6.40$, $p = .48$, $d_z = 0.002$; Brookshire et al. high words, *M* difference = 1.19 ms, $t(9240) = 0.20$, $SE = 5.89$, $p = .42$, $d_z = 0.006$; low words, *M* difference = 1.18 ms, $t(9157) = 0.20$, $SE = 5.83$, $p = .42$, $d_z = 0.006$).

For the omnibus main effects, group significantly predicted RT2, $F(4,100) = 28.07$, $p < .001$; the faster RTs in the control and Brookshire et al. replication conditions probably reflected absence of an initial judgment. The omnibus effect of congruency was also significant, $F(1,9196) = 9.89$, $p = .002$, with incongruent RTs generally being slower than congruent RTs across groups. The omnibus interaction for congruency with verticality was also significant $F(1,9201) = 20.70$, $p < .001$, with high words showing an overall congruency effect across groups, and low words not. No other omnibus main effects or interactions were significant (verticality, $F(1,95) = 0.18$, $p = .67$; group × congruency, $F(4,9187) = 1.22$, $p = .30$). The random effects for participant and word intercepts each accounted for significant variance (for participants, estimated variance = 2,692.01, $SE = 393.83$, Wald $Z = 6.84$, $p < .001$; for words, estimated variance = 124.42, $SE = 30.19$, Wald $Z = 4.12$, $p < .001$, respectively).

3.3.2. Error rates

In every group, the average error rates were 5% or less (verticality, $M = 2\%$, $SD = 1.80$; concreteness, $M = 3\%$, $SD = 2.72$; abstractness, $M = 2\%$, $SD = 1.84$; no-judgment control, $M = 2\%$, $SD = 2.48$; Brookshire et al. replication, $M = 5\%$, $SD = 3.53$). The Brookshire et al. replication group had a significantly higher percentage of errors than the other groups, $F(4,100) = 5.91$, $p < .001$, $n_p^2 = 0.19$ (Tukey's HSD comparisons all $p < .05$). However, "true error rates" (defined earlier) did not differ significantly across groups, $F(4,100) = 0.25$, $p = .907$, $n_p^2 = 0.01$ (verticality, $M = 1\%$, $SD = 1.13$; concreteness, $M = 1\%$, $SD = 1.15$; abstractness, $M = 1\%$, $SD = 0.74$; no-judgment control, $M = 1\%$, $SD = 1.43$; Brookshire et al. replication, $M = 1\%$, $SD = 1.32$). All erroneous trials were again removed when computing congruency effects on RTs.

3.4. Discussion

We failed to obtain a congruency effect in two conditions: the no-judgment control and the Brookshire et al. replication. Again, we predicted these null effects, as we did for the related null effect in Experiment 1. In all three cases, our theoretical position did not predict a congruency effect, because task conditions did not draw attention to spatial features, nor activate them in some other manner. We will see a fourth null effect in Experiment 3, again predicted.

Because significant congruency effects occurred following all three semantic orienting tasks, any kind of deep processing appears sufficient for making the spatial features central to the meanings of the target words accessible, thereby producing congruency effects. We address implications of this broad effect in the General Discussion. More specifically, however, the benefits of an orienting task were strongest following verticality judgments, with both high and low words exhibiting patterns associated with congruency. In contrast, concreteness judgments only exhibited congruency for high words (with low words exhibiting interference), whereas abstractness judgments only exhibited congruency for low words (as in Experiment 1). No account that we know of explains this complex set of patterns for high and low words, including polarity correspondence (Lakens, 2012).

4. Experiment 3

4.1. Overview

Perhaps we did not observe congruency effects in the no-judgment control and Brookshire et al. replication groups of Experiment 2 because participants only processed the surface characteristics of the words (in particular, their font color). Participants may not have actually read the words because the task did not necessitate it. Although the "emory" trials in Experiment 1 indicated that participants were reading the words, Experiment 2 did not include these trials. To rule out this alternative hypothesis, we repeated the Brookshire et al. replication condition in Experiment 3, but this time required participants to take a surprise "old/new" memory test on the words.

If participants do not read the words when processing them in the color decision task, then they will not remember them reliably later on the memory test. If, however, participants do read the words, then they will remember them later above chance. Additionally, if participants remember the words, but do not show a congruency effect, then this will suggest that the null results in Experiment 2 did not occur because participants failed to read the words. Instead, this pattern would again indicate that spatial features only become active with sufficiently deep processing.

4.2. Methods

4.2.1. Participants

Participants were 17 students from Emory University, 7 males and 10 females, ages 18–30 years (M = 23.12). The sample was 35% Asian, 35% Caucasian, 12% African American, 12% other, and 6% Hispanic. We used the same selection criteria as Experiments 1 and 2, obtained informed consent, and treated the participants in accordance with the ethical standards of the American Psychological Association.

4.2.2. Design and materials

We used the same design as the Brookshire et al. replication group in Experiment 2, with the following adjustments. The 96 spatial words were divided into three groups of 32, with each group containing half high words and half low words. The three groups did not differ significantly on the number of letters, syllables, concreteness, or Kucera–Francis written frequencies of the words. Two of these word groups appeared during an initial testing phase, as participants completed the standard color judgment task without an orienting task. In a second testing phase, participants again made color judgments, but this time, one of the word groups from the first phase was mixed randomly with words from the third word group (not seen during the first testing phase). Additionally, in the second testing phase, participants also indicated whether each target word had or had not been presented in the first phase, after indicating the word's color. Thus, the second phase included a forced choice recognition test on half the words presented in Phase 1. If participants processed these words in Phase 1, then they should demonstrate memory for them in Phase 2. Three different versions of the materials were constructed that rotated the three word groups through Phases 1 and 2, such that each word group appeared equally often as targets and foils on the recognition test. Each participant received a random order of words in Phases 1 and 2.

Because the purpose of this experiment was to rule out shallow surface processing, we did not replicate all of the counter-balanced conditions from Experiment 2 in full. Instead, we randomly chose one condition for use in Experiment 1: up responses to red words, and down responses to blue words. Half the high words and half the low words were randomly assigned to the red font color.

4.2.3. Procedure

Participants completed two testing phases. Phase 1 was identical to the Brookshire et al. replication group in Experiment 2, except that participants only saw 64 of the 96 words. In Phase 2, 32 words were repeated from Phase 1, and 32 words were new. Participants performed the red vs. blue font color judgment task in both phases. In Phase 2, however, they also indicated whether the target was old (press the ENTER key) or new (press the SPACE bar), after making a color response.

4.3. Results and discussion

4.3.1. Congruency RTs and recognition d'

We hypothesized that if participants read the words, then they would remember them later. Additionally, we hypothesized that if participants did not process the words deeply enough to activate spatial features, then they would not show congruency effects. These

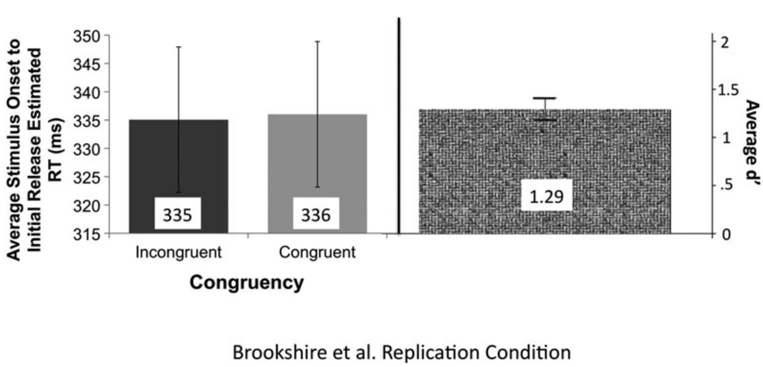

Fig. 4. The left panel displays the congruency effect in Experiment 3 (estimated marginal means from the mixed effects analysis on RT2 during Phase1). The right panel displays the average *d*' value for forced-choice recognition test during Phase 2. Error bars in both panels represent ± 1 *SE*.

two hypotheses combined—memory for the words, but no congruency effects—would be evidence that spatial features only become active when these words receive sufficiently deep processing.

Table 1 presents descriptive statistics from a mixed analysis of RT2 during Phase 1 for the color judgments. Reaction times that were three standard deviations above or below a participant's mean were removed. Fig. 4 presents the key findings for congruency in Phase 1 and for memory in Phase 2. Consistent with Experiment 2's results, the Brookshire et al. replication group again did not display a significant congruency effect, M difference = −0.96 ms, $t(1032) = -0.29$, $SE = 3.35$, $p = .39$, $d_z = 0.01$, one-tailed. Without an orienting task that made spatial features accessible, a congruency effect did not occur. Across three experiments, this is the fourth time that we predicted and obtained a null effect under these conditions. Furthermore, the Bayes factor was 19.89, indicating strong odds in favor of the null hypothesis (N = 1,088; r scale = .5).

Participants did, however, remember the words well above a chance level, $d' = 1.29$, $t(16) = 10.38$, $SE = 0.12$, $p < .001$, $d_z = 2.52$. The hit rate was 62%, and the false alarm rate was 18% (SDs = 8.73 and 7.85, respectively). These results demonstrate that participants clearly processed the target words and were not simply responding to font color.

None of the fixed effects in the RT analysis were significant (congruency: $F(1,1032) = 0.08$, $p = .78$; verticality: $F(1,1032) = 1.29$, $p = .26$; congruency × verticality: $F(1,1032) = 1.03$, $p = .31$). The random effect of participant intercepts was significant (estimated variance = 2,713.65, SE = 976.14, Wald $Z = 2.78$, $p = .01$); the random effect of word intercepts was redundant.

This pattern of results supports our hypothesis that participants do indeed read the target words during the color judgment task but do not access spatial features for the words automatically. Instead, spatial features only become active and produce congruency effects when some factor in the task context makes the relevant spatial features accessible.

4.3.2. Error rates

Error rate was calculated for Phase 1 in the same way as in Experiments 1 and 2, and again was very low (M = 2%, SD = 3.35). The true error rate was even lower (M = 1%, SD = 1.25). All erroneous trials were removed when computing congruency effects on RTs.

5. General discussion

5.1. Summary of results

Based on the broad lack of evidence for automaticity across diverse literatures, we predicted that grounded congruency effects would not occur invariantly across task conditions. Instead, congruency effects should only occur when task conditions make relevant features salient. Findings from three experiments supported this hypothesis.

In Experiment 1, a congruency effect failed to emerge when initial task instructions did not make participants aware of spatial features. Conversely, when initial task instructions made participants aware of spatial features, a congruency effect emerged that bordered on significance. If the semantic access of spatial features had been automatic and context-independent, a congruency effect should have occurred regardless of whether participants were told explicitly about spatial features or not. The failure for congruency effects to occur across *both* conditions indicates that spatial features did not become active automatically.

In Experiment 2, orienting tasks drew attention to spatial features on each trial, attempting to make them more salient than did the initial instructions manipulation in Experiment 1. Under these conditions, congruency effects emerged, regardless of whether the orienting task focused participants' attention on verticality, concreteness, or abstractness. All three orienting tasks appeared sufficiently deep to access spatial features central to the target words' meanings. Conversely, two conditions that did not use orienting tasks to make spatial features salient again did not exhibit a congruency effect. If the semantic access of spatial features had been automatic, congruency effects should have occurred regardless of whether an orienting task was performed prior to making a color response.

In Experiment 3, when no orienting task was used, congruency effects again did not occur, further demonstrating the necessity of context for activating spatial features. Notably, however, participants remembered the words on a surprise forced-choice memory test later, indicating that they had read and processed them. In this task context, reading the words did not activate their spatial features, such that congruency effects with motor responses could occur.

Across experiments, a congruency effect failed to occur a total of four times, in four different groups of participants. Although null results should be interpreted with caution, we predicted each absence of a congruency effect (with all of them exhibiting high JZS Bayes factors, consistent with true null results). Whenever spatial features were not made salient, a congruency effect did not occur. Congruency effects only occurred when spatial features became salient.

5.2. Reconciling conflicting results

Our four null effects contrast with congruency effects obtained under similar task conditions (Brookshire et al., 2010; Casasanto, 2008). Any number of minor design differences could have increased the accessibility of spatial features dynamically in these previous studies. First, these researchers positioned their response apparatus directly below the computer screen. If this arrangement caused participants to constantly glance up and down from the computer screen to the response apparatus, eye movements in the vertical dimension may have reinforced the salience of verticality in these studies enough to elicit congruency effects. Indeed, recent work has demonstrated congruency effects with eye movements (Dudschig et al., 2013). In contrast, our response apparatus was positioned directly to the right of the computer screen to ensure that the response buttons lay above and below the target words at eye level on the computer screen. If this caused participants to glance from left to right, their eye movements may not have made verticality salient, and might have actually masked it. As a result, congruency effects may only occur when instructions or an orienting task activates verticality, consistent with the results observed here.

Second, previous studies used abstract words associated with metaphorical spatial meaning, whereas we used concrete words associated with literal spatial meaning. Perhaps this difference in stimuli led to the discrepancy in congruency effects, although there is evidence that concrete words exhibit a more robust effect than abstract words in a similar verticality congruency paradigm (Bergen, Lindsay, Matlock, & Narayanan, 2007).

Third, other unknown cues in previous experiments may have inadvertently made participants aware that verticality was manipulated in the stimulus materials, thereby influencing their response times enough to cause a congruency effect. Previous studies do not report whether participants noticed anything about the words, especially about the verticality of their semantic content.

Two other recent articles reported that grounded congruency effects were obtained across diverse task conditions, concluding that they reflect automatic processing (Dudschig et al., 2013; Lachmair et al., 2011). Again, however, for processing to occur automatically, it must occur generally across task conditions. Because we consistently failed to find congruency effects repeatedly under certain well-defined conditions here, it follows that congruency effects are not automatic. If further follows that these researchers may have used task conditions that made spatial features salient across all the task conditions they examined.

In Dudschig et al. (2013), for example, their use of vertically oriented eye movements, coupled with a lack of nonspatial fillers, may have made verticality salient, such that a congruency effect emerged. Under other task conditions, their congruency effect might well be strengthened or weakened, especially when one considers the dynamic character of congruency effects across all the literatures we review.

Similarly, in Lachmair et al.'s (2011) Experiment 3, the majority of their participants were aware that verticality was being manipulated, suggesting that verticality may have often become salient implicitly. Also, Lachmair et al. positioned their response box below the screen, again encouraging up-down eye movements, similar to Casasanto (2008) and Brookshire et al. (2010).

In summary, our predicted failure to obtain congruency effects four times suggests that the congruency effects in these other articles may not be as automatic as claimed, and that they may disappear under other processing conditions. An important goal for future research is to replicate and explain these conflicting results within a single theoretical framework. And again, it is essential to integrate these findings with those in the other literatures reviewed earlier that similarly concern themselves with automaticity vs. context-dependent processing.

5.3. Revisiting conceptual cores

Consistent with the literature reviewed earlier, our findings suggest that conceptual cores do not exist in word meanings. Because spatial features are central for words whose meanings depend heavily on spatial position (e.g., *sky*, *basement*), they are likely candidates for core features. As our results illustrate, however, the spatial features of these words are dynamic and context-dependent, with their availability varying across task contexts. Many findings, across literatures, now demonstrate clearly that features potentially viewed as core are actually context-dependent.

Together all of these results question the gold-standard status that researchers often place on automatically activated semantic information. Instead, these results suggest that context-dependent activation of semantic information is the norm, not some peripheral, irrelevant, uninteresting phenomenon, and that automatic activation is an idealized phenomenon that may never occur (also see Gawronski & Cesario, 2013; Kiefer et al., 2012).

One potential counter-argument is that conceptual cores are automatically activated initially but are then suppressed rapidly for some reason. Certainly, this is possible, but it is a significant retreat from the typical view that conceptual cores become active immediately to provide fundamental information for conceptual processing. If conceptual cores do not play this role and just disappear, then what use are they? Another potential counter-argument is that cores only become active when a concept is processed deeply, with our task conditions not providing sufficiently deep processing to activate them. Again, this is possible but constitutes a significant retreat from the view that cores play central roles in processing (and is essentially the cores-last view presented earlier). Also, if our orienting tasks were not deep enough to activate cores, then what tasks would be?

5.4. Revisiting modal versus amodal features

Sometimes, grounded congruency effects do not occur in an experiment, even when predicted (e.g., Besner & Stolz, 1999; Borghi et al., 2004; Bub et al., 2008; Fias et al., 2001; Rotteveel & Phaf, 2004; Sato et al., 2008; van Elk & Blanke, 2011). As a result, researchers sometimes conclude that the absence of a grounded congruency effect implies that word meanings are not grounded, or at least do not have grounded conceptual cores (cf. Dove, 2009; Machery, 2007; Mahon & Caramazza, 2008).

Importantly, however, the absence of a grounded congruency effect does not imply that grounded information does not exist for a word's meaning. Nor does it necessitate that word meanings must have amodal cores instead, as these theorists conclude. Instead, the absence of a grounded congruency effect most likely demonstrates that grounded features associated with a word's meaning are simply not active in the current task context. Because all features, grounded or otherwise, are dynamic and context dependent, they are not always active. When grounded features are relevant to the task at hand, they may be activated quickly and produce congruency effects. Other times, when grounded features are irrelevant, they may not be activated.

Grounded features may not become active for a variety of reasons. Under some task conditions, processing may be relatively shallow, such that only linguistic forms become active (Barsalou et al., 2008; Glaser, 1992; Louwerse & Connell, 2011). Alternatively, amodal representations could possibly become active instead, and suffice for processing under these task conditions. Still another possibility is that grounded representations become active, just not the ones predicted to become active. In the control and replication conditions of our experiments, for example, participants may have simulated objects without their verticality being represented or salient (e.g., as attentional orienting responses).

As Barsalou (1999) noted, simulations may often exclude much potentially relevant situational information, only including a few aspects of experience relevant to current task conditions. If spatial information was excluded in some conditions here, then the simulations constructed would not be expected to interact with responses and produce verticality-based congruency effects.

Furthermore, the absence of a congruency effect does not imply that some other kind of (possibly amodal) information constitutes the word's conceptual core (Dove, 2009; Machery, 2007; Mahon & Caramazza, 2008). Most likely, the word has no core. Based on the literature review earlier and the findings reported here, the accessibility of all information in a word's meaning varies dynamically as a function of context. To the extent that amodal features play some role in a word's meaning, their accessibility, too, is likely to be dynamic and context-dependent. To argue that amodal cores exist for concepts, one must first provide positive direct evidence that amodal features exist (as opposed to arguing negatively that they exist because grounded features sometimes do not become active). One must then provide additional evidence that these amodal features are active automatically across contexts. As much evidence indicates, this appears increasingly implausible.

5.5. Issues related to the composition and time course of semantic processing

Much remains to be understood and established about the dynamic activation of conceptual information. Future research needs to specify how situational context and task demands control the context-dependent construction of semantic representations (Wilson-Mendenhall, Barrett, Simmons, & Barsalou, 2011). One issue, implicated by the results of Experiment 2, concerns the interaction of context and accessibility. In Experiment 2, all three deep semantic-orienting tasks activated spatial features (verticality, concreteness, abstractness). A tentative hypothesis that follows is that any sufficiently deep orienting task will tend to activate highly accessible features, such as the spatial features for the words used here. A related hypothesis is that the less accessible features of a concept will require much more specific orienting tasks to become active. Consider Barsalou's (1982) discussion of activating the feature *floats* for basketballs. Because *floats* is a relatively inaccessible feature of basketballs (relative to a feature like *round* or *bounces*), it probably requires a very specific context like "was thrown into the water" to become active. Specific deep processing may mostly be necessary when a feature is not consistently mapped to a concept, such as *floats* for basketballs (Shiffrin & Schneider, 1977). Conversely, general deep processing may be sufficient to activate features that are frequently and consistently mapped to a concept, such as *round* for basketballs.

A final issue concerns the time course of activation for a word's meaning over the course of comprehension. From the perspective that automatic conceptual cores constitute the gold standard, researchers often seem to assume that fast early information plays the most important central role when comprehending a word (core-first theories). As reviewed earlier for core-last theories, however, important conceptual information can become active more slowly than superficial information. As comprehension evolves, important conceptual information could become active at many later points in the process. Why should central information only become active initially? From the perspective of comprehension, central information could be added whenever it happens to become relevant. Just because information is added late does not mean that it is unimportant.

The distinction between shallow vs. deep comprehension underlines the potential importance of semantic information added late in the comprehension process. When reading a text initially, only shallow comprehension and partial integration of the text may occur; for example, processing only occurs to ascertain how the text fits together superficially (the Minimalist account; McKoon & Ratcliff, 1992). Conversely, much deeper and more accurate processing of a text's meaning may develop as readers go beyond initial superficial processing (the Constructivist account; Graesser, Singer, & Trabasso, 1994). As readers spend more time processing a text, they are increasingly likely to establish elaborative and integrative inferences that are fundamentally important to understanding it accurately and deeply (e.g., related to a goal, theme, or background knowledge). It is unlikely that this additional processing is less important than superficial processing that occurs initially.

In general, we have much to learn about how the information activated early vs. late for a word's meaning enters into the time course of comprehension. As we explore this important issue, we should be careful not to impulsively confer gold-standard status on semantic information accessed initially. As we have seen, central features in a word's meaning are not activated automatically in a context-independent manner, but instead vary dynamically as a function of context. Furthermore, it is highly likely that important conceptual information becomes active late during comprehension. Finally, we should avoid making overly quick assumptions about whether information accessed early or late is grounded, linguistic, amodal, etc. Careful theorizing that takes the entire comprehension process into account will be essential for understanding the representation of word meaning in context, together with empirical investigation that assesses these accounts.

Acknowledgments

We thank Kevin Holmes, David Kemmerer, Daniel Lakens, Esther Papies, and Charles Folk for enlightening discussions of congruency effects; Taryn Colton, Hailey Friedman, Rob St. Louis, Moses Tincher, Bridget Warren, and Irene Vitale for assistance with data collection; Nancy Bliwise and Owen Devine for assistance with mixed-effect regression modeling; Julio Santiago, Daniel Lakens, Dermot Lynott, and Zachary Estes for reviews that significantly strengthened the work and the manuscript.

Notes

1. For words with multiple meanings, theories sometimes assume that each meaning has its own conceptual core. All arguments we have made so far apply to each of these meanings, namely, each could have a core that is activated early or late.

2. Based on this definition, SNARC, perceptual priming, and Simon effects are also grounded congruency effects. We present them separately, however, because they were not initially developed to test grounded cognition. In this section, we focus on congruency effects that aim to establish the role of simulation in conceptual processing.
3. We do not assume that linguistic forms are the same as amodal symbols, but assume that amodal symbols constitute a form of conceptual representation to which linguistic forms can refer. See Barsalou et al. (2008) for further discussion.
4. http://www.psy.uwa.edu.au/MRCDataBase/uwa_mrc.htm
5. All effect size calculations were made with Lakens's (2013) open-source spreadsheet.
6. These trials were also not included in the calculation of congruency effects on RTs. Although these responses were in the correct direction, the keys pressed were at varying distances from the home key, such that including these responses could result in RTs that differed simply because of the distance required to travel with one's finger, irrespective of congruency. Additionally, we do not report error rates by condition (e.g., congruent high words, congruent low words, etc.). Because error rates across groups in all three experiments were so low (5% or less; 1% for true errors), comparisons between conditions produce little contrast.

References

Baayen, R. H., Davidson, D. J., & Bates, D. M. (2008). Mixed-effects modeling with crossed random effects for subjects and items. *Journal of Memory and Language*, *59*, 390–412.

Bargh, J. A. (1997). The automaticity of everyday life. In R. S. Wyer, Jr (Ed.), *The automaticity of everyday life: Advances in social cognition* (Vol. 10, pp. 1–61). Mahwah, NJ: Erlbaum.

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

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

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

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

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

Barsalou, L. W. (2008). Grounded cognition. *Psychology*, *59*(1), 617.

Barsalou, L. W., Santos, A., Simmons, W. K., & Wilson, C. D. (2008). Language and simulation in conceptual processing. In M. De Vega, A. M. Glenberg & A. C. Graesser (Eds.). *Symbols, embodiment, and meaning* (pp. 245–283). Oxford, UK: Oxford University Press.

Barsalou, L. W., & Wilson, C. D., & Hasenkamp. (2010). On the vices of nominalization and the virtues of contextualizing. In B. Mesquita, L. Feldman Barrett & E. Smith (Eds.), *The mind in context* (pp. 334–360). New York: Guilford Press.

Bergen, B. K., Lindsay, S., Matlock, T., & Narayanan, S. (2007). Spatial and linguistic aspects of visual imagery in sentence comprehension. *Cognitive Science*, *31*, 733–764.

Besner, D., & Stolz, J. A. (1999). What kind of attention modulates the Stroop effect? *Psychonomic Bulletin & Review*, *6*(1), 99–104.

Besner, D., Stolz, J. A., & Boutilier, C. (1997). The Stroop effect and the myth of automaticity. *Psychonomic Bulletin & Review*, *4*(2), 221–225.

Blanchette, I., & Dunbar, K. (2000). How analogies are generated: The roles of structural and superficial similarity. *Memory & Cognition*, *28*(1), 108–124.

Borghi, A. M., Glenberg, A. M., & Kaschak, M. P. (2004). Putting words in perspective. *Memory & Cognition*, *32*(6), 863–873.

Borgmann, K. W. U., Risko, E. F., Stolz, J. A., & Besner, D. (2007). Simon says: Reliability and the role of working memory and attentional control in the Simon task. *Psychonomic Bulletin and Review*, *14*, 313–319.

Brooks, L. R., & Hannah, S. D. (2006). Instantiated features and the use of "rules." *Journal of Experimental Psychology: General*, *135*, 133–151.

Brookshire, G., Casasanto, D., & Ivry, R. (2010). *Modulation of motor-meaning congruity effects for valenced words*. Paper presented at the Proceedings of the 32nd Annual Conference of the Cognitive Science Society, Portland, OR.

Bub, D. N., Masson, M. E. J., & Cree, G. S. (2008). Evocation of functional and volumetric gestural knowledge by objects and words. *Cognition*, *106*, 27–58.

Casasanto, D. (2008). *Body-specific abstract concepts: mental metaphors without linguistic metaphors*. Paper presented at the International Conference on Language, Communication, and Cognition, Brighton, UK.

Chan, A. M., Baker, J. M., Eskandar, E., Schomer, D., Ulbert, I., Marinkovic, K., Cash, S. S., & Halgren, E. (2011). First-pass selectivity for semantic categories in human anteroventral temporal lobe. *Journal of Neuroscience*, *31*, 18119–18129.

Coltheart, M. (1981). The MRC psycholinguistic database. *Quarterly Journal of Experimental Psychology*, *33A*, 497–505.

Conrad, C. (1978). Some factors involved in the recognition of words. In J. W. Cotton, & R. L. Klatzky (Eds.), *Semantic factors in cognition* (pp. 103–120). Hillsdale, NJ: Erlbaum.

de la Vega, I., de Filippis, M., Lachmair, M., Dudschig, C., & Kaup, B. (2012). Emotional valence and physical space: Limits of interaction. *Journal of Experimental Psychology: Human Perception and Performance*, *38*, 375–385.

Dehaene, S., Bossini, S., & Giraux, P. (1993). The mental representation of parity and number magnitude. *Journal of Experimental Psychology: General*, *122*, 371–396.

Dove, G. (2009). Beyond perceptual symbols: A call for representational pluralism. *Cognition*, *110*, 412–431.

Dudschig, C., Souman, J., Lachmair, M., de la Vega, I., & Kaup, B. (2013). Reading "sun" and looking up: The influence of language on saccadic eye movements in the vertical dimension. *PLoS ONE*, *8*, 1–7.

Duschere, K., Holender, D., & Molenaar, E. (2008). Revisiting the affective Simon effect. *Cognition and Emotion*, *22*(2), 193–217.

Estes, Z., Verges, M., & Barsalou, L. W. (2008). Head up, foot down: Object words orient attention to the objects' typical location. *Psychological Science*, *19*, 93–97.

Evans, V. (2009). *How words mean: Lexical concepts, cognitive models, and meaning construction*. New York: Oxford University Press.

Fias, W., Lauwereyns, J., & Lammertyn, J. (2001). Irrelevant digits affect feature-based attention depending on the overlap of neural circuits. *Cognitive Brain Research*, *12*(3), 415–423.

Folk, C. L., Remington, R. W., & Johnston, J. C. (1992). Involuntary covert orienting is contingent on attentional control settings. *Journal of Experimental Psychology: Human Perception and Performance*, *18* (4), 1030–1044.

Forbus, K. D., Gentner, D., & Law, K. (1995). MAC/FAC: A model of similarity-based retrieval. *Cognitive Science*, *19*, 141–205.

Gawronski, B., & Cesario, J. (2013). Of mice and men: What animal research can tell us about context effects on automatic responses in humans. *Personality and Social Psychology Review*, *17*, 187–215.

Gelman, S. A. (2003). *The essential child: Origins of essentialism in everyday thought*. Oxford, UK: Oxford University Press.

Gentner, D., & Forbus, K. D. (1991). MAC/FAC: A model of similarity-based access and mapping. In *Proceedings of the Thirteenth Annual Conference of the Cognitive Science Society*, 504–509. Hillsdale, NJ: Lawrence Erlbaum.

Glaser, W. R. (1992). Picture naming. *Cognition*, *42*, 61–105.

Glenberg, A. M., & Kaschak, M. P. (2002). Grounding language in action. *Psychonomic Bulletin & Review*, *9*, 558–565.

Goldfarb, L., Aisenberg, D., & Henik, A. (2011). Think the thought, walk the walk—Social priming reduces the Stroop effect. *Cognition*, *118*, 193–200.

Graesser, A. C., Singer, M., & Trabasso, T. (1994). Constructing inferences during narrative text comprehension. *Psychological Review*, *101*(3), 371–395.

Greenspan, S. L. (1986). Semantic flexibility and referential specificity of concrete nouns. *Journal of Memory and Language*, *25*, 539–557.

Heck, R. H., Thomas, S. L., & Tabata, N. (2014). *Multilevel and longitudinal modeling with IBM SPSS* (2nd ed.). New York: Routledge.

Herring, D. R., White, K. R., Jabeen, L. N., Hinojos, M., Terrazas, G., Reyes, S. M., Taylor, J. H., & Crites, S. L. (2013). On the automatic activation of attitudes: A quarter century of evaluative priming research. *Psychological Bulletin*, *139*, 1062–1089.

Heyes, C. (2011). Automatic imitation. *Psychological Bulletin*, *137*, 463–483.

Jacoby, L. L., Lindsay, D. S., & Hessels, S. (2003). Item-specific control of automatic processes: Stroop process dissociations. *Psychonomic Bulletin and Review*, *10*, 638–644.

James, W. (2008). *The meaning of truth*. Rockville, MD: Arc Manor. (Original work published 1909).

Jax, S. A., & Buxbaum, L. J. (2010). Response interference between functional and structural actions linked to the same familiar object. *Cognition*, *115*(2), 350–355.

Jeffreys, H. (1961). *Theory of probability* (3rd ed.). Oxford, UK: Oxford University Press.

Jonides, J. (1981). Voluntary vs. automatic control over the mind's eye's movement. In J. B. Long & A. D. Baddeley (Eds.), *Attention and performance IX* (pp. 187–203). Hillsdale, NJ: Erlbaum.

Kass, R. E., & Raftery, A. E. (1995). Bayes factors. *Journal of the American Statistical Association*, *90*, 773–795.

Keil, F. C., & Batterman, N. (1984). A characteristic-to-defining shift in the acquisition of word meaning. *Journal of Verbal Learning and Verbal Behavior*, *23*, 221–236.

Keus, I. M., & Schwarz, W. (2005). Searching for the functional locus of the SNARC effect: Evidence for a response-related origin. *Memory & Cognition*, *33*(4), 681–695.

Kiefer, M., Adams, S. C., & Zovko, M. (2012). Attentional sensitization of unconscious visual processing: Top-down influences on masked priming. *Advances in Cognitive Psychology*, *8*, 50–61.

Kornblum, S., Hasbroucq, T., & Osman, A. (1990). Dimensional overlap: Cognitive basis for stimulus–response compatibility—A model and taxonomy. *Psychological Review*, *97*, 253–270.

Labruna, L., Fernandez-del-Olmo, M., Landau, A., Duque, J., & Ivry, R. B. (2011). Modulation of the motor system during visual and auditory language processing. *Experimental Brain Research*, *211*, 243–250.

Lachmair, M., Dudschig, C., De Filippis, M., de la Vega, I., & Kaup, B. (2011). Root versus roof: Automatic activation of location information during word processing. *Psychonomic Bulletin Review*, *18*, 1180–1188.

Lakens, D. (2012). Polarity correspondence in metaphor congruency effects: Structural overlap predicts categorization times for bipolar concepts presented in vertical space. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *38*, 726–736.

Lakens, D. (2013). Calculating and reporting effect sizes to facilitate cumulative science: A practical primer for t-tests and ANOVAs. *Frontiers in Psychology*, *4*, 863–900.

Lammertyn, J., Notebaert, W., Gevers, W., & Fias, W. (2007). The size of the Simon effect depends on the nature of the relevant task. *Experimental Psychology*, *54*(3), 202–214.

Landauer, I. K., & Dumais, S. T. (1997). A solution to Plat's problem: The latent semantic analysis theory of acquisition, induction, and representation of knowledge. *Psychological Review*, *104*, 211–240.

Louwerse, M., & Connell, L. (2011). A taste of words: Linguistic context and perceptual simulation predict the modality of words. *Cognitive Science*, *35*, 381–398.

Lu, C., & Proctor, R. W. (1995). The influence of irrelevant location information on performance: A review of the Simon and spatial Stroop effects. *Psychonomic Bulletin & Review*, *2* (2), 174–207.

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

MacLeod, C. M. (1991). Half a century of research on the Stroop effect: An integrative review. *Psychological Bulletin*, *109*, 163–203.

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

Malt, B. C. (1994). Water is not H_2O. *Cognitive Psychology*, *27*, 41–70.

McCloskey, M., & Glucksberg, S. (1979). Decision processes in verifying category membership statements: Implications for models of semantic memory. *Cognitive Psychology*, *11*, 1–37.

McKoon, G., & Ratcliff, R. (1992). Inference during reading. *Psychological Review*, *99*, 440–466.

Meier, B. P., & Robinson, M. D. (2004). Why the sunny side is up: Associations between affect and vertical position. *Psychological Science*, *15*, 243–247.

Meyer, D. E., & Schvaneveldt, R. W. (1971). Facilitation in recognizing pairs of words: Evidence of a dependence between retrieval operations. *Journal of Experimental Psychology*, *90*(2), 227–234.

Moors, A., & De Houwer, J. (2006). Automaticity: A theoretical and conceptual analysis. *Psychological Bulletin*, *132*, 297.

Moseley, R. L., Pulvermüller, F., & Shtyrov, Y. (2013). Sensorimotor semantics on the spot: brain activity dissociates between conceptual categories within 150 ms. *Scientific reports*, 3.

Paivio, A. (1971). *Imagery and verbal processes*. New York: Holt, Rinehart, & Winston.

Papeo, L., Negri, G. A. L., Zadini, A., & Rumiati, R. I. (2010). Action performance and action-word understanding: Evidence of double dissociations in left-damaged patients. *Cognitive Neuropsychology*, *27* (5), 428–461.

Papeo, L., & Rumiati, R. I. (2012). Lexical and gestural symbols in left-damaged patients. *Cortex*, *49*, 1668–1678.

Papeo, L., Rumiati, R. I., Cecchetto, C., & Tomasino, B. (2012). On-line changing of thinking about words: The effect of cognitive context on neural responses to verb reading. *Journal of Cognitive Neuroscience*, *24*(12), 2348–2362.

Papeo, L., Vallesi, A., Isaja, A., & Rumiati, R. I. (2009). Effects of TMS on different stages of motor and non-motor verb processing in the primary motor cortex. *PLoS ONE*, *4*(2), 1–11.

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

Proctor, R. W., & Cho, Y. S. (2006). Polarity correspondence: A general principle for performance of speeded binary classification tasks. *Psychological Bulletin*, *132*, 416–442.

Ralph, M. A. L., Cipolotti, L., Manes, F., & Patterson, K. (2010). Taking both sides: Do unilateral anterior temporal lobe lesions disrupt semantic memory? *Brain*, *133*(11), 3243–3255.

Remington, R. W., Johnston, J. C., & Yantis, S. (1992). Involuntary attentional capture by abrupt onsets. *Perception & Psychophysics*, *51*, 279–290.

Rotteveel, M., & Phaf, R. H. (2004). Automatic affective evaluation does not automatically predispose for arm flexion and extension. *Emotion*, *4*(2), 156–172.

Rouder, J. N., Speckman, P. L., Sun, D., Morey, R. D., & Iverson, G. (2009). Bayesian t tests for accepting and rejecting the null hypothesis. *Psychonomic Bulletin & Review*, *16*(2), 225–237.

Santana, E., & de Vega, M. (2011). Metaphors are embodied, and so are their literal counterparts. *Frontiers in Psychology*, *2*, 1–12.

Santiago, J., Ouellet, M., Roman, A., & Valenzuela, J. (2012). Attentional factors in conceptual congruency. *Cognitive Science*, *36*, 1051–1077.

Santiago, J., Roman, A., & Ouellet, M. (2011). Flexible foundations of abstract thought: A review and a theory. In T. Schubert, & A. Mass (Eds.), *Spatial dimensions of social thought* (pp. 39–108). Berlin: Mouton de Gruyter.

Sato, M., Mengarelli, M., Riggio, L., Gallese, V., & Buccinoet, G. (2008). Task related modulation of the motor system during language processing. *Brain and Language*, *105*, 85–90.

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

Shiffrin, R. M., & Schneider, W. (1977). Controlled and automatic human information processing: II. Perceptual learning, automatic attending, and a general theory. *Psychological Review*, *84*(2), 127–190.

Shtyrov, Y., Butorina, A., Nikolaeva, A., & Stroganova, T. (2014). Automatic ultrarapid activation and inhibition of cortical motor systems in spoken word comprehension. *Proceedings of the National Academy of Sciences*, *111*, E1918–E1923.

Simon, J. R. (1969). Reactions towards the source of stimulation. *Journal of Experimental Psychology*, *81*, 174–176.

Simon, J. R. (1990). The effects of an irrelevant directional cue on human information processing. In R. W. Proctor & T. G. Reeva (Eds.), *Stimulus-response compatibility: An integrated perspective* (pp. 31–86). Amsterdam: North-Holland.

Smith, E. E., Shoben, E. J., & Rips, L. J. (1974). Structure and process in semantic memory: A featural model for semantic decisions. *Psychological Review*, *81*, 214–241.

Snijders, A. B., & Bosker, R. J. (2012). *Multilevel analysis: An introduction to basic and advanced multilevel modeling* (2nd ed.). London: Sage Publications.

Spruyt, A., De Houwer, J., Hermans, D., & Eelen, P. (2007). Affective priming of nonaffective semantic categorization responses. *Experimental Psychology*, *54*(1), 44–53.

Storbeck, J., & Robinson, M. D. (2004). Preferences and inferences in encoding visual objects: A systematic comparison of semantic and affective priming. *Personality and Social Psychology Bulletin*, *30*, 80–93.

Stroop, J. R. (1935). Studies of interference in serial verbal reactions. *Journal of Experimental Psychology*, *18*(6), 643–662.

Swinney, D. A. (1979). Lexical access during sentence comprehension: (Re)consideration of context effects. *Journal of Verbal Learning and Verbal Behavior*, *18*, 645–659.

Tanenhaus, M. K., Leiman, J. M., & Seidenberg, M. S. (1979). Evidence for multiple stages in the processing of ambiguous words in syntactic contexts. *Journal of Verbal Learning and Verbal Behavior*, *18*, 427–440.

Thornton, T., Loetscher, T., Yates, M. J., & Nicholis, M. E. R. (2013). The highs and lows of the interaction between word meaning and space. *Journal of Experimental Psychology: Human Perception and Performance*, *39*, 964–973.

Tomasino, B., & Rumiati, R. I. (2013). At the mercy of strategies: The role of motor representations in language understanding. *Frontiers in Psychology*, *4*, 27 (1–13).

van Dam, W. O., van Dijk, M., Bekkering, H., & Rueschemeyer, S. (2011). Flexibility in embodied lexical-semantic representations. *Human Brain Mapping*, *33*(10), 2322–2333.

van Elk, M., & Blanke, O. (2011). Manipulable objects facilitate cross-modal integration in peripersonal space. *PLoS ONE*, *6*(9), 1–7.

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

Wilson-Mendenhall, C. D., Barrett, L. F., Simmons, W. K., & Barsalou, L. W. (2011). Grounding emotion in situated conceptualization. *Neuropsychologia*, *49*, 1105–1127.

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

Zwaan, R. A., & Madden, C. J. (2005). Embodied sentence comprehension. In D. Pecher & R. Zwaan (Eds.), *Grounding cognition: The role of perception and action in memory, language, and thought* (pp. 224–245). New York: Cambridge University Press.

Zwaan, R. A., & Yaxley, R. H. (2003). Hemispheric differences in semantic-relatedness judgments. *Cognition*, *87*, B79–B86.

Appendix: Verticality ratings for the materials used in the three experiments, using a verticality scale that ranged from −9 (low vertical position) to +9 (high vertical position)

Word	Mean	*SE*	Word	Mean	*SE*	Word	Mean	*SE*	Word	Mean	*SE*
High words											
airship	6.47	0.54	apex	6.53	0.56	arch	5.00	0.35	balloon	5.35	0.46
belfry	2.82	0.49	branch	4.29	0.44	ceiling	5.18	0.54	cloud	7.35	0.31
collar	2.53	0.42	crown	4.35	0.48	eagle	6.47	0.43	float	4.06	0.63
fly	6.44	0.46	forehead	2.56	0.48	fountain	3.53	0.33	hair	3.41	0.57
hat	2.88	0.38	hawk	6.47	0.41	head	3.47	0.55	helmet	3.18	0.46
hill	4.71	0.42	hood	2.88	0.36	kite	6.06	0.42	lark	4.76	0.55
lift	5.35	0.54	lightning	6.65	0.46	missile	6.24	0.42	monument	5.41	0.49
moon	8.18	0.23	mountain	7.06	0.33	plane	7.53	0.31	platform	3.24	0.55
pulpit	3.65	0.55	rocket	7.71	0.33	roof	5.12	0.34	sky	8.18	0.18
skyscraper	7.41	0.35	smoke	5.24	0.46	spire	4.88	0.53	star	8.47	0.19
steeple	5.13	0.33	sun	8.18	0.30	tip	3.76	0.66	tornado	5.59	0.54
tower	6.29	0.40	umbrella	3.41	0.48	volcano	6.59	0.33	wig	2.76	0.45
Low words											
anchor	−5.18	0.46	ankle	−2.18	0.41	asphalt	−2.76	0.34	base	−3.41	0.55
basement	−4.06	0.42	basin	−3.41	0.44	boot	−2.29	0.36	burrow	−4.18	0.44
carpet	−2.94	0.40	casket	−4.53	0.67	cave	−3.47	0.53	cellar	−3.94	0.46
chasm	−5.65	0.62	dip	−2.88	0.35	dirt	−3.13	0.36	ditch	−3.82	0.50
dive	−4.41	0.54	drain	−3.13	0.38	dungeon	−5.94	0.39	earth	−3.76	0.67
fall	−4.59	0.69	feet	−2.76	0.42	floor	−2.71	0.40	foundation	−4.00	0.44
grass	−2.35	0.45	ground	−3.76	0.53	heel	−2.76	0.42	hole	−3.47	0.48
land	−3.00	0.42	mat	−2.88	0.42	pit	−5.13	0.48	puddle	−2.59	0.49
root	−4.65	0.51	rug	−3.00	0.49	sand	−2.47	0.39	sewer	−4.88	0.56
shoe	−2.44	0.38	sink	−3.94	0.57	slipper	−2.00	0.30	sock	−2.18	0.37
soil	−3.06	0.42	submarine	−6.71	0.34	sunset	−4.12	0.64	toe	−2.65	0.38
tomb	−5.47	0.54	tumble	−2.94	0.30	valley	−4.63	0.43	worm	−3.06	0.42
Closed-class non-spatial words											
a			everyone			himself			likewise		
of			or			ours			where		