Published in final edited form as:
J Exp Psychol Gen. 2015 December ; 144(6): e116–e141. doi:10.1037/xge0000125.

Just Out of Reach: On the Reliability of the Action-Sentence Compatibility Effect

Megan H. Papesh
Louisiana State University

Abstract

The action-sentence compatibility effect (ACE; Glenberg & Kaschak, 2002), a hallmark finding in Embodied Cognition, implicates the motor system in language comprehension. In the ACE, people process sentences implying movement toward or away from themselves, responding with actions toward or away from their bodies. These processes interact, implying a linkage between linguistic and motor systems. From a theoretical perspective, the ACE has been extremely influential, being widely-cited evidence in favor of embodied cognition. The present study began as an attempt to extend the ACE in a new direction, but eventually became a series of attempts to simply replicate the effect. Across eight experiments, I tested whether the ACE extends to a novel mouse-tracking method and/or is susceptible to higher-order cognitive influences. In three experiments, attempts were made to “disembody” the ACE by presenting participants' names on the computer screen (as in Markman & Brendl, 2005). In each experiment, the ACE could not be disembodied, because the ACE did not occur. In further experiments, the ACE was not observed in reading times, regardless of response mode (mouse movements versus button-presses) or stimuli, including those from the original research. Similarly, no ACE was observed in physical movement times. Bayes Factor analyses of the current experiments, and the previous ACE literature, suggest that the evidence for the ACE is generally weak: Many studies considered as positive evidence actually support the null hypothesis, and very few published results offer strong evidence for the ACE. Implications for the embodiment hypothesis are discussed.

Keywords

Embodied cognition; reading; action-sentence compatibility

In the emerging field of Embodied Cognition (EC), the central theoretical claim holds that the mind and body are not merely intertwined, but that the physical characteristics of the body *directly* shape cognitive activity (Glenberg, Witt, & Metcalfe, 2013). Specifically, EC proposes that the representational format of all cognitive processes is inherently sensorimotor. As a result, cognition can only be fully understood by considering the “big picture” – the physical ensemble of a person situated in some environment. This assertion makes EC an exciting idea, one that has inspired new theoretical debates about classic cognitive science (see Glenberg, 2015; Mahon, 2015a, 2015b). But what does it mean for cognition to be embodied, and why is embodiment even necessary?

Correspondence: M.H. Papesh, Department of Psychology, Louisiana State University, Baton Rouge, LA, 70803. mpapesh@lsu.edu.

Proponents of the EC hypothesis offer it as an alternative to “traditional” cognitive science, which is typically portrayed using the computer metaphor of mind, a caricature of cognitive science that refers to the early 1970s (e.g., Newell & Simon, 1972). This metaphor, of the mind as a computing device, suggests that cognitive processes are computed in the brain using abstract, a modal symbols, independent of the systems involved in perception and action (Barsalou, 2008). To illustrate the putative problems with such abstract, amodal symbols, Glenberg and Kaschak (2002) described Harnad's (1990) adaptation of Searle's (1980) “Chinese Room” argument: Suppose a foreigner lands at a Chinese airport, knowing none of the local language, but carrying a full Chinese dictionary. When trying to interpret the airport signs, the traveler will become stuck in an endless loop of abstract symbols, as every definition in his dictionary references other symbols. This problem underscores the need for abstract symbols to be *grounded* in the environment (e.g., Lakoff, 1987), and motivates the core assumption of EC: Cognitive processes are composed of modality-specific, sensorimotor interactions with the environment.

Masson (2015) recently highlighted two bodies of evidence supporting the EC hypothesis, from neuroimaging and behavioral studies. Ample neuroimaging evidence suggests that sensorimotor and cognitive processes are interactive, and often recruit similar (or overlapping) brain regions. For example, reading effector-specific verbs (e.g., *kick*) activates brain regions involved in generating leg movements (Hauk et al., 2004; Tettamanti et al., 2005). Similarly, reading about visual motion yields MT/V5 activity, whereas reading about static visual images does not (Rueschemeyer, Glenberg, Kaschak, Mueller, & Friederici, 2010). Although studies of patients with damage to sensorimotor brain areas provide only weak evidence of concomitant deficits in concept comprehension (see discussions in Hickok, 2009; Mahon & Caramazza, 2008), studies using transcranial magnetic stimulation (TMS) suggest a closer relationship. For example, Pulvermüller, Hauk, Nikulin, and Ilmoniemi (2005) applied TMS to hand or foot motor areas contralateral to participants' dominant hands during lexical decisions, and observed faster decisions to effector-specific words. Such interactions between sensorimotor activity and cognitive processes have led to the strong embodied hypothesis that sensorimotor experiences not only contribute to cognitive processes, but they constitute cognition (Glenberg, 2015).

In addition to the neuroimaging work suggesting a link between sensorimotor systems and cognitive processes, the embodied hypothesis is supported by behavioral findings, often in the domain of language[1]. Traditional cognitive theories describe linguistic processing as the activation and combination of stored knowledge. For example, a theory may describe how basic phonemic and lexical units combine to produce meaningful utterances, and how those meaningful utterances combine to produce a vast range of linguistically lawful statements. According to a strong view of embodiment, however, linguistic content is not different from perceptual content, such that comprehending the sentence “Jeff caught the ball” is essentially the neural equivalent of perceiving Jeff catching the ball. Embodied accounts of

---

[1] Language studies do not constitute the sole evidence in support of EC. “Embodied” effects are also seen in object identification (e.g., grip-size congruency effects; Tucker & Ellis, 2001), perception (e.g., hill-slant estimation; Proffitt, 2006), and education (e.g., Kontra, Lyons, Fischer, & Beilock, 2015), with recent evidence suggesting a role for embodiment in cultural effects (Soliman, Gibson, & Glenberg, 2013).

language often cite Gibsonian views of perception as a framework to understand cognition, relying on the concept of affordances[2] (e.g., Barsalou, 1999; Glenberg & Robertson, 1999, 2000), and actively denying any role for mental representations (e.g., Chemero, 2011; Wilson & Golonka, 2013).

By embodied accounts of language processing, perceivers are not only sensitive to potential physical interactions with linguistically referenced objects (i.e., affordances). Instead, perceivers are theorized to *simulate* sentence-implied actions, such that achieving motor resonance leads to action understanding (Zwaan & Taylor, 2006). In a now-classic demonstration of such sentence simulation, Glenberg and Kaschak (2002) had participants make sensibility judgments to sentences describing action either toward or away from themselves, and sentences describing *transfer* to or from oneself (nonsense sentences, without directional components, were also included). For example, two imperative action sentences were “Open the drawer” and “Close the drawer,” implying motion toward and away from the body, respectively. Similarly, transfer sentences could include giving or receiving either concrete objects (e.g., *the bottle*) or abstract concepts (e.g., *the compliment*). As predicted by EC, if perceivers automatically simulate sentence-implied actions, then processing transfer verbs should be influenced by the motor component of the physical response, producing facilitation when the verb and motoric action are congruent, and interference when they are incongruent.

Glenberg and Kaschak (2002) developed an innovative method for collecting sensibility judgments, allowing them to examine whether overt motor *behaviors* interact with (theorized) motoric *simulation* during language processing. Participants made “yes/no” sensibility decisions to sentences on an elongated response-box; a central key served as a “launching point,” and response buttons were located near and far from the body. With this set up, responding “sensible” could involve moving the arm either toward or away from oneself. When the movement described by the sentence matched the movement involved in responding, participants' reading times were facilitated. (“Reading times” were defined as the latency between sentence onset and the participant releasing the central, start key.) When the sentences and intended movements were incompatible, participants' reading times were slowed (a comparable result is obtained when responses are made by turning a dial; Zwaan & Taylor, 2006).

The interaction between bodily movement and implied sentence direction is called the *action-sentence compatibility effect* (ACE), and it has been cited as evidence that language comprehension is inherently embodied (sensorimotor), rather than symbolic. It is difficult to overstate how influential the ACE has been, in terms of motivating and validating EC as a counter-theory to information-processing approaches in cognitive science. As of September, 2015, the article by Glenberg and Kaschak (2002) had been cited 1538 times (*Google Scholar*). After more than a decade since its original publication, the ACE is widely considered scientific fact, with deep theoretical implications. For example, Weiskopf (2010), noted that interactions between motoric and linguistic processes are critical for the EC hypothesis: They imply that sentence comprehension involves motor simulation, such that

---

[2] From J.J. Gibson's ecological psychology, affordances are potentials for interaction with a given object.

language comprehension and physical action are cognitively and neurally inseparable (Glenberg et al., 2008). Indeed, compatibility effects, such as the ACE, constitute the majority of behavioral evidence in support of embodiment (Masson, 2015).

The embodied view of language processing stands in stark contrast to the “amodal” view, often cast as a theoretical competitor. According to amodal views, language processes map abstract symbols (e.g., words) to their semantic and conceptual referents, without necessarily relying on sensory-motor systems (for reviews, see Horchak et al., 2014). Instead, these are viewed as separate, but interacting systems (e.g., Mahon, 2015a). The debate surrounding the format of cognitive representation, whether they are sensorimotor (embodied) or abstract, shows no sign of abating (e.g., Glenberg, 2015; Mahon, 2015a). Thus, cognition-action couplings such as ACE are theoretically critical, and also provide opportunities to potentially constrain embodied theories.

Importantly, the temporal dynamics of ACE suggest that motor simulation occurs online, during language comprehension, rather than as post-comprehension translation of sentential actions (e.g., off-line priming). In fact, temporal properties of the ACE can be rather nuanced: de Vega et al. (2011) observed interference when congruent actions and verbs occurred within 100-200 ms of each other, suggesting that the language and action competed for neural resources, but facilitation when the action and verb occurred 350 ms apart. Others have observed that compatibility effects are only observed when participants know the response mapping prior to reading the sentence, suggesting that the action features must be concurrently active to produce facilitation (Borreggine & Kaschak, 2006). Similar effects for motor involvement in language comprehension have been repeatedly observed for both words and sentences (for a review, see Pulvermüller, 1999). In the current study, the general method from Glenberg and Kaschak (2002) was adopted, such that verbs and actions were temporally extended from one another, and response mapping was known prior to all sentences.

Interactions between motor activity and cognition exist beyond the ACE; evaluative judgments are often affected by ongoing physical activity, such as unrelated facial expressions. For example, Strack, Martin, and Stepper (1988) found that participants who were forced to hold a frowning facial expression evaluated cartoons as less funny than their smiling counterparts. Numerous studies on the impact of unrelated brow furrowing (activation of the corrugator muscle, which is commonly observed during difficult tasks, requiring intense concentration) have been reported. Brow furrowing changes performance, relative to other facial expressions, in judgments of difficulty (Stepper & Strack, 1993), fluency/confidence (Alter et al., 2007), preference (Tamir et al., 2004), and fame judgments (Strack & Neumann, 2000). In each case, furrowing one's brow leads people to interpret tasks as more challenging, or their perceptions as more disfluent, reflecting an interaction between bodily states and ongoing evaluative decisions.

The literature on brow furrowing suggests that bodily movements can be used to manipulate ongoing cognitive processing. In the broader psychological science literature, body movements can also *reveal* ongoing processes. For example, arm movements (e.g., mouse-tracking) can reflect cognitive processes in social perception (Freeman & Ambady, 2009), language (Spivey, Grosjean, & Knoblich, 2005), and memory (Papesh & Goldinger, 2012). The speed and force of arm movements during valence judgments can reveal preferences and desires: Pulling motions, for example, are related to approach-related desires for an object or word under scrutiny, while pushing motions are related to avoidance, as in pushing something away from one's body (e.g., Cacioppo, Priester, & Bernston, 1993; Chen & Bargh, 1999; Solarz, 1960; but see Wentura, Rothermund, & Bak, 2000). The EC interpretation of such findings rests on the (reasonable) assumption that people's representation of "self" is situated in their physical bodies. Using a strikingly simple manipulation, Markman and Brendl (2005) found that participants' senses of self could be decoupled from their physical bodies by printing their names on the computer screen. While reading positively and negatively valenced words, participants either pulled or pushed a lever toward or away from their names. Depending on the condition, this movement was either compatible with their physical bodies (e.g., pulling the lever toward the name was consistent with pulling it toward oneself) or was incompatible. Markman and Brendl (2005) found that response times (RTs) to positive words were reliably faster when the lever movement was toward participants' names, not their physical bodies, suggesting that the sense of self can be "disembodied." Likewise, RTs to negative words were faster when the movement direction "pushed" the word away from the participants' names, even if it simultaneously pulled the word toward their physical bodies. These results constrain EC theories by suggesting that action-cognition links, while clearly being tied to perceptual and motor processes, may also rely on higher-order cognitive processes, such as symbolic information about how the self is conceptualized in physical space.

In the present study, the original motivation was to further examine the ACE, testing response patterns when the "disembodiment" manipulation from Markman and Brendl (2005) was added to the design. This manipulation was intended to reveal the ACE, and determine whether the effect can be reversed by decoupling participants' representation of "self" from their physical bodies. To document the ACE, one needs to observe a double dissociation in either reading or movement times, which manifests statistically as a crossover interaction between movement direction and implied sentence direction. It is not sufficient for an effect to emerge in only half of the design (e.g., if away sentences facilitate away movements); both halves must produce reliable effects. A disembodied ACE, on the other hand, would produce the standard ACE effect with an additional interaction factor, the location of the participant's name on the computer screen. As will become evident, the original goal of disembodying the ACE was thwarted: In order to test whether the ACE can be modulated by a new manipulation, a necessary precondition is that the ACE must be observed. Establishing this basic effect, however, proved challenging. After repeated failures, the eventual goal changed, becoming instead a focused attempt to replicate the ACE, and then taking a closer look at the prior ACE literature. To ensure adequate power, all experiments maintained a sample size above the recommended value obtained in G*Power (Faul, Erdfelder, Lang, & Buchner, 2007) for large effects (.40)[3]. For between-

---

[3] A large effect size was used as the estimate due to the large theoretical implications of the ACE, and because Glenberg and Kaschak (2002) did not report sufficient statistics to calculate effect sizes. To further justify sample sizes, I examined the average sample size for all studies reported in Table 2 (see General Discussion). Sample sizes for published ACE studies ranged from 9 to 89, with a median of 48; only two of the current experiments do not exceed this median value (Experiments 3 and 4B).

subjects comparisons, the minimum N was 36, and for within-subjects comparisons, the minimum N was 10.

Experiments 1-3

Experiments 1 and 2 were conducted to determine whether the ACE could be constrained by higher-order cognitive variables, specifically a "disembodied" sense of self. If so, then sentence judgment times should produce a three-way interaction between the location of the name on the computer screen, the location of the response option, and the direction implied by the sentence. For example, when a participant's name is at the top of the screen with the "sensible" response option, she should be faster to verify "Abby gave you a coin," relative to "You gave Abby a coin," because the physical response movement would be toward her name. On the other hand, if the ACE is truly embodied to the physical self, the location of the name on the computer screen should not affect RTs; effects should be driven solely by the interaction between the implied movement direction and the physical movement of the response (i.e., the aforementioned double dissociation of sentence and movement directions). Experiment 3 was conducted to replicate the ACE using the same general method, but without the disembodiment manipulation.

General Method: Experiments 1-3

Experiments 1-3 followed the same basic mouse-tracking method, in which physical arm movements were recorded by the computer mouse, and are presented together for brevity[4]. Procedural changes are noted below.

Participants

Across all experiments, participants were recruited from Psychology classes, and they participated in exchange for partial course credit. In addition, all participants were right-handed and native English speakers. The breakdown of participants by experiment and condition can be found in Table 1.

Stimuli

The 240 sentences used in Experiments 1 through 3 contained an equal number of sensible and nonsense sentences. For both the sensible and nonsense sentences, 60 implied "toward the body" motion and 60 implied "away from the body" motion. Within each subset of 60 toward and away sentences were 30 abstract and 30 concrete sentences. Toward and away sentences were created by rearranging the "giver" and "receiver" roles for the sentences. Only one version of each sentence was shown within a given experimental block, and the name of the actor was changed across sentence versions, to reduce the likelihood that participants would notice the overlapping content. All sentences can be found in Appendixes A-C.

---

[4] Although the original ACE experiment (Glenberg & Kaschak, 2002) involved movements over a custom response box, follow-up studies involved varied physical responses, including knob turning (Zwaan & Taylor, 2006) foot pedals (Buccino et al., 2005), and keyboards (Borreggine & Kaschak, 2006). There was no reason to suspect, a priori, that arm movements with the computer mouse would not show the same effects. As will be discussed, when the ACE failed to generalize to the mouse-movement approach, it motivated more direct replication attempts.

Procedure

Prior to each experiment, the computer mouse speed was changed to “very slow,” so that small changes in the location of the mouse cursor were associated with relatively large arm movements. After participants indicated that they understood the task, they completed 20 practice trials to familiarize themselves with the response mechanics, followed by 240 experimental trials, with a 3-minute break in the middle. Before resuming the experiment after the break, all participants completed more practice trials. Throughout the experiment, each trial began with a sentence printed in the vertical and horizontal center of the screen, and the mouse cursor centered directly below the text. As shown in Figure 1, this paradigm permitted two RT measures, reading time and movement time. When participants were ready to issue their sensibility judgments, they clicked on the sentence (this did not require a mouse movement), which then turned into a black box. This procedure was meant to encourage participants to finish reading before initiating a movement; latency to click the sentence was used as a measure of reading time. To indicate a decision, participants dragged the black box to a sensible or nonsense response area, located at the top or bottom of the screen (counter balanced by participants). When the mouse button was released and the black box was inside the response area, the trial ended; the mouse release was used as a measure of movement time.

In Experiments 1 and 2, participants' names were also printed on the computer screen, such that name location was compatible with physical location in half the trials (see Figure 2). Starting name location was counterbalanced across participants, and the location switched after the break. Although participants were not told the motivation for the name manipulation, they were made aware of it. The only procedural difference between Experiments 1 and 2 occurred during the inter-trial interval: In Experiment 2, participants were required to center the mouse in a black box drawn on the desk after every trial; this was not the case in Experiment 1. (The location of the box was the same for all participants, such that the tip of the mouse was located exactly 25.4 cm from the front edge of the desk. The box was located 8.89 cm from the right edge of the desk and was drawn to 7.62 × 12.7 cm.).

Data Analysis

With regard to data analysis, two points merit special mention. First, for experiments similar to those reported here (which include variability in both subjects and items), there is a growing trend to conduct analyses using linear mixed-effects models (e.g., Barr, Levy, Scheepers, & Tily, 2014). In the current study, however, data were analyzed using mixed-model, repeated-measures ANOVAs. There were several reasons for this approach. One was design-based: Because all sentences were used equally in all conditions, item-level variation was not a key concern. Another reason was to maintain direct comparability to previous research. Finally, and most important, upon seeing the data, it became clear that different analyses would not yield substantially different results.

Second, the literature on the ACE is a bit unclear, regarding the proper data to examine. The present experiments were modeled after Glenberg and Kaschak (2002), such that reading time and movement time could each be recorded. With that being the case, the approach taken by the original authors might surprise some readers. In the opening paragraph of their seminal paper, Glenberg and Kaschak (2002, p. 558) wrote:

> "We demonstrate that merely comprehending a sentence that implies action in one direction (e.g., "Close the drawer" implies action away from the body) interferes with real action in the opposite direction (e.g., movement toward the body). These data are consistent with the claim that language comprehension is grounded in bodily action, and they are inconsistent with abstract symbol theories of meaning."

Given this quote, it is surprising to note that Glenberg and Kaschak never actually examined movement times (i.e., "real actions"). Their analyses focused only on trimmed reading times, the latency between sentence onset and release of the center button. Undoubtedly, these are closely related measures, but lifting one's finger may not be isomorphic with movement toward either response button. Moreover, in other ACE studies, researchers have focused on actual movement time as the key dependent variable (e.g., Nazir et al., 2008). With these inconsistencies in mind, and in hopes of giving the ACE every possible opportunity to emerge in the results, both reading and movement times are reported for every experiment.

Results and Discussion

Across all analyses, alpha was maintained at .05, and multiple comparisons were Bonferroni-corrected. Because analyses were on millisecond timescale data, mean square error (MSe) is reported in seconds (raw MSe divided by 1000). For any participant, outliers were defined as reading or movement times that exceeded their respective means by more than 2.5 standard deviations; outliers were replaced with the cutoff value (Winer, 1971). Only trials with correct sensibility judgments were analyzed, resulting in an average loss of 9 trials (7.5%) per participant in Experiment 1, 8 trials (6.6%) in Experiment 2, and 10 trials (8.4%) in Experiment 3[5]. Decision times to nonsense and no-transfer sentences, by their nature, cannot show compatibility effects and are thus not reported for any of the experiments.

Reading Time: Experiment 1

Reading times during correct "sensible" judgments were analyzed in a 2 (Name Location: Near/Far) × 2 (Sentence Direction: Toward/Away) × 2 (Concreteness: Abstract/Concrete) × 2 (Response Direction: Sensible is near/far) mixed-model RM ANOVA, with Response Direction as the between-subjects factor[6]. Although the hallmark of the ACE is an interaction between Sentence Direction and Response Direction, a "disembodied" ACE would require a three-way interaction that also included Name Location. This three-way interaction was not observed, $F(1, 122) = 1.17$, $MSe = 68.76$, $p = .28$, $n^2_p = .01$, although a

---

[5] Readers familiar with the ACE literature will note that this data trimming procedure does not follow that used in the original study by Glenberg and Kaschak (2000). In their study, "trimmed reading times" were derived by dropping participants with uneven error rates across design cells, dropping the first block of trials per condition, and dropping the fastest and slowest trials in each condition. The present approach is more inclusive, and standard for reading studies, based on statistical screening of outliers (Winer, 1971). This approach resulted in dropping fewer than 10% of trials per experiment (mainly due to errors), and has been used by other researchers in the ACE literature (e.g., Borreggine & Kaschak, 2006).

[6] Although Concreteness was manipulated in all experiments, subsequent analyses collapse across this factor. The data were examined with and without Concreteness, and it did not change any patterns reported.

main effect of Name Location revealed that participants were faster to read sentences when their names appeared at the top of the screen (*M* = 1775 ms, *SE* = 47), relative to the bottom (*M* = 2053 ms, *SE* = 50), *F*(1, 122) = 46.22, *MSe* = 416.2, $p < .0001$, $n^2_p = .28$. This effect cannot be attributed to practice effects, as participants were randomly assigned to begin with either near/far name locations.

The ACE interaction (Sentence Direction × Response Direction) was reliable, *F*(1, 122) = 6.86, *MSe* = 73.05, $p = .01$, $n^2_p = .05$. As shown in the left panel of Figure 3, the observed pattern for the "toward" sentences contradicted the ACE prediction, with substantially faster responses in the "away" direction. However, for the "away" sentences, the observed pattern was consistent with the ACE prediction (i.e., faster responses when the sentences and actions were congruent), with an even larger disparity, leading to the reliable interaction (the between-subjects main effect of Response Direction was not reliable, *F*(1, 122) = 3.24, *MSe* = 113.89, $p = .07$, $n^2_p = .03$). As predicted by embodied accounts of language processing, participants were faster to verify "You gave Abby a coin" when the physical response direction was away from their body. It is challenging to interpret this effect, however, when the other half of the design produced a difference in the "wrong" direction. In the original Glenberg and Kaschak (2002) study, both key experiments (1 and 2A) produced crossover interactions, with each half of the design being consistent with the ACE prediction. As in the original ACE demonstration (Glenberg & Kaschak, 2002), the movement compatibility effect was not limited to concrete sentences, but extended to abstract transfer sentences as well (participants were, however, faster to judge concrete (*M* = 1859 ms, *SE* = 43), relative to abstract (*M* = 1969 ms, *SE* = 47), sentences, *F*(1, 122) = 26.41, *MSe* = 1898.59 $p < .001$, $n^2_p = .18$). Although they do not represent a full replication of the original ACE, these results suggest that the present stimuli are capable of eliciting movement compatibility effects, albeit in a more restricted manner than in the original demonstration.

Movement Time: Experiment 1

Movement times (right panel of Figure 3) were analyzed in the same manner as reading times. There was again no evidence for a disembodied ACE; the critical three-way interaction between Name Location, Movement Direction, and Sentence Direction was not reliable, *F*(1, 122) = 0.12, *MSe* = 14.28, $p = .73$, $n^2_p = .001$. Participants were, however, faster to issue physical responses when their names appeared at the bottom of the screen (*M* = 973 ms, *SE* = 19), relative to the top (*M* = 1060 ms, *SE* = 20), *F*(1, 122) = 21.61, *MSe* = 86.68, $p < .001$, $n^2_p = .15$. There was also a small (20 ms), but reliable, trend for participants to judge "away" sentences more quickly than "toward" sentences (*M* = 1027 ms, *SE* = 18), *F*(1, 122) = 6.72, *MSe* = 14.33, $p = .01$, $n^2_p = .05$. Inconsistent with the reading times, there was no interaction of Sentence Direction × Movement Direction, *F*(1, 122) = 2.74, *MSe* = 14.33, $p = .10$, $n^2_p = .02$ (i.e., no ACE effect). This result is surprising, given the theoretical locus of the ACE in motor simulation during sentence comprehension. If participants simulated the motor components of the transfer sentences, and if this simulation affected reading times, one would expect to see parallel effects in motor response execution, as has been observed by Borreggine and Kaschak (2006).

Taken together, Experiment 1 produced a partial replication of the ACE in reading times (i.e., half the data contradicted the prediction, but the other half confirmed it to a slightly larger degree), and no ACE in movement times. The failure to observe an effect in movement times suggests either that motor simulation during language comprehension is transient (e.g., Pulvermüller et al., 2005), or that the current method was insufficiently sensitive to reveal such effects, should they exist. In reading times, participants always issued the same response to indicate that sentences were comprehended, left-clicking the computer mouse. To issue the sensibility judgment, however, participants moved the blackened sentence to the top or bottom of the screen, involving mouse movements toward or away from themselves. After each response, however, participants' hand positions were uncontrolled, which may have created variable starting positions across trials, adding noise to the results. To more precisely control these mechanics, participants in Experiment 2 centered the mouse in a designated desk area following every trial.

Reading Time: Experiment 2

Reading times were analyzed in a 2 (Name Location: Near/Far) × 2 (Sentence Direction: Toward/Away) × 2 (Response Direction: Sensible is near/far) mixed-model RM ANOVA, with Response Direction as the between-subjects factor. No main effects or interactions emerged in the analyses (all *p*s > .05), including the ACE interaction between Sentence Direction and Movement Direction, $F(1, 52)= .04$, $MSe = 34.72$, $p= .84$, $n^2_p = .001$. As shown in the left panel of Figure 4, there was a numerical trend for facilitation in movement-congruent toward sentences, but this pattern appears to be driven by a general tendency for toward-movements to be faster, regardless of sentence direction. No pair wise comparisons were statistically reliable.

Movement Time: Experiment 2

Although no ACE was observed in the reading times, the goal of Experiment 2 was to improve the estimates of movement time. As such, movement times (right panel of Figure 4) were analyzed in the same manner as reading times. As shown, there was no ACE interaction between Sentence Direction and Movement Direction, $F(1, 52)= .04$, $MSe = 9.26$, $p= .85$, $n^2_p = .001$, nor was there a three-way interaction with Name Location that would indicate a disembodied ACE, $F(1, 52)= .24$, $MSe = 6.03$, $p= .63$, $n^2_p = .005$. There was a main effect of Name Location, $F(1, 52) = 6.47$, $MSe = 61.72$, $p = .01$, $n^2_p = .11$, showing that participants completed mouse movements more quickly when their name was printed at the top ($M = 1049$ ms, $SE = 37$), relative to the bottom ($M = 1135$ ms, $SE = 43$). This effect did not interact with any other variables and is therefore difficult to interpret. Theoretically, there is no reason to expect movement times to be facilitated by the location of one's name on the computer screen, particularly in the absence of an interaction with another factor.

In general, Experiment 2 failed to replicate the (partial) ACE observed in reading times from Experiment 1. It also failed to elicit an ACE in the movement data, despite the increased control over movement consistency. Although the goal of discovering a “disembodied ACE” had thus far failed, it also proved difficult to elicit a consistent ACE. Before an effect can be manipulated and explored, it must first be replicable. The goal of Experiment 3 was to replicate the ACE with the same stimulus set as Experiments 1 and 2, using the mouse-movement paradigm, but without any names on-screen.

Reading Time: Experiment 3

Experiment 3 reading times were analyzed in a 2 (Response Direction) × 2 (Sentence Direction) mixed-model RM ANOVA, with Response Direction as the between-subjects factor. No main effects or interactions were observed, all *p*s > .05, including the hallmark ACE interaction of Sentence Direction and Movement Direction, $F(1, 35)= .05$, $MSe = 18.77$, $p= .82$, $n^2_p = .001$. As shown in the left panel of Figure 5, the results looked similar to those from Experiment 1, with an apparent congruity effect for the "away" sentences, but an equivalent backwards effect for the "toward" sentences. This pattern qualitatively replicated Experiment 1, but with no statistical (or numerical) evidence for the ACE.

Movement Time: Experiment 3

The movement time data (right panel of Figure 5) were analyzed in the same manner as reading times. As shown, no main effects or interactions were observed, all *p*s > .05, including the ACE interaction, $F(1, 35) = .35$, $MSe= 3.31$, $p = .56$, $n^2_p = .01$, suggesting that physical movement times were not influenced by linguistic variables. Because embodied effects on language processing have been used to fuel a theoretical perspective with a stated end-game to unify all of psychological science (Glenberg, 2010), one should expect an important phenomenon to emerge under varied conditions and with varied materials. Thus far, such replication has proven challenging.

Overall, Experiment 3 failed to replicate the modest ACE observed in reading times from Experiment 1, and all experiments thus far have failed to replicate the original ACE reported by Glenberg and Kaschak (2002). Two primary differences between the original study and the current paradigm could potentially explain the lack of replication. First, the current study used different stimulus sentences. Whereas the original study included sentences describing motion in imperative and transfer sentences, Experiments 1-3 only included transfer sentences. Perhaps more salient, the current paradigm involved mouse clicks and movements, whereas the original study used an elongated response-box to encourage arm movements. In order to determine whether the stimuli or the response mechanism may explain the current failures to replicate, Experiment 4 was conducted to more closely mimic the physical response parameters from Glenberg and Kaschak (2002), but still using the materials from Experiments 1-3 (to avoid changing both factors at once).

Experiments 4A and 4B

As noted, the original ACE experiments involved arm movements over a response-box, or at least involved their planning. Although movement times were not reported by Glenberg and Kaschak, it is plausible that planning largearm-movements is more cognitively demanding than planning the mouse-movements in Experiments 1-3. Moving one's index finger from one remote button to another is a fairly precise maneuver, perhaps requiring greater attention than the mouse movements required thus far. Although this explanation is undermined by movement compatibility effects in paradigms using other response procedures (e.g., knob turning, Zwaan & Taylor, 2006), it remains possible that the mouse method of Experiments 1-3 was poorly chosen to reveal the ACE. Alternatively, the failure to observe an effect could simply mean that the ACE is elusive, calling into question its legitimacy as a hallmark EC effect. Given the clear implications for embodiment theory, Experiment 4 was designed to closely approximate the physical mechanics of the original ACE study, without (yet) changing the stimuli. Rather than respond using mouse-movements, computer keyboards were modified using Othello® game pieces to reproduce the characteristics of the response-box used by Glenberg and Kaschak (2002). Although the present studies used a modified keyboard, rather than a custom-built response box, left/right orientation biases for keyboards were mitigated by orienting the keyboard sideways, such that the longer dimension projected outward from participants' bodies. Notably, Borreggine and Kaschak (2006) also used a sideways keyboard to investigate the ACE. The ACE was originally documented in reading times (i.e., the latency to lift one's finger from the starting location); it is unlikely that apparatus differences across studies would produce different latencies to lift one's finger.

Method

Participants—In Experiment 4A, 52 native English speaking, right-handed participants were randomly divided into equal-sized groups based on Response Direction (sensible is near versus far). In Experiment 4B, 22 students participated in both Response Direction conditions, reversing the response mapping mid-way through the experiment.

Procedure—As described by Glenberg and Kaschak (2002, p. 559), their response button box was "…approximately 28 × 18 × 6 cm. The box was held in the lap, with the longest dimension projecting outward from the body. Three critical response buttons were arrayed on the top surface…and they differed in distance from the body: near, middle, and far." To replicate this, a standard Gateway® keyboard was modified with Othello® pieces, and turned such that it sat on the participants' lap with the spacebar on the right. As shown in Figure 6 (from the instructions provided to research assistants), the keyboards had extra raised buttons added to the 'q', 'p', and '9' (on the number pad) keys, which made them easy to locate without looking down. As in Glenberg and Kaschak (2002; Borreggine & Kaschak, 2006), participants initiated each trial by pressing (and holding) the middle (start) key. To make a sensibility decision, participants released the start key (giving a measure of reading time) and pressed either the A or B button, which were designated as 'sensible' and 'nonsense' response options (counterbalanced by participant in Experiment 4A and by block in Experiment 4B). Movement times were recorded as the latency to press the A or B key. The computer monitor displayed the same visual information as Experiments 1-3, with the exception that sentences disappeared when participants released the 'start' key, rather than turn black for participants to manipulate. As in previous experiments, participants completed 20 practice trials prior to beginning the experimental trials. In Experiment 4B, response mapping switched mid-way through the experiment (followed by more practice trials). Response mapping was consistent throughout Experiment 4A.

Results and Discussion

Results were processed in the same manner as prior experiments. In Experiment 4A, an average of 8 trials (6.6%) were lost per participant due to errors, and in Experiment 4B, the average was9 trials (7.5%).

Reading Time: Experiment 4A

Reading times (left panel of Figure 7) were analyzed in a 2 (Response Direction) × 2 (Sentence Direction) mixed-model, RM ANOVA, with Response Direction as the between-subjects variable. There was a main effect of Response Direction, $F(1, 50) = 4.12$, MSe =441.55, $p = .048$, $n^2_p = .08$; participants in the "sensible is near" condition read sentences more quickly (M = 1597 ms, SE = 92) than participants in the "sensible is far" condition (M = 1861 ms, SE = 92). The ACE interaction between Response Direction and Sentence Direction was not reliable, $F(1, 50) = 0.01$, MSe =13.13, $p = .92$, $n^2_p = .00$. Although the physical response parameters were nearly identical to those used in the original ACE experiments, Experiment 4A failed to elicit any evidence of the effect on reading times.

Movement Time: Experiment 4A

Movement times (right panel of Figure 7) were analyzed in the same manner as reading times, and were similarly uninformative. No main effects or interactions were reliable (all $ps > .05$), including the critical ACE interaction, $F(1, 50) = 0.04$, MSe =2.88, $p = .84$, $n^2_p = .001$.

Reading Time: Experiment 4B

Because the original ACE was observed in a within-subjects design, Experiment 4B was conducted to more closely approximate both the response mechanics and the experimental design. Reading times (left panel of Figure 8) were analyzed in a 2 (Response Direction) × 2 (Sentence Direction) within-subjects RM ANOVA. The main effect of Response Direction, $F(1, 21) = 29.16$, MSe = 68.98, $p < .001$, $n^2_p = .58$, was in the opposite direction of the effect observed in Experiment 4A (although the effect size in 4B was larger). Participants were faster to complete the reading task when the "sensible" decision was away from their bodies (M = 1400 ms, SE = 60), relative to near it (M = 1702 ms, SE = 76). There were no other main effects or interactions, including the predicted ACE interaction, $F(1, 21) = 0.88$, MSe = 10.91, $p = .36$, $n^2_p = .04$. In short, even when the response parameters and design more closely mirrored those in the original ACE study, Experiment 4B again yielded no support for cognition-action coupling in sentence processing.

Movement Time: Experiment 4B

As before, movement times (right panel of Figure 8) were examined in parallel with reading times, in another 2 (Response Direction) × 2 (Sentence Direction) within-subjects RM ANOVA. As in Experiment 4A, there were no reliable main effects or interactions, including the ACE interaction ($F(1, 21) = 0.29$, MSe = 2.60, $p = .59$, $n^2_p = .01$), again suggesting that either the stimuli were not appropriate to elicit the ACE, or that physical movements are generally unaffected by linguistic processes.

Because the ACE did not emerge in either Experiment 4A or 4B, despite closely approximating the original study, two main possibilities remain regarding the generality of the effect: On one hand, the results may suggest that the ACE is limited to specific stimuli, or that the current stimuli were somehow inappropriate. This conclusion would carry considerable implications for embodied accounts of language, forcing the question: How meaningful is the ACE if it only arises with certain stimuli? (Examination of the transfer sentences in Appendix A suggests that they were appropriately crafted.) On the other hand, the results may indicate that the effect is elusive, and that existing demonstrations may reflect *publication bias*, the tendency for published reports to unduly favor positive results. As noted by Ferguson and Heene (2012; also de Bruin, Treccani, & Della Sala, 2014) publication bias can arise from many sources (e.g., researchers choosing to emphasize those experiments that worked, reviewers or editors asking for null effects to be removed from manuscripts to maximize “news value,” etc.). In the present case, published ACE findings may represent only a subset of studies, with a larger body of null effects lost in a file drawer (where the current studies could have ended up). Before considering such issues, it is important to establish whether the ACE might replicate when different materials are used. Experiments 5 and 6 were designed to replicate the ACE using the original stimuli, generously provided by Arthur Glenberg.

General Method: Experiments 5 and 6

Experiments 4A and 4B were designed to test whether the mouse response method precluded the observation of an ACE, yet both yielded null effects. Experiments 5 and 6 were designed to test whether the stimuli caused the failure to replicate, while reverting back to the mouse method from Experiments 1-3. Should the ACE be observed, it would suggest that the stimuli in Experiments 1-4 were not well-designed to elicit the congruency effect, and that the ACE can generalize to movement dynamics beyond button-presses.

Stimuli

Stimuli consisted of 160 sentences (80 sensible), from Glenberg and Kaschak (2002). The sensible sentences were comprised of 40 imperative transfer sentences (e.g., “close/open the drawer”), and 20 each of concrete and abstract transfer sentences.

Procedure

The response mechanics were identical to those in Experiment 1, but with 160 trials instead of 240. Unlike Experiment 1, the response mapping (sensible is near/far) was manipulated within-subjects, such that midway through the experiment, the response mapping switched. All participants completed a second round of practice trials following this switch. As in Glenberg and Kaschak (2002), trials were divided into 10 blocks consisting of 16 sentence judgments each (8 nonsense, 4 imperative transfer, 2 concrete transfer, and 2 abstract transfer, with equal representation of toward/away implied transfer directions). Changes from the general procedure are noted on a per-experiment basis.

Experiment 5

In keeping with the original aim of these experiments, Experiment 5 was designed to determine (a) whether the original stimuli from Glenberg and Kaschak (2002) would elicit the ACE, and (b) whether such movement compatibility effects in language processing can be disembodied, using the method from Markman and Brendl (2005) with participants' names on-screen.

Method

Participants—Seventy-one native English-speaking right-handed students participated for partial course credit. By random assignment, 31 participants saw their names at the top of the screen, and 40 saw their names at the bottom of the screen.

Procedure—The mouse response mechanics were identical to Experiment 1. The only on-screen change was that participants' names were resized and “walls” were added to create a stronger illusion of distance for names at the top of the screen (as in Markman & Brendl, 2005; see Figure 9).

Results and Discussion

Outliers were processed in the same manner as previous experiments. On average, 9 trials (11.3%) were dropped for inaccurate sensibility decisions.

Reading Time—Reading times (left panel of Figure 10) were analyzed in a 2 (Sentence Type: Imperative/Transfer) × 2 (Sentence Direction: Away/Toward) × 2 (Response Direction: Away/Toward body) × 2 (Name Location: Near/Far) mixed-model, RM ANOVA, with Name Location as the between-subjects factor. There was a main effect of Sentence Type, $F(1, 69) = 187.09$, $MSe = 201.50$, $p < .001$, $n^2_p = .73$, such that participants were faster to finish reading imperative sentences ($M = 1683$ ms, $SE = 65$), relative to transfer sentences ($M = 2202$ ms, $SE = 69$). This was expected, as imperative sentences are shorter than transfer sentences. Participants also finished reading 113 ms faster when subsequently issuing “toward” responses, relative to “away” responses, $F(1, 69) = 4.97$, $MSe = 361.85$, $p = .03$, $n^2_p = .07$. Unlike previous experiments, there was a reliable three-way interaction between Name Location × Response Direction × Sentence Direction, $F(1, 69) = 4.75$, $MSe = 85.33$, $p = .03$, $n^2_p = .06$, such that the patterns depicted in the left panel of Figure 10 were magnified when participants' names appeared at the bottom of the screen. Contrary to the ACE prediction, the Response Direction × Sentence Direction interaction was again null, $F(1, 69) = 0.15$, $MSe = 85.33$, $p = .70$, $n^2_p = .002$. Although this experiment used the same stimuli as Glenberg and Kaschak (2002), there was still no ACE in reading times.

Movement Time—Movement times (right panel of Figure 10)were analyzed in the same manner as reading times. In contrast to the reading times, in which imperative sentences were read faster than transfer sentences, the movement times revealed the opposite: Participants finished their physical movements more quickly for transfer sentences ($M = 1012$ ms, $SE = 20$), relative to imperative sentences ($M = 1137$ ms, $SE = 30$), $F(1, 69) = 39.26$, $MSe = 55.41$, $p < .001$, $n^2_p = .36$. This may suggest that physical responses were primed by transfer-specific language, which could be taken as tentative support for the embodied view of language processing, although it would not explain why imperative sentences did not similarly prime motion. A more likely interpretation is that the shorter, imperative sentences gave people less time to adequately create motor plans for responding. In either case, there was no interaction between Sentence Direction and Response Direction, $F(1, 69) = 0.27$, $MSe = 25.70$, $p = .87$, $n^2_p = .00$, again contradicting the ACE prediction.

Experiment 6

The null results in Experiments 1, 2, and 5 suggest that attempting to disembody the ACE is not a fruitful avenue. More important, the lack of crossover interactions between implied sentence direction and physical movement direction in Experiments 1 through 5 suggests that reproducing the ACE may also not be a fruitful avenue. (Notably, even a less rigorous potential outcome, such as an ACE-like pattern in half of the design, has only been observed once thus far, for “away” sentences in Experiment 1.) A finding as important as the ACE, with such an elaborate theoretical interpretation, should be robustly observable using procedures that reasonably approximate the original study. For example, mirror effects in recognition memory are observed in various paradigms. The own-race bias in face perception is observed using learning, recognition, and line-ups. Classic cognitive effects, such as Stroop interference and serial-position effects, are easily demonstrated in a classroom, with no instrumentation at all. It seems reasonable to expect that the ACE, which has provided ample behavioral evidence in favor of the EC hypothesis, should also be broadly verifiable. The goal of Experiment 6 was to test, without additional disembodying manipulations, whether the ACE could be observed in the mouse-movement paradigm, using the stimuli from Glenberg and Kaschak (2002).

Method

Participants—Because the ACE had proven so elusive in Experiments 1-5, Experiment 6 used a larger sample size to better detect its presence or absence. Ninety-two native English-speaking, right-handed students participated for partial course credit.

Procedure—The procedure was identical to Experiment 5, but names were not shown on the computer screen.

Results and Discussion

Results were processed in the same manner as prior experiments. On average, 9 trials (11.3%) were dropped per participant because of incorrect sensibility judgments. One person was dropped for having an excessive error rate (> 50%), leaving 91 participants in the final analysis.

Reading Time—The reading times (left panel of Figure 11) were analyzed in a 2 (Sentence Type: Imperative/Transfer) × 2 (Sentence Direction: Away/Toward) × 2 (Response Direction: Away/Toward body) within-subjects, RM ANOVA. As in previous experiments, there was a main effect of Sentence Type, $F(1, 90) = 88.08$, $MSe = 293.43$, $p < .001$, $n^2_p = .50$; participants read the shorter, imperative sentences ($M = 1652$ ms, $SE = 49$)

faster than longer, transfer sentences (M = 2029 ms, SE = 55). There was also an interaction between Sentence Type and Sentence Direction, $F(1, 90) = 6.04$, $MSe = 73.20$, $p = .02$, $n^2_p = .06$. For imperative sentences, participants read “toward” sentences (M = 1613 ms, SE = 49) more quickly than “away” sentences (M = 1690 ms, SE = 52); this pattern was not reliable for the transfer sentences. As shown in Figure 9, there was no hint of the ACE, with a null Sentence Direction × Movement Direction interaction, $F(1, 90) = 0.01$, $MSe = 68.79$, $p = .91$, $n^2_p = .00$. Taken together with the previous results, Experiment 6 suggests that the ACE does not occur in the mouse-movement paradigm. This null result is surprising, given the apparently robust history of the ACE in the EC literature. The mouse-movement paradigm involves physical movements toward and away from the participant's body, which should be capable of eliciting the effect, should one be present.

Movement Time—As before, movement times (right panel of Figure 11) were analyzed in the same manner as the reading times. There was a main effect of Sentence Type, $F(1, 90) = 25.19$, $MSe = 140.22$, $p < .001$, $n^2_p = .22$, which revealed that imperative sentences elicited slower movement times (M = 1133 ms, SE = 39), relative to transfer sentences (M = 994 ms, SE = 20). As before, the ACE would have manifested as an interaction between Response Direction and Sentence Direction, which was not reliable, $F(1, 90) = 0.74$, $MSe = 20.61$, $p = .39$, $n^2_p = .01$. As shown in Figure 11, there was no trend toward such an action-sentence congruency effect.

Experiment 7

Because the ACE was not observable using the mouse-movement paradigm, and given its theoretical importance, Experiment7was designed to assess whether the effect could be observed in a paradigm that very closely approximated Glenberg and Kaschak (2002). The experiment used their original stimuli, and the keyboard modified to approximate their response-box (as before, the keyboard was turned, so participants made toward/away movements, not left/right movements; see also Borreggine & Kaschak, 2006). Although six consecutive null results do not bode well for such a theoretically important effect, failures to replicate can have many explanations, especially when using modified procedures. Experiment 7 restored all the original procedures from Glenberg and Kaschak (2002) as closely as possible. Should the ACE emerge in this experiment, it would at least provide limited support for the EC account of language perception, revealing a relationship between sentence comprehension and motoric action. On the other hand, if the ACE does not emerge in this experiment, it would call into question the reliability of the effect, and whether it should (partly) motivate a radical reframing of cognitive science.

Method

Participants—Fifty-nine native English-speaking, right handed students participated in exchange for partial course credit.

Procedure—Participants completed the button-press version of the experiment described in Experiment 4, now with stimuli from Glenberg and Kaschak (2002). As before, buttons were affixed to the keyboard, which was turned sideways to require toward/away arm movements, as in Glenberg and Kaschak (2002).

Results and Discussion

Results were processed in the same manner as prior experiments. On average, 8 trials (10%) were dropped per participant because of incorrect sensibility judgments.

Reading Time—As the original ACE was observed in reading times, this analysis was expected to reveal a Sentence Direction × Response Direction interaction, with facilitation for congruent trials and inhibition for incongruent trials. Although this would replicate the original ACE, in light of the previous findings, it would suggest that the effect arises only in very unique circumstances. Reading times (left panel of Figure 12) were analyzed in a 3 (Sentence Type: Imperative/Abstract/Concrete) × 2 (Sentence Direction: Away/Toward) × 2 (Response Direction: Away/Toward body) with in-subjects RM ANOVA[7]. Consistent with previous findings, there was a main effect of Sentence Type, $F(2, 116) = 194.08$, $MSe = 82.53$, $p < .001$, $n^2_p = .77$, such that reading times were faster for imperative transfer sentences ($M = 1374$ ms, $SE = 49$), relative to abstract ($M = 1829$ ms, $SE = 62$) and concrete ($M = 1821$ ms, $SE = 62$) sentences, which did not differ from each other. There was a main effect of Response Direction, $F(1, 58) = 21.93$, $MSe = 262.15$, $p < .001$, $n^2_p = .27$; participants were faster to finish reading when their subsequent response movement would be toward their bodies($M = 1585$ ms, $SE = 63$),rather than away ($M = 1765$ ms, $SE = 57$). As in all previous experiments, however, the ACE effect (Response Direction × Sentence Direction interaction) did not occur, $F(1, 58) = 3.68$, $MSe = 79.77$, $p = .06$, $n^2_p = .06$. As before, although planned comparisons suggest an action-sentence congruency effect in half the design (in this case, the “toward” sentences), the opposite pattern was observed in the other half of the design. As noted earlier, observing facilitation in half of the design is not sufficient to conclude that action language is comprehended via motor simulation or resonance. To support the embodiment hypothesis, reading times must show facilitation for action-verb congruency in both directions. In the present case, reading times were always faster for subsequent toward-body motion, irrespective of the implied sentence direction.

Movement Time—As before, movement times (right panel of Figure 12) were analyzed in the same manner as reading times. No main effects or interactions were observed; planned comparisons revealed no trace of the ACE, $F(1, 58) = 0.00$, $MSe = 7.73$, $p = .99$, $n^2_p = .00$.

Experiment 8

One potential criticism of Experiment 7 as a replication attempt is that keyboards may induce an orientation bias: Although the keyboard was turned sideways, with buttons added to three response keys, it was still clearly a keyboard. Experiment 8 was conducted with more profoundly modified keyboards, to more closely approximate the original response apparatus from Glenberg and Kaschak (2002). Although it was not possible to recreate their original response boxes, Experiment 8 represented the most direct replication attempt as possible, using the original ACE stimuli and disguised keyboards as the input device (and a larger sample size to afford greater power).

---

[7] Results do not change when Sentence Type is collapsed into 2 levels, imperative and transfer.

Method

Participants—Eighty-eight native English-speaking, right handed students participated in exchange for partial course credit. Two participants were dropped prior to analysis for failing to follow instructions.

Procedure—The procedure was identical to Experiment 7, with the exception that participants' responses were issued via disguised keyboards. As shown in Figure 13, all keys were removed, leaving only the “start” button and the two response options. The keyboards were covered in black tape, which both created the appearance of a flat surface, and removed the possibility that extraneous keys might affect motor planning. Raised buttons were glued to the response options, and were labeled as in previous experiments.

Results and Discussion

The data were processed in the same manner as prior experiments. On average, fewer than 8 trials (10%) were dropped per participant because of incorrect sensibility judgments.

Reading Time—Reading times (left panel of Figure 14) were analyzed in a 3 (Sentence Type: Imperative/Abstract/Concrete) × 2 (Sentence Direction: Away/Toward) × 2 (Response Direction: Away/Toward body) within-subjects RM ANOVA. As in prior experiments, there was a main effect of Sentence Type, $F(2, 170) = 254.81$, $MSe = 93.36$, $p < .001$, $n^2_p = .75$. Reading times were fastest for imperative transfer sentences ($M = 1344$ ms, $SE = 44$), followed by abstract ($M = 1764$ ms, $SE = 56$) and then concrete ($M = 1828$ ms, $SE = 63$) sentences, with all pairwise comparisons reliably different. There was a main effect of Response Direction, $F(1, 85) = 31.53$, $MSe = 364.95$, $p < .001$, $n^2_p = .27$. Contrary to Experiment 7, participants were faster to finish reading when their subsequent response movements were away from their bodies ($M = 1586$ ms, $SE = 53$),rather than toward their bodies ($M = 1705$ ms, $SE = 56$). Consistent with Experiment 7, however, no ACE interaction (Response Direction × Sentence Direction) was observed, $F(1, 85) = 0.14$, $MSe = 10.68$, $p = .71$, $n^2_p = .002$. As in prior experiments, planned comparisons suggested an action-sentence congruency effect in half the design (in this case, the “away” sentences). However, the opposite pattern was observed in the “toward” sentences (in fact, the reading time facilitation for subsequent away movements was numerically larger for the incongruent, toward sentences). As previously noted, ACE facilitation in half of the design is insufficient evidence for the hypothesis that language comprehension relies on motor simulation or resonance; the ACE must be observed as a crossover interaction. In Experiment 8, reading times were always faster for subsequent away motions, irrespective of implied sentence directions.

Because Experiment 8 was designed as a near-direct replication of the original ACE experiment, “trimmed reading times” were also computed, consistent with the statistical approach taken by Glenberg and Kaschak (2002; see Footnote [5]). These reading times were computed by dropping the first block of trials for both the “move toward” and “move away” conditions, as well as the fastest and slowest reading times in each of the 12 conditions (created by combinations of sentence types and movement compatibility). The trimmed reading times were analyzed in the same manner as the standard reading times, and no ACE interaction was observed, $F(1, 75) = 0.02$, $MSe = 2.65$, $p = .89$, $n^2_p = .000$. It should be noted that this trimming procedure eliminated 10 participants for missing data and ultimately excluded over 40% of the total trials from analysis. These proportions of excluded data were similar to those reported by Glenberg and Kaschak (2002).

Movement Time—Movement times (right panel of Figure 14) were analyzed in the same manner as reading times. There was a main effect of Response Direction, $F(1, 85) = 6.65$, $MSe = 240.58$, $p = .01$, $n^2_p = .07$, which revealed that participants completed away movements ($M = 394$ ms, $SE = 12$) faster than toward movements ($M = 425$ ms, $SE = 13$).No other main effects or interactions were observed, including the critical ACE interaction, $F(1, 85) = 0.95$, $MSe = 7.61$, $p = .33$, $n^2_p = .01$.Although a strong embodied account predicts the effect in movement times (Borreggine & Kaschak, 2006; Zwaan & Taylor, 2006), the present experiments consistently showed no such evidence.

General Discussion

Eight experiments explored extensions and replications of the action-sentence compatibility effect (ACE), first reported by Glenberg and Kaschak (2002) and later reported by others (e.g., Borreggine, & Kaschak, 2006; de Vega & Urrutia, 2011; de Vega et al., 2013; Glenberg et al., 2008; Kaschak & Borreggine, 2008; Taylor & Zwaan, 2008; Zwaan & Taylor, 2006), as evidence in favor of an embodied account of language comprehension. The original motivation was to determine whether higher-order cognitive processes, such as manipulating participants' sense of “self” in the physical environment (as in Markman & Brendl, 2005), would modulate the ACE. Ultimately, repeated failures to observe the ACE motivated more direct replication attempts. Experiments 6 through 8 did not reveal an ACE resembling any of the existing literature, despite using the same stimuli and nearly identical response methods (in Experiments 7 and 8) as the original finding. Although weak congruency effects were observed for transfer sentences in Experiments 1, 7, and 8, no other experiments revealed this effect. Moreover, each time that ACE-consistent evidence was observed in half the sentences, the exact opposite (i.e., ACE-inconsistent evidence) was observed in the other half. Together, the present data motivate the question: How reliable is the ACE?

In recent years, replication attempts have garnered considerable attention in cognitive and social psychology (Brandt et al., 2014). Whereas replications were once seen as a within-lab necessity prior to publication, but not necessarily worth journal pages on their own, recent changes in the scientific community have produced a bigger push for replication, particularly of classic or noteworthy findings (e.g., the Open Science Framework's Reproducibility Project and Registered Replication Reports in *Psychological Science*). In cognitive science, there has been a simultaneous push for an appreciation of action-cognition couplings, in the relatively new field called Embodied Cognition. Embodiment's purported role in language processing has motivated unique theoretical perspectives, including Perceptual Symbol Systems (Barsalou, 2008), the Indexical Hypothesis (Glenberg & Robertson, 1999, 2000), and the Linguistic Focus Hypothesis (Taylor & Zwaan, 2008). Despite differences, each theory holds a common view that language comprehension is rooted in the body and the physical environment, such that comprehending a word or phrase relies heavily on the perceiver's ability to “ground” that phrase in the real world and use the same sensory-motor neural mechanisms necessary for perception. The core assumption of the embodiment hypothesis is that the physical body and environment are critically involved in all cognitive events, such that the format of cognitive processes is essentially sensorimotor (e.g., Glenberg, 2010). This contrasts with a non-embodied hypothesis that cognition and action are representationally distinct, albeit connected through distributed neural networks (see Mahon, 2015a). The aim in this article was not to adjudicate between embodied and non-embodied hypotheses, but to evaluate a hallmark finding that both motivated and accelerated the EC “revolution,” the action-sentence compatibility effect (ACE; see e.g., Horchak et al., 2014).

When critically examining the present study, a natural desire (particularly among EC advocates) may be to request still more attempts to replicate the ACE, perhaps with more participants, a different response mode, etc. Presumably, the rationale for such a request would be framed as follows: “Given all the existing evidence for the ACE, the burden of proof should be conservatively high before its validity is besmirched.” This may be a fair point, although the ACE was given numerous opportunities to emerge in the present study. One might take issue, however, with the antecedent in the foregoing argument, specifically the implication that the prior literature makes a compelling case for the ACE. How valid is this assertion?

Although the current study provided little evidence for the ACE, previous work has revealed the effect, and has described its characteristic time-course. For example, Zwaan and Taylor (2006) found that knob-turning speed was faster while participants processed sentences describing direction-congruent actions (and vise-versa). This facilitation was observed over a longer time-course by temporally extending the action with an adverb, but not an adjective (e.g., *turned the knob nervously*; Taylor & Zwaan, 2008). de Vega, Morena, and Castillo (2013) further investigated the time-course of motor activity during sentence comprehension by manipulating the stimulus onset asynchrony (SOA) between the transfer verb in a sentence (e.g., *gave* in “I gave you the pencil”) and the motion cue, which prompted participants to move their hands toward or away from their bodies. At short verb-action lags (100-200 ms), they observed interference, but at the long lags (350 ms), they observed facilitation (priming). They interpreted these results using a resonance analogy, suggesting that action verbs are processed automatically, yet briefly, in motor areas of the brain (e.g., Bub & Masson, 2010; Taylor & Zwaan, 2008; Zwaan & Taylor, 2006).

Pulvermüller et al. (2005) used magneto encephalography to show that processing action verbs describing the mouth, hand, and leg resulted in motor and premotor activity within 200 ms of word onset. Applied to de Vega et al.'s (2013) data, this would suggest that motor interference in the 100-200 ms SOA conditions reflects an early, automatic processing stage, during which lexical access and response preparation compete for neural resources. When the two responses no longer temporally overlap, as when the action cue is provided 350 ms after the action verb, facilitation is observed. These results may support a mirror neuron explanation (e.g., Chersi et al., 2010) or a multi-stage processing model (de Vega et al., 2008). In either case, they suggest that motor planning interacts with ongoing language processing. Notably, all such research (testing the ACE time-course) presumes that the ACE is robust enough to be replicated and extended. By way of analogy, a researcher may test whether target valence influences the duration of the "attentional blink" during RSVP, having well-founded confidence that the attentional blink is robustly observed across numerous experiments (e.g., Shapiro, Raymond, & Arnell, 1997). The current results suggest that treating the ACE as such a benchmark phenomenon is premature.

Is the ACE robust?

Similar to the ongoing push for more replications in psychological science, there is great current interest regarding data analysis, with many researchers suggesting that traditional null-hypothesis significance testing (NHST) does not suffice (e.g., Cumming, 2013; Kline, 2004; Rouder et al., 2009). Rather, they have argued in favor of Bayesian analyses (e.g., Kruschke, 2010), which take base rates into account when determining whether any given data lend greater support to *either* the null or alternative hypotheses. As argued by Rouder et al. (2009), a key benefit to Bayesian analyses is that the null hypothesis is no longer merely "rejected" or "not rejected," but can actually be supported. Sometimes the logical interpretation of data is that an effect does not exist, and Bayesian analyses allow researchers to make such inferences. Of greater importance to the present article, even when reported results are classified as "significant" using NHST, some Bayesian tests allow researchers to evaluate the strength of the evidence, expressed as an odds ratio (of the null versus alternative hypotheses), rather than merely classifying any given outcome as "significant or not." Such Bayesian analyses often reveal that "significant" findings provide little more than anecdotal evidence for their conclusions (Wagenmakers et al., 2011).

In the present study, all experiments were analyzed using traditional NHST methods, in keeping with the prior ACE literature. The sheer volume of null effects, however, raises a serious question. Specifically, they motivate a reconsideration of the prior literature. Is the existing literature actually convincing? To more closely examine the ACE literature, *Scaled JZS Bayes factor* values were computed (using the calculator available at http://pcl.missouri.edu/bayesfactor) for all the present experiments, as well as numerous prior studies[8]. Although the process is simple, a brief explanation of the steps is helpful. For any reported ACE effect in the literature, either a t-value is provided, or can easily be derived from a reported F-ratio. As a relevant example, in Glenberg and Kaschak (2002, Experiment 1), the ACE interaction was reported as significant, with $F(1,34) = 7.75$. This corresponds to a t-value of 2.78. To apply the online calculator, this t-value and sample size (35) is entered. Finally, one must choose a value for the "scaled-information prior," which partly controls the interpretation of small effects. As suggested by Rouder et al. (2009, p. 233), the scaling value is set as r = 1 as a default, which is typically appropriate.[9] Continuing with the example from Glenberg and Kaschak (2002), with $t = 2.78$, $N = 35$, and r = 1, the resultant

---

[8] Studies were selected for inclusion by clicking the "cited by" link in Google Scholar for the original ACE demonstration (Glenberg & Kaschak, 2002), as well as several closely related follow-up studies (Borreggine & Kaschak, 2006; Kaschak & Borreggine, 2008). Inclusion criteria were simple: Only articles describing reading or movement facilitation (or interference, see Table 3) resulting from action-verb congruency were included.

[9] Rouder et al. (2009) recommend using a scale of r = 1, unless researchers have some prior reason to expect small or large effects. For small and large anticipated effects, r is adjusted downward or upward, respectively. In the present case, the ACE is a relatively new phenomenon, with no *a priori* reason to expect relatively small effects. In addition, because effect sizes could not be computed for several key ACE findings (e.g., Glenberg & Kaschak, 2002; Kaschak & Borreggine, 2008), it would be inappropriate to arbitrarily scale the JZS Bayes Factors. Therefore, the recommended default scale value (r = 1) was selected for all analyses.

Scaled JZS Bayes Factor = 3.97. This value is directly interpretable as an odds ratio: The alternative hypothesis (i.e., that a bona fide ACE was observed) is supported, with 4:1 odds, relative to the null hypothesis. This general approach was followed for many studies in the ACE literature, deriving t-values and computing Scaled JZS Bayes Factors, with r = 1.

Before considering the results, it is important to address the nebulous “strength of evidence” concept. As noted, the odds ratio from Glenberg and Kaschak (2002, Experiment 1), was 4:1 in favor of the alternative hypothesis. How should such a result be evaluated? There is no clear answer to this, although recommendations exist in the literature, mainly stemming from Jeffreys (1961; see Rouder et al., 2009; Morey & Rouder, 2011; Wagenmakers et al., 2011). Briefly, Jeffreys recommended that odds smaller than 3:1 should be considered “anecdotal evidence,” that odds between 3:1 and 10:1 should be considered “substantial evidence,” and that odds greater than 10:1 should be considered “strong evidence.” By this classification system (which is not sacrosanct), the 4:1 finding from Glenberg and Kaschak would be considered “substantial evidence.” For reference, the figures below indicate thresholds for odds ratios equal to 2:1, 3:1, and 7:1. These thresholds do not have particular importance, but were chosen to help readers visually appreciate the evidence strengths that were observed.

With these considerations in mind, JZS Bayes Factors were computed for reports of ACE facilitation (i.e., faster responding when movements and sentences express congruent directions). The results are shown in Figure 15, with further details available in Table 2. In Figure 15, any left-going bars signal support for the null hypothesis (i.e., no ACE), and right-going bars signal support for the alternative hypothesis (i.e., ACE). The magnitudes of the bars indicate the strength of evidence in either direction. Starting from the bottom of the figure, the first ninebars (A through I) represent the present experiments. Note that Experiment 1 provided evidence for the alternative (ACE) hypothesis, but with very weak 2:1 odds. The remaining experiments all provided evidence for the null hypothesis, with six experiments exceeding 7:1 odds, and three of those reaching or exceeding 10:1 odds.

In the remainder of Figure 15 (bars J through AC), JZS Bayes Factor results from other published findings are shown. Several points should be noted: First, the bars marked by asterisks (L, U, and Z) indicate findings that were originally reported as null effects, with p-values ranging from .09 to .15. As shown, these all correspond to JZS Bayes values that favor the null hypothesis, although weakly in every case. Of greater interest are the remaining 17 bars in Figure 15, each corresponding to effects that were originally reported as reliable ACE findings, with $p < .05$, or beyond. Out of these 17 experiments, nine actually show more evidence for the null hypothesis than the alternative, despite their original analyses rejecting the null hypothesis. Of the remaining eight experiments that actually provided evidence for the alternative hypothesis, only four exceed the 3:1 threshold (for “substantial evidence”) and one exceeds the 7:1 threshold. As anironic point, in traditional statistics, many assume that, given alpha = .05, one out of 20 experiments can be expected to generate a false-positive result. Out of the 20 experiments spanning bars I through AB, exactly one produced strong positive results, and even that one did not reach Jeffrey's (1961) 10:1 threshold for “strong evidence.”

On the whole, Figure 15 shows that almost all prior evidence for ACE facilitation is either weakly positive, or actually supports the null hypothesis. Even without considering the present experiments, the ACE is only weakly convincing 20% of the time. At first blush, the series of null effects reported in Experiments 1-8 may appear incompatible with the extensive literature on ACE facilitation, but this first impression is mistaken. When examined with statistics that measure support for either the null or the alternative hypotheses, the current data appear consistent with the literature at large (or at least 80% of it), which suggests that facilitative language-movement congruency effects are the exception, rather than the norm.

Finally, one other point should be mentioned about the results in Figure 15. As already noted, out of 20 previous experiments, only three were reported as null effects, despite the generally weak evidence seen in nearly all of them. In a series of recent articles, Francis (2012; 2013) noted that, when experiments have relatively low power, they tend to produce both false-positives and false-negatives. Thus, when a series of relatively low-power experiments is conducted, we should expect occasional null results, even if the underlying effect is real. In the 20 experiments shown in bars J through AC, 17 (85%) of them were reported as significant, which suggests a pooled, overall power of approximately .85. Given the general weakness of nearly all the results, it strains credulity to imagine that power was so high in these prior studies. In less esoteric terms, there is also a striking visual difference between the present experiments (bars A through I) and all the remaining experiments in Figure 15. What are the odds that the present experiments would produce such strongly null results over and over, whereas the previous literature never had one such occurrence? Such an outcome is nearly impossible, and strongly suggests that publication bias has skewed our understanding of the ACE.

One curious aspect of the ACE is its ability to manifest as either facilitation or interference, depending on task parameters. As summarized by Kaschak and Borreggine (2008, p. 883), "…the ACE occurs when the motor response is executed at an early point in the comprehension of the sentence, disappears for a time, and then reappears when the motor response is executed right before the end of the sentence." In several ACE studies (e.g., Boulenger et al., 2006; Buccino et al., 2005; Nazir et al., 2008; Sato, Mengarelli, Riggio, Gallese, & Buccino, 2008), the compatibility between movement effectors (e.g., hands) and the effectors implied by the action in the sentence (e.g., "He grabbed the bar") has produced interference, rather than facilitation. For example, Buccino et al. (2005) tested participants with sentences implying hand or foot actions, and participants completed a go/no-go task after each sentence. If the sentence implied an action, half of the participants responded by hand-pressing a button; the other half responded with a foot pedal. If no action was implied, participants remained still. Buccino et al. observed interference: When sentences and the required movements referenced the same effector, RTs were slowed, relative to trials wherein the implied and actual effectors differed. Similar results were observed in a different paradigm, in which participants listened to action verbs or nouns while performing arm-reaching movements. Action verbs inhibited arm movements within 200 ms of word onset, but facilitated arm movements when words were fully processed prior to action initiation (Boulenger et al., 2006; Nazir et al., 2008).

The observation of inhibition or facilitation seems to be guided by both timing parameters and task demands. Borreggine and Kaschak (2006) observed facilitation from action-sentence compatibility when movement directions were cued before the onset of sentences for processing. This facilitation disappeared when motion was cued 50-ms *after* the sentence, and did not return (orreverse) when cues were delayed 500- or 1000-ms post-sentence. Later, Kaschak and Borreggine (2008) observed facilitation with a 500-ms SOA, and suggested that task demands explained the discrepancy. Whereas participants in the 2006 study made sensibility judgments, participants in the 2008 study determined whether pictures were congruent with preceding sentences, thus requiring sentences to be kept in working memory. The congruence effect in the 2008 study is consistent with the *linguistic focus hypothesis* (LFH) proposed by Taylor and Zwaan (2008). According to the LFH, motor resonance should be observed for linguistic content that is still in current focus, such as when an adverb modifies a preceding verb (e.g., "…turned the knob *slowly*…"). Once processing moves on to later sentence components, this facilitation should disappear. Inconsistent results, however, have also been observed. For example, de Vega et al. (2011) observed interference when movements and verbs implied the same action at short (100-200 ms) verb-cue SOAs, but facilitation when the SOA was extended to 350 ms. They interpreted this temporal pattern as reflecting competition for neural resources, suggesting that verbs and physical responses automatically activate overlapping sensory-motor regions of the brain, and this competition causes interference (see also Pulvermüller et al., 2005). Subsequent work supported this conclusion; decisions to factual and counterfactual sentences showed interference at short delays, and facilitation at longer delays (de Vega, Moreno, & Castillo, 2013; de Vega & Urrutia, 2011).

The empirical inconsistency regarding the time course of the ACE, and whether it will materialize as facilitation or interference, remains unresolved. Whereas some (Taylor & Zwaan, 2008; Zwaan & Taylor, 2006) have observed facilitation when the verb and the action cue appear in close temporal proximity, others have observed interference (e.g., de Vega et al., 2011; de Vega & Urrutia, 2011). From a broad, EC perspective, either alternative is possible, but differences in materials (e.g., judging sentences versus words), timing parameters, and tasks (e.g., sensibility judgments, lexical decisions, etc.) render true comparison of the effects challenging. Rather than attempt to reconcile the disparities, scaled JZS Bayes factors for inhibitory ACE evidence (see Figure 16 and Table 3) were computed in the same manner as before. None of the present experiments are represented in Figure 16; all the represented studies are from prior literature (see Footnote [8] for selection criteria).

Examining Figure 16, the alternative hypothesis (i.e., evidence for language-motor interference) appears largely supported. Of the 14 included experiments, nine (64%) show evidence in favor of inhibitory effects, with five strong effects. As with the ACE, language-motor interference effects have been cited as strong support for embodied language processing, such that action verbs automatically activate effector-specific motor areas during comprehension (Barsalou, 1999; Gallese & Lakoff, 2005; Glenberg, 1997; Glenberg & Robertson, 2002; Pulvermüller, 2002; Zwaan & Taylor, 2006). In fact, Sato et al. (2008) suggested that this recruitment is a necessary step for language comprehension, not merely an epiphenomenal accompaniment, because they observed verb-motor interference during a semantic task, but not during lexical decision.

Upon closer examination of the data, however, the apparent support for language-motor interference is weaker than it initially appears. To construct Figures 15 and 16, an inclusive approach was taken, with a goal to retain as many relevant studies as possible, without seeking reasons to disqualify any. Nevertheless, the results from Sato et al. (2008; bars G and H in Figure 16) demanded a closer look. Together, Figures 15 and 16 show results from 43 different experiments: For 41 of them, the JZS Bayes factors range from approximately 1.5 to 14. The two experiments from Sato et al. (2008) elicited JZS Bayes factors of 408 and 51,071, vastly outside the range of all other experiments. When the procedures from that experiment are examined, however, a design flaw is easily appreciated, making the experiment uninterpretable.

Stated briefly, in Sato et al. (2008), participants either read or heard a series of Italian verbs, expressing hand-related actions (e.g., “applaud”), foot-related actions (e.g., “walk”) or non-actions (abstract verbs, e.g., “love”). In the key conditions of the experiment, participants performed a go/no-go task: For action verbs, they pressed a button with the right index finger as quickly as possible; for abstract verbs, they withheld responding. (Other conditions involved waiting an extra second for the “go” signal, or performing lexical decision.) In all experiments, only 10 hand-related and 10 foot-related verbs were presented, which means one or two challenging items could dramatically affect RTs. The principal finding was that RTs were slower to hand-related verbs, relative to foot-related verbs, which Sato et al. interpreted as linguistic-motor interference, arguing that processing hand-related verbs selectively interacted with hand-related motor planning.

The problem with Sato et al.'s (2008) interpretation is that their experiment only included half of the necessary design. The experiment did not show linguistic-motor interference; it only showed that people were slower to classify hand-related verbs, relative to foot-related verbs. In order to draw firm conclusions, Sato et al. needed a crossover interaction, including a condition where in participants made responses to the same materials, but using a foot-pedal. If the embodied hypothesis were correct, such a condition would flip the results, now showing slower RTs for foot-related verbs. Because this condition was not included, Sato et al.'s results are not interpretable. (It bears mentioning that a previous study conducted in the same laboratory, Buccino et al., 2005, included both hand- and foot-response conditions.) Returning to Figure 16, if the two outlier results (bars G and H) are excluded from consideration, there remain five experiments supporting the null hypothesis, three with “substantial” evidence, and seven experiments supporting the alternative hypothesis, again with three showing “substantial” evidence. On balance, it seems premature to draw strong conclusions about language-motor interference.

Concluding Remarks

The present study showed that, across eight experiments, the action-sentence compatibility effect (ACE), a key finding supporting the embodiment hypothesis for language comprehension, could not be replicated. It did not emerge in a mouse-movement paradigm with new sentences, or with the original sentences from Glenberg and Kaschak (2002), nor did it emerge in a paradigm that closely matched the original study. Moreover, examination of the prior literature reveals that reported findings of language-motor interactions are far weaker than the published evidence seems to suggest. (The strength or weakness of non-reported findings is anyone's guess.)

Upon closely evaluating the ACE literature, the observed pattern of null and weak results is quite surprising, because various investigators have afforded themselves tremendous latitude to find reliable effects. Across studies, language-motor interactions are inferred both from facilitated responses and from inhibited responses. When effects vacillate across conditions, they are attributed to complex patterns of resonance, waxing and waning between language and motor planning (e.g., Kaschak & Borreggine, 2008, Nazir et el., 2008; Taylor & Zwaan, 2008). Across studies, the results that validate embodied theories of language are apparently bidirectional. The core hypothesis is that *"language processing is inherently linked to motor system activity*,"suggesting that the ACE arises from patterns of sensorimotor activity. When the sensorimotor representation for language is active simultaneously with sensorimotor action planning, interesting effects emerge. As a theoretical position, this is perfectly sound, but what is the proper null hypothesis? In practice, the prediction is that both facilitation and interference indicate embodied language processing. This only leaves a null hypothesis that *"there is no relationship between language processes and the motor system."* This null hypothesis, which would not even be predicted by non-embodied theories of language processing, requires results in a narrow band of values centered at zero, a band that will shrink as sample sizes increase. It is essentially a straw-man hypothesis, rendered even less meaningful when observed values equivalent to zero can be readily explained.

If a theory is flexible enough to explain any effect that differs from zero, it has almost no meaning. And yet, even with such unconstrained theorizing, the predicted ACE effects rarely differ convincingly from zero. Looking back at Figures15 and 16 again, it is clear that the *published* evidence is not compelling for either ACE facilitation or inhibition. Bear in mind, most of the depicted experiments were claimed as support for embodied language processing. Nevertheless, roughly half the time, the available evidence supports the null hypothesis, and positive results rarely provide "strong" evidence. As a matter of practice, the strength of theoretical claims should align with the strength of empirical evidence. For an embodied hypothesis of language processing, that claim would be, "Under certain circumstances, and in certain linguistic tasks, comprehending an effector-specific verb while simultaneously using that effector will produce interference, but will sometimes produce facilitation. It depends."

Acknowledgments

Author Notes: Support provided by NIH / NICHD grant R01 HD075800-01. I am grateful to Stephen Goldinger for helpful discussions, and comments on earlier drafts. I also thank Anthony Barnhart of providing comments, and Arthur Glenberg for generously sharing his stimuli. A special thanks to Melissa Miola, for ensuring that all research assistants followed the same protocols. Thank you to Tresa Marchi, Kirsten Keane, Ken Myers, Shawna Peterson, Jay Todorow, Joel Chabrier, Laura Berman, Megan Trog, Jessie Moecia, Kelsey Grodner, Hannah Messick, Ryan Shepard, Suhani Mehrotra, Rachelle Friedman, Anne Kuni, Juan D. Guevara Pinto, Eduardo Hernandez, Diana Rose, Amelia-Kate Johnson, Sophie Johnson, Julian Bernard, Kimberly Perry, and Biannca Pierre for assistance with data collection.

Appendix A: Sensible Transfer Sentences

You sold the land to [Random Name].	… sold the land to you.
You dedicated the song to [Random Name].	… dedicated a song to you.
You sang [Random Name] a song.	… sang you a song.
You pitched [Random Name] the idea.	… pitched the idea to you.
You paid [Random Name] tribute.	… paid tribute to you.
You gave [Random Name] some writing tips.	… gave you writing tips.
You confessed your secret to [Random Name].	… confessed a secret to you.
You lavished [Random Name] with praise.	… lavished you with praise.
You received the complaint from [Random Name].	… complained to you.
You devoted your time to [Random Name].	… devoted time to you.
You blew [Random Name] a kiss.	… blew you a kiss.
You transmitted the order to [Random Name].	… transmitted orders to you.
You gave [Random Name] another chance.	… gave you another chance.
You told [Random Name] the story.	… told you the story.
You transferred responsibility to [Random Name].	… transferred responsibility to you.
You sent [Random Name] your regards.	… sent you regards.
You gave [Random Name] a piece of your mind.	… gave you a piece of [his/her] mind.
You bestowed the honor upon [Random Name].	… bestowed the honor upon you.
You radioed the message to [Random Name].	… radioed the message to you.
You conveyed the message to [Random Name].	… conveyed the message to you.
You threw [Random Name] the pen.	… threw you a pen.
You awarded a medal to [Random Name].	… awarded the medal to you.
You kicked [Random Name] the soccer ball.	… kicked you the soccer ball.
You bought [Random Name] ice cream.	… bought you ice cream.
You slid [Random Name] the cafeteria tray.	… slid you the cafeteria tray.
You handed [Random Name] the notebook.	… handed you the notebook.
You dealt the cards to [Random Name].	… dealt the cards to you.
You forked over the cash to [Random Name]	… forked over the cash to you.
You donated money to [Random Name].	… donated money to you.
You shot [Random Name] the rubberband.	… shot you the rubberband.
You poured [Random Name] some water.	… poured you some water.
You rolled [Random Name] the marble.	… rolled you the marble.
You slipped [Random Name] a note.	… slipped you a note.
You kicked the football to [Random Name].	… kicked the football to you.
You entrusted the key to [Random Name].	… entrusted the key to you.
You delivered the pizza to [Random Name].	… delivered the pizza to you.
You handed the puppy to [Random Name].	… handed you the puppy.
You drove the car to [Random Name].	… drove the car to you.
You hit [Random Name] the baseball.	… hit you the baseball.
You dispensed the rations to [Random Name].	… dispensed the rations to you.

Appendix B: Sensible No-Transfer Sentences

[Random Name] talked about pizza with you.
[Random Name] combed the puppy with you.
You and [Random Name] took the lesson.
[Random Name] felt honored to be with you.
[Random Name] counted the rations with you.
[Random Name] heard the radioed message with you.
[Random Name] played baseball with you.
[Random Name] listened to the message with you.
You and [Random Name] regarded the problem.
[Random Name] washed the car with you.
[Random Name] sang the song with you.
You bought the soccer ball with [Random Name].
You ate ice cream with [Random Name].
[Random Name] thought about your idea.
You and [Random Name] sang a song.
[Random Name] watched the tribute with you.
[Random Name] looked for the pen with you.
You and [Random Name] discussed some writing tips.
You cleaned the medal for [Random Name].
[Random Name] looked at the land with you.
You discussed shooting rubberbands with [Random Name].
You played cards with [Random Name].
You and [Random Name] shared a secret.
[Random Name] drank some water with you.
[Random Name] and you shared the notebook.
You read about donated money with [Random Name].
[Random Name] held the cafeteria tray with you.
You discuss the cash with [Random Name].
You heard the praise with [Random Name].
You complained about [Random Name].
[Random Name] took a chance with you.
[Random Name] ducked responsibility with you.
You spent your time around [Random Name].
[Random Name] read the story about you.
You found the key with [Random Name].
[Random Name] watched football with you.
[Random Name] looked for the marble with you.
[Random Name] read the orders with you.
[Random Name] watched people blow kisses with you.
[Random Name] wrote the note with you.

Appendix C: Nonsense Sentences

You barked the football to [Random Name].
You kissed the time to [Random Name].
You loafed the coffee cup to [Random Name].
[Random Name] cleaned honor upon you.
You held the chance to [Random Name].
You harpooned [Random Name] the sheet.
You flew on the note to [Random Name].
[Random Name] pickled praise on you.
[Random Name] rolled you adoration.
You perfumed [Random Name] accolades.
You hoarded your love to [Random Name].
You heated [Random Name] the blame.
You tasted the papers to [Random Name].
You tickled the orders with [Random Name].
You swam [Random Name] the truth.
You blew a car to [Random Name].
[Random Name] festered relief to you.
You glued the story with [Random Name].
[Random Name] washed you the thought.
[Random Name] medicated commands to you.
[Random Name] smiled the key to you.
[Random Name] faltered the obligations to you.
You parked the memo to [Random Name].
You sang the marble with [Random Name].
You sneezed [Random Name] secrets.
You cleaned responsibility with [Random Name].
[Random Name] tossed you with the paintball.
[Random Name] shambled you loyalty.
[Random Name] blanketed you the chance.
[Random Name] posted you the flowers.
You flexed [Random Name] a moment.
You cleaned the honor for [Random Name].
You sing the pizza to [Random Name].
You ingested the car with [Random Name].
You snored justice for [Random Name].
You carpeted directions to [Random Name].
You cleaned the pizza to [Random Name].
You gargled integrity to [Random Name].
You mowed [Random Name] your opinion.
[Random Name] ingested you instructions.
[Random Name] mingled the complaint to you.
[Random Name] flew the house with you.

[Random Name] snored the frame with you.
You bordered [Random Name] the chain.
You fell the message to [Random Name].
[Random Name] flushed you the appeal.
[Random Name] choked the lesson with you.
[Random Name] jousted you the marker.
[Random Name] wedged an homage to you.
You retaliated [Random Name] the opportunity.
[Random Name] paddled fairness to you.
[Random Name] hanged you more time.
You radioed the floor with [Random Name].
[Random Name] scratched the hat to you.
[Random Name] ate the regards to you.
You genuflected [Random Name] orders.
You bled the rations to [Random Name].
You filed the box to [Random Name].
You drank the baseball with [Random Name].
You cleaned the puppy to [Random Name].
You sunk [Random Name] the monitor.
You cooked [Random Name] duties.
You drank your idea to [Random Name].
[Random Name] trudged you the concept.
You hanged honesty with [Random Name].
You laughed the pen to [Random Name].
[Random Name] parted you the trailer.
[Random Name] frowned the door for you.
You danced the land to [Random Name].
[Random Name] bit the message to you.
[Random Name] tasted the soccer ball to you.
[Random Name] parked the string to you.
You smelled the song with [Random Name].
[Random Name] rehearsed the medal to you.
[Random Name] locked the lint roller to you.
You sunk [Random Name] your viewpoint.
[Random Name] joked the ice cream to you.
[Random Name] drank the shovel with you.
You floundered the train to [Random Name].
[Random Name] broke the writing tips on you.
You forked the bottle to [Random Name].
[Random Name] laundered the bench to you.
You licked the jacket to [Random Name].
[Random Name] bruised you the hamburger.

You looted the blanket to [Random Name].	[Random Name] stapled the song with you.
[Random Name] droned you the pretzel.	[Random Name] slept you responsibilities.
You cried the jack to [Random Name].	[Random Name] tackled the cloud to you.
You harvested [Random Name] your idea.	[Random Name] threw you the street.
You built the water with [Random Name].	[Random Name] pounded you the sunglasses.
You swam [Random Name] the truth.	You slid the tribute on [Random Name].
You smoked the cafeteria tray to [Random Name].	[Random Name] matched the phone to you.
[Random Name] pickled the secret to you.	You drank [Random Name] your regards.
[Random Name] smelled the money to you.	You roasted [Random Name] the notion.
You pocketed the bike to [Random Name].	[Random Name] watered you a tribute.
You painted the hammer to [Random Name].	You fertilized [Random Name] advice.
[Random Name] swam the cards with you.	[Random Name] smelled the pencil to you.
[Random Name] ate you the poem.	You blanketed [Random Name] to the store.
You dove the rubberband with [Random Name].	[Random Name] dances the cash to you.
You drank the notebook to [Random Name].	You tore [Random Name] the shoe.
You boxed [Random Name] the keyboard.	[Random Name] sniffed the praise from you.
[Random Name] boiled information to you.	[Random Name] recited the mouse for you.
You sang the computer with [Random Name].	You shrugged [Random Name] the hard drive.

References

Alter AL, Oppenheimer DM, Epley N, Eyre RN. Overcoming intuition: Metacognitive difficulty activates analytic reasoning. Journal of Experimental Psychology: General. 2007; 136:569–576.10.1037/0096-3445.136.4.569 [PubMed: 17999571]

Barr DJ, Levy R, Scheepers C, Tily HJ. Random effects structure for confirmatory hypothesis testing: Keep it maximal. Journal of Memory and Language. 2013; 68(3) 10.1016/j.jml.2012.11.001. 10.1016/j.jml.2012.11.001

Barsalou LW. Perceptual symbol systems. Behavioral and Brain Sciences. 1999; 22:577–609.10.1017/S0140525X99532147 [PubMed: 11301525]

Barsalou LW. Grounded cognition. Annual Review of Psychology. 2008; 59:617–645.10.1146/annurev.psych.59.103006.093639

Borreggine KL, Kaschak MP. The action-sentence compatibility effect: It's all in the timing. Cognitive Science. 2006; 30:1097–1112.10.1207/s15516709cog0000_91 [PubMed: 21702848]

Bub DN, Masson ME. Grasping beer mugs: on the dynamics of alignment effects induced by handled objects. Journal of Experimental Psychology: Human Perception and Performance. 2010; 36(2): 341–358.10.1037/a0017606 [PubMed: 20364923]

Brandt MJ, IJzerman H, Dijksterhuis A, Farach FJ, Geller J, Giner-Sorolla R, Grange JA, Perugini M, Spies JR, Veer A. The replication recipe: What makes for a convincing replication? Journal of Experimental Social Psychology. 2014; 50:217–224.10.1016/j.jesp.2013.10.005

Buccino G, Riggio L, Melli G, Binkofski F, Gallese V, Rizzolatti G. Listening to action-related sentences modulates the activity of the motor system: A combined TMS and behavioral study. Cognitive Brain Research. 2005; 24:355–363.10.1016/j.cogbrainres.2005.02.020 [PubMed: 16099349]

Cacioppo JT, Priester JR, Bernston GG. Rudimentary determinants of attitudes: II. Armflexion and extension have different effects on attitudes. Journal of Personality and Social Psychology. 1993; 65:5–17.10.1037/0022-3514.65.1.5 [PubMed: 8355142]

Chemero, A. Radical embodied cognitive science. MIT press; 2011.

Chen M, Bargh JA. Consequences of automatic evaluation: Immediate behavioral predispositions to approach or avoid the stimulus. Personality and Social Psychology Bulletin. 1999; 25(2):215–224.10.1177/0146167299025002007

Chersi F, Thill S, Ziemke T, Borghi AM. Sentence processing: linking language to motor chains. Frontiers in Neurorobotics. 2010; 4:4.10.3389/fnbot.2010.00004 [PubMed: 20725506]

Cumming G. The new statistics: Why and how. Psychological Science. 2013; 25:7–29.10.1177/0956797613504966 [PubMed: 24220629]

de Bruin A, Treccani B, Della Salla S. Cognitive advantage in bilingualism: An example of publication bias? Psychological Science. 2014 doi:0956797614557866.

de Vega M, Urrutia M. Counterfactual sentences activate embodied meaning: An action–sentence compatibility effect study. Journal of Cognitive Psychology. 2011; 23(8):962–973.10.1080/20445911.2011.590471

de Vega M, Moreno V, Castillo D. The comprehension of action-related sentences may cause interference rather than facilitation on matching actions. Psychological Research. 2013; 77:20–30.10.1007/s00426-011-0356-1 [PubMed: 21755368]

Faul F, Erdfelder E, Lang AG, Buchner A. G*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. Behavior Research Methods. 2007; 39:175–191. [PubMed: 17695343]

Ferguson CJ, Heene M. A vast graveyard of undead theories publication bias and psychological science's aversion to the null. Perspectives on Psychological Science. 2012; 7(6):555–561.10.1177/1745691612459059 [PubMed: 26168112]

Fitts PM, Seeger S. S-R compatibility: Spatial characteristics of stimulus and response codes. Journal of Experimental Psychology. 1953; 46:199–210. [PubMed: 13084867]

Francis G. Too good to be true: Publication bias in two prominent studies from experimental psychology. Psychonomic Bulletin & Review. 2012; 19:151–156.10.3758/s13423-012-0227-9 [PubMed: 22351589]

Francis G. Replication, statistical consistency, and publication bias. Journal of Mathematical Psychology. 2013; 57:153–169.

Freeman JB, Ambady N. Motions of the hand expose the partial and parallel activation of stereotypes. Psychological Science. 2009; 20(10):1183–1188.10.1111/j.1467-9280.2009.02422.x [PubMed: 19686295]

Gallese V, Lakoff G. The brain's concepts: The role of the sensory-motor system in conceptual knowledge. Cognitive Neuropsychology. 2005; 22(3-4):455–479.10.1080/02643290442000310 [PubMed: 21038261]

Glenberg AM. What memory is for: Creating meaning in the service of action. Behavioral and Brain Sciences. 1997; 20:41–50.10.1017/S0140525X97470012

Glenberg AM. Embodiment as a unifying perspective for psychology. Wiley Interdisciplinary Reviews: Cognitive Science. 2010; 1:586–596.10.1002/wcs.55 [PubMed: 26271505]

Glenberg AM. Few believe the world is flat: How embodiment is changing the scientific understanding of cognition. Canadian Journal of Experimental Psychology. 2015; 69:165–171. [PubMed: 26010024]

Glenberg AM, Kaschak MP. Grounding language in action. Psychonomic Bulletin & Review. 2002; 9:558–565.10.3758/BF03196313 [PubMed: 12412897]

Glenberg AM, Robertson DA. Indexical understanding of instructions. Discourse Processes. 1999; 28:1–26.10.1080/01638539909545067

Glenberg AM, Robertson DA. Symbol grounding and meaning: A comparison of high-dimensional and embodied theories of meaning. Journal of Memory & Language. 2000; 43:379–401.10.1006/jmla.2000.2714

Glenberg AM, Sato M, Cattaneo L. Use-induced motor plasticity affects the processing of abstract and concrete language. Current Biology. 2008; 18:R290–R291.10.1016/j.cub.2008.02.036 [PubMed: 18397734]

Glenberg AM, Sato M, Cattaneo L, Riggio L, Palumbo D, Buccino G. Processing abstract language modulates motor system activity. The Quarterly Journal of Experimental Psychology. 2008; 61:905–919.10.1080/17470210701625550 [PubMed: 18470821]

Glenberg AM, Witt JK, Metcalfe J. From the revolution to embodiment: 25 years of cognitive psychology. Perspectives on Psychological Science. 2013; 8:573–585.10.1177/1745691613498098 [PubMed: 26173215]

Harnad S. The symbol grounding problem. Pyusica D. 1990; 42:335–346.

Hickok G. Eight problems for the mirror neuron theory of action understanding in monkeys and humans. Journal of Cognitive Neuroscience. 2009; 21(7):1229–1243.10.1162/jocn.2009.21189 [PubMed: 19199415]

Horchak OV, Giger JC, Cabral M, Pochwatko G. From demonstration to theory in embodied language comprehension: A review. Cognitive Systems Research. 2014; 29–30:66–85. http://dx.doi.org/10.1016/j.cogsys.2013.09.002. [PubMed: 19594449]

Jeffreys, H. Theory of probability. 3rd. Oxford: Oxford University Press; 1961.

Kaschak MP, Borreggine KL. Temporal dynamics of the action-sentence compatibility effect. The Quarterly Journal of Experimental Psychology. 2008; 61:883–895.10.1080/17470210701623852 [PubMed: 18470819]

Kline, RB. Beyond significance testing: Reforming data analysis methods in behavioral research. Washington, DC: APA Books; 2004.

Kontra, Lyons, Fischer, Beilock. Physical experience enhances science learning. Psychological Science. 2015; 26:737–749.10.1177/0956797615569355 [PubMed: 25911125]

Kruschke, JK. Doing Bayesian data analysis: A tutorial with R and BUGS. New York, NY: Academic Press; 2010.

Lakoff, G. Women, fire, and dangerous things: What categories reveal about the mind. Chicago: University of Chicago Press; 1987.

Mahon BZ. The burden of embodied cognition. Canadian Journal of Experimental Psychology. 2015a; 69:172–178. [PubMed: 26010025]

Mahon BZ. Response to Glenberg: Conceptual content does not constrain the representational format of concepts. Canadian Journal of Experimental Psychology. 2015b; 69:179–180. [PubMed: 26010026]

Mahon BZ, Caramazza A. A critical look at the embodied cognition hypothesis and a new proposal for grounding conceptual content. Journal of Physiology. 2008; 102:59–70.10.1016/j.jphysparis.2008.03.004 [PubMed: 18448316]

Markman AB, Brendl CM. Constraining theories of embodied cognition. Psychological Science. 2005; 16:6–10.10.1111/j.0956-7976.2005.00772.x [PubMed: 15660844]

Morey RD, Rouder JN. Bayes factor approaches for testing interval null hypotheses. Psychological Methods. 2011; 16(4):406.10.1037/a0024377 [PubMed: 21787084]

Nazir TA, Boulenger V, Roy A, Silber B, Jeannerod M, Paulignan Y. Language-induced motor perturbations during the execution of a reaching movement. The Quarterly Journal of Experimental Psychology. 2008; 61:933–943.10.1080/17470210701625667 [PubMed: 18470823]

Papesh MH, Goldinger SD. Memory in motion: Movement dynamics reveal memory strength. Psychonomic Bulletin & Review. 2012; 19:906–913.10.3758/s13423-012-0281-3 [PubMed: 22711579]

Proffitt DR. Embodied perception and the economy of action. Perspectives on Psychological Science. 2006; 1(2):110–122.10.1111/j.1745-6916.2006.00008.x [PubMed: 26151466]

Pulvermüller F. Words in the brain's language. Behavioral and Brain Sciences. 1999; 22:253–279.10.1017/S0140525X9900182X [PubMed: 11301524]

Pulvermüller F, Hauk O, Nikulin VV, Ilmoniemi RJ. Functional links between motor and language systems. European Journal of Neuroscience. 2005; 21:793–797.10.1111/j.1460-9568.2005.03900.x [PubMed: 15733097]

Rouder JN, Speckman PL, Sun D, Morey RD, Iverson G. Bayesian t tests for accepting and rejecting the null hypothesis. Psychonomic Bulletin & Review. 2009; 16(2):225–237.10.3758/PBR.16.2.225 [PubMed: 19293088]

Sato M, Mengarelli M, Riggio L, Gallese V, Buccino G. Task related modulation of the motor system during language processing. Brain and Language. 2008; 105:83–90.10.1016/j.bandl.2007.10.001 [PubMed: 18054379]

Searle JR. Minds, brains, and programs. Behavioral & Brain Sciences. 1980; 3:417–424.10.1017/S0140525X00005756

Shapiro KL, Raymond JE, Arnell KM. The attentional blink. Trends in Cognitive Sciences. 1997; 1(8): 291–296.10.1016/S1364-6613(97)01094-2 [PubMed: 21223931]

Shapiro, LA. Embodied cognition. London: Routledge; 2011.

Solarz AK. Latency of instrumental responses as a function of compatibility with the meaning of eliciting verbal signs. Journal of Experimental Psychology. 1960; 59(4):239–245.10.1037/h0047274 [PubMed: 13832584]

Soliman T, Gibson A, Glenberg AM. Sensory motor mechanisms unify psychology: The embodiment of culture. Frontiers in Psychology. 2013; 4:885.10.3389/fpsyg.2013.00885 [PubMed: 24348439]

Spivey MJ, Grosjean M, Knoblich G. Continuous attraction toward phonological competitors. Proceedings of the National Academy of Sciences of the United States of America. 2005; 102:10393–10398.10.1073/pnas.0503903102 [PubMed: 15985550]

Stepper S, Strack F. Proprioceptive determinants of emotional and nonemotional feelings. Journal of Personality and Social Psychology. 1993; 64:211–220.10.1037/0022-3514.64.2.211

Strack F, Martin LL, Stepper S. Inhibiting and facilitating conditions of the human smile: A nonobtrusive test of the facial feedback hypothesis. Journal of Personality and Social Psychology. 1988; 54(5):768.10.1037/0022-3514.54.5.768 [PubMed: 3379579]

Strack F, Neumann R. Furrowing the brow may undermine perceived fame: The role of facial feedback in judgments of celebrity. Personality and Social Psychology Bulletin. 2000; 26:762–768.10.1177/0146167200269002

Tamir M, Robinson MD, Clore GL, Martin LL, Whitaker DJ. Are we puppets on a string? The contextual meaning of unconscious expressive cues. Personality and Social Psychology Bulletin. 2004; 30:237–249.10.1177/0146167203259934 [PubMed: 15030636]

Taylor LJ, Zwaan RA. Motor resonance and linguistic focus. The Quarterly Journal of Experimental Psychology. 2008; 61(6):896–904.10.1080/17470210701625519 [PubMed: 18470820]

Tucker M, Ellis R. The potentiation of grasp types during visual object categorization. Visual cognition. 2001; 8(6):769–800.10.1080/13506280042000144

Wagenmakers EJ, Wetzels R, Borsboom D, van der Maas HLJ. Why psychologists must change the way they analyze their data: The case of psi: Comment on Bem (2011). Journal of Personality and Social Psychology. 2011; 100:426–432. [PubMed: 21280965]

Weiskopf DA. Embodied cognition and linguistic comprehension. Studies In History and Philosophy of Science Part A. 2010; 41(3):294–304.10.1016/j.shpsa.2010.07.005

Wentura D, Rothermund K, Bak P. Automatic vigilance: the attention-grabbing power of approach- and avoidance-related social information. Journal of Personality and Social Psychology. 2000; 78(6):1024.10.1037/0022-3514.78.6.1024 [PubMed: 10870906]

Wilson AD, Golonka S. Embodied cognition is not what you think it is. Frontiers in Psychology. 2013; 4(58):1–13.10.3389/fpsyg.2013.00058 [PubMed: 23382719]

Wilson M. Six views of embodied cognition. Psychonomic Bulletin & Review. 2002; 9:625–636.10.3758/BF03196322 [PubMed: 12613670]

Winer, BJ. Statistical Principles in Experimental Design. 2nd. New York: McGraw-Hill; 1971.

Zwaan RA, Taylor LJ. Seeing, acting, understanding: Motor resonance in language comprehension. Journal of Experimental Psychology: General. 2006; 135:1–11.10.1037/0096-3445.135.1.1 [PubMed: 16478313]

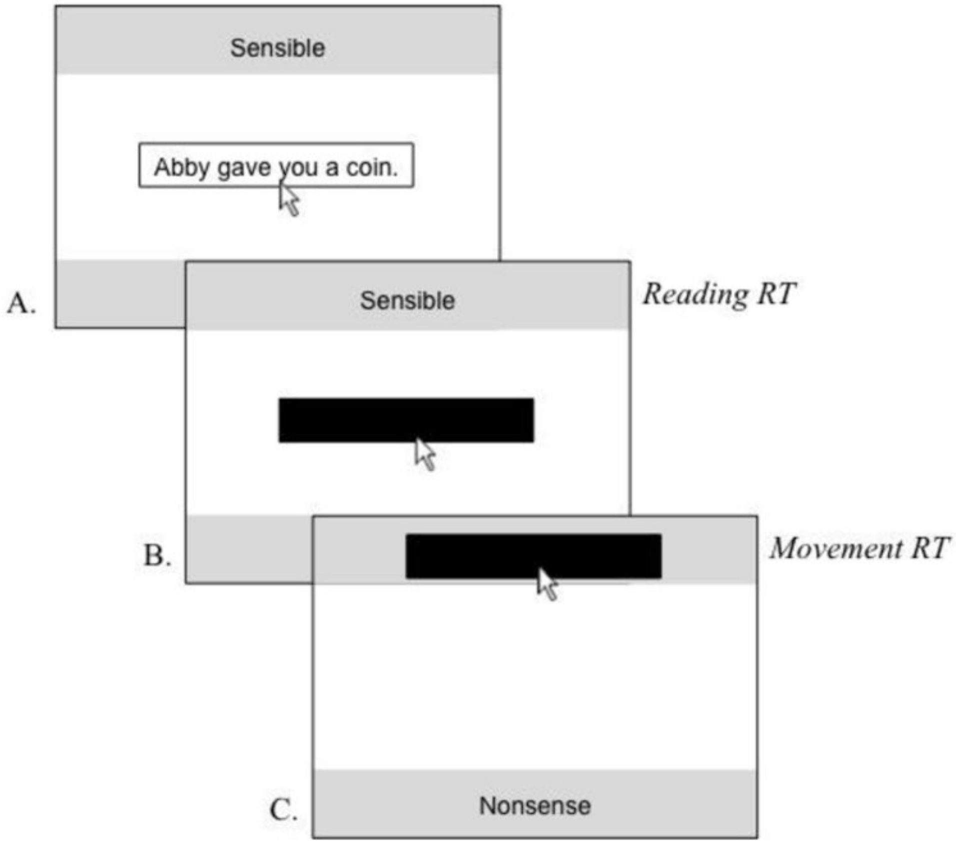

Figure 1.

Schematic outline of the mouse method used in Experiments 1 – 3, and Experiments 5 and 6. (A) Self-paced reading. (B) Participants clicked the sentence to indicate that they finished reading; the sentence turned into a black box. Click time was a measure of reading duration. (C) Participants moved the black box to the appropriate response area and released the left mouse button when finished. The click release was a measure of movement duration.

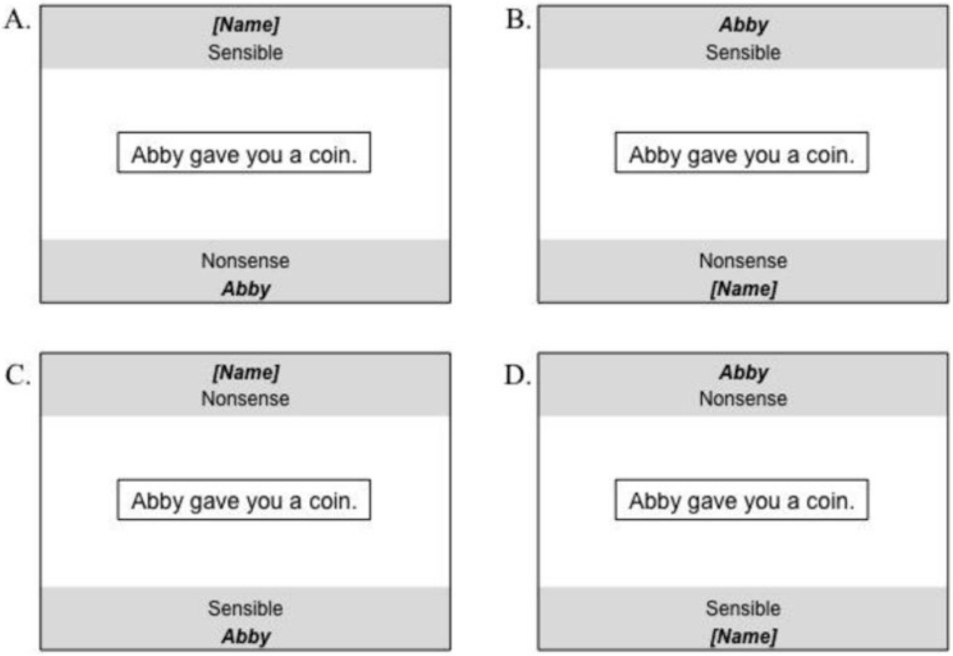

Figure 2.

Experimental setup. Panels (A) and (C) represent the “disembodied” conditions, during which participants' names were displayed away from their physical bodies. Panels (B) and (D) are conditions in which participants' name location was redundant with their physical body. Panels (A) and (B) depict the “Sensible is far” (from the body) condition, and panels (C) and (C) depict the “Sensible is near” (to the body) condition.

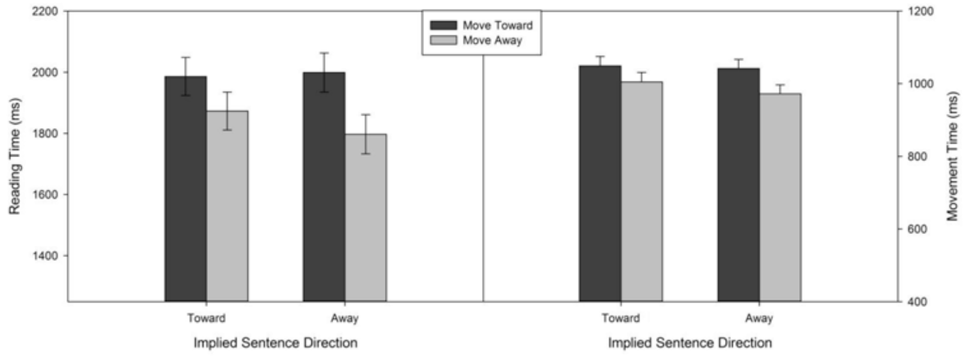

Figure 3.

Average reading times (left panel) and movement times (right panel) in Experiment 1 (error bars represent standard error).

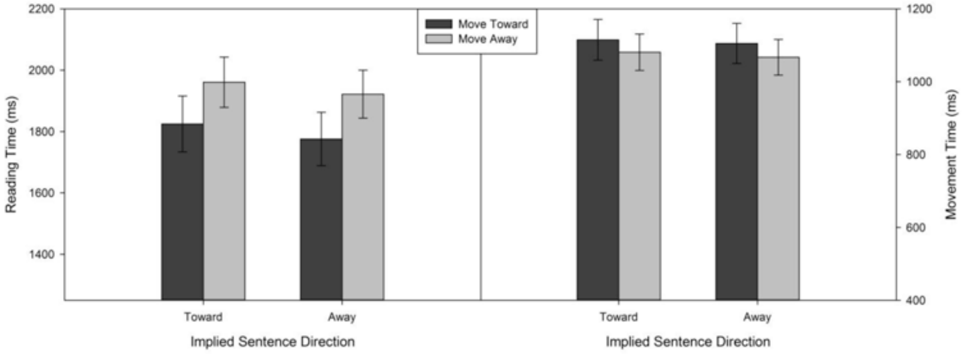

Figure 4.

Average reading times (left panel) and movement times (right panel) in Experiment 2 (error bars represent standard errors).

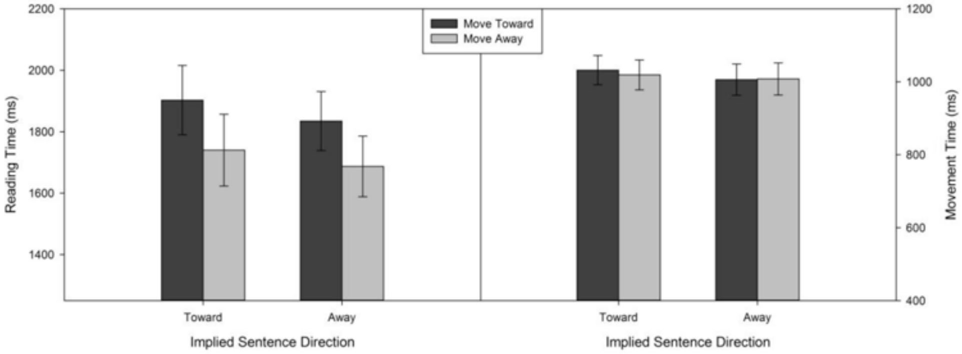

Figure 5.

Average reading times (left panel) and movement times (right panel) in Experiment 3 (error bars represent standard errors).

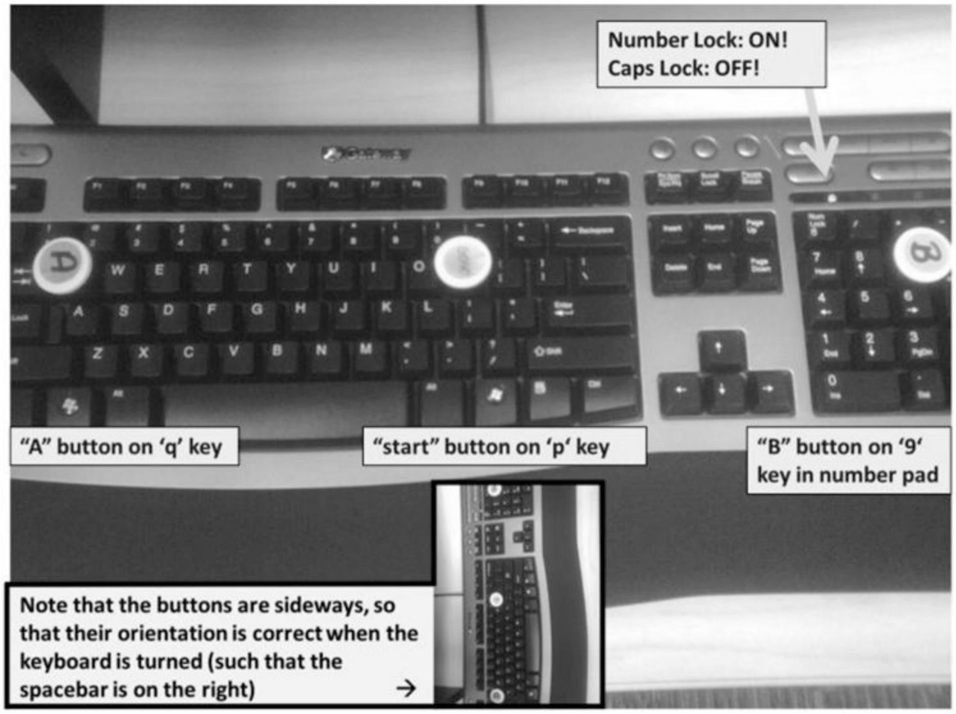

Figure 6.

Depiction of the keyboard setup for Experiments 4A, 4B, and 7. The inset picture (thick black outline) shows the final orientation of the keyboard from a participant's perspective.

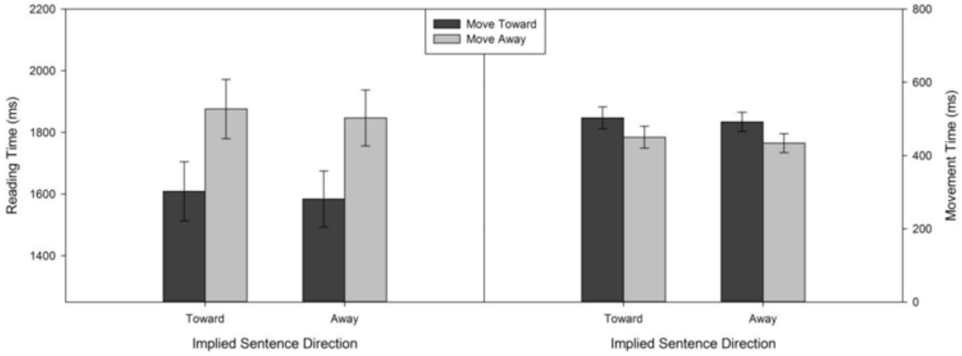

Figure 7.

Average reading times (left panel) and movement times (right panel) for Experiment 4A (error bars represent standard error). *Note:* Average movement times with keyboard input were faster than those with mouse input, so the movement time scale differs from mouse input experiments.

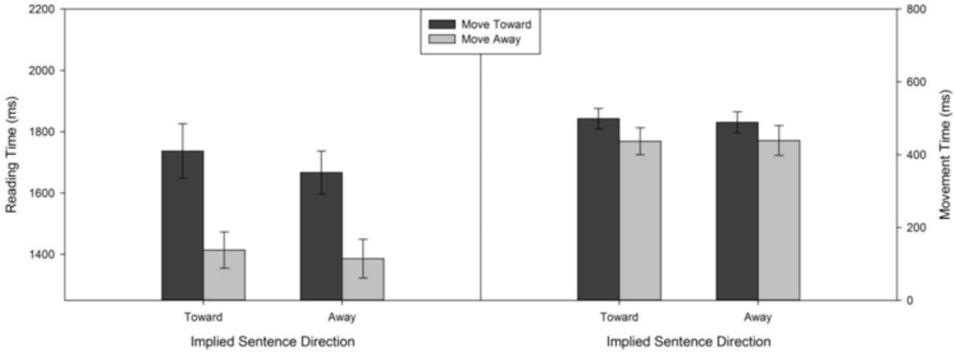

Figure 8.

Average reading times (left panel) and movement times (right panel) for Experiment 4B (error bars represent standard error).

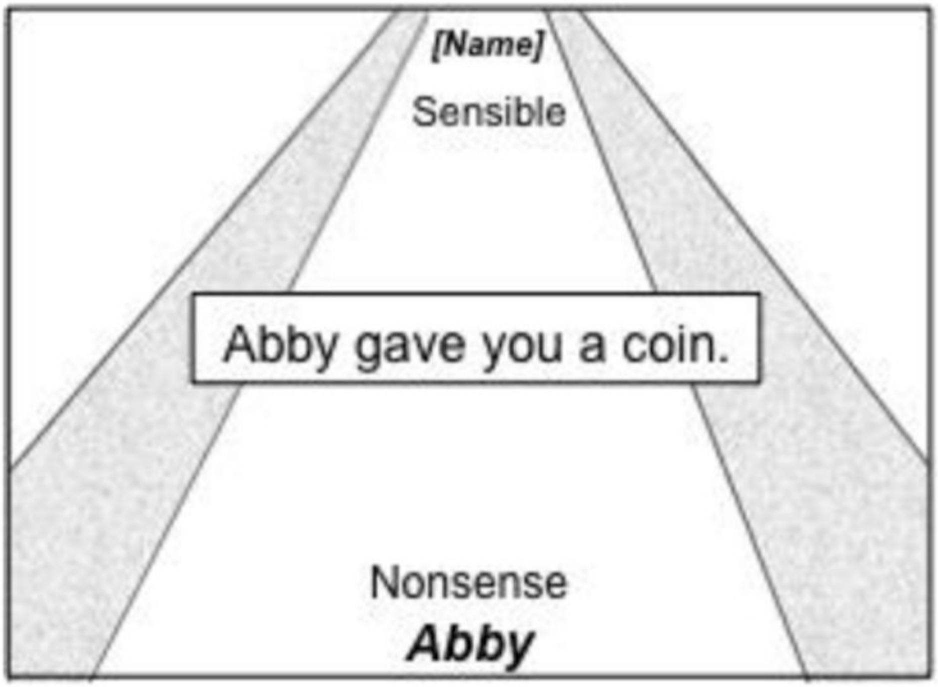

Figure 9.

Depiction of the screen setup for Experiments 5 and 6. Participants clicked the sentence to initiate their decision, which turned the sentence into a black box. Participants then dragged the black box to the Nonsense or Sensible response area. In Experiment 6, no names appeared on screen.

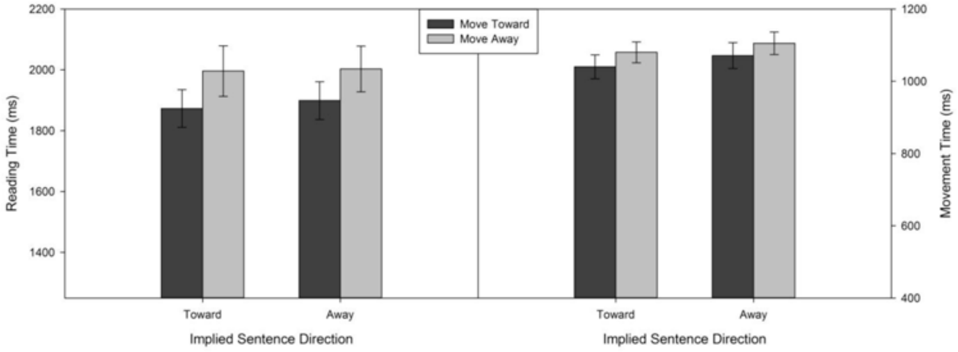

Figure 10.

Mean reading times (left panel) and movement times (right panel) in Experiment 5 (error bars represent standard errors).

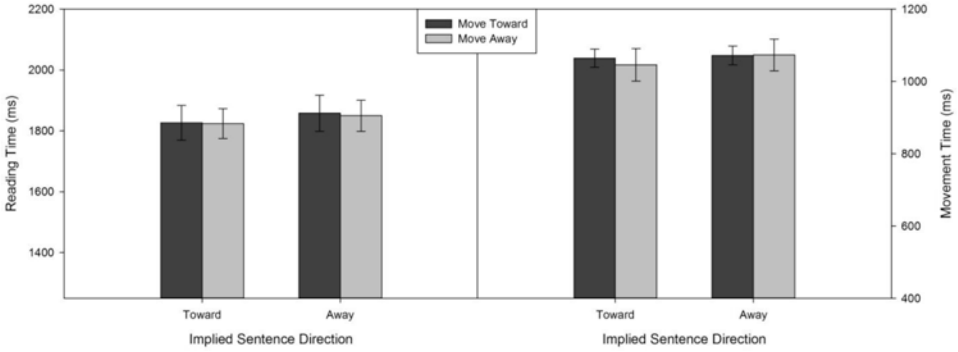

Figure 11.

Mean reading times (left panel) and movement times (right panel) in Experiment 6 (error bars represent standard errors).

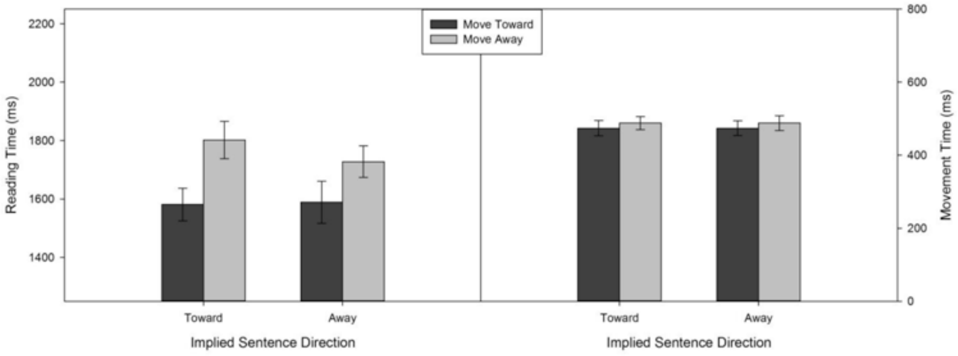

Figure 12.

Mean reading times (left panel) and movement times (right panel) in Experiment 7 (error bars represent standard errors).

Figure 13.
Example disguised keyboard used in Experiment 8.

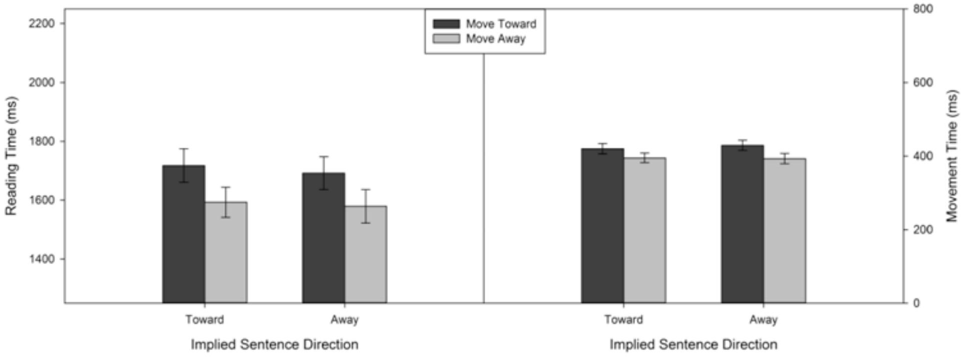

Figure 14.

Mean reading times (left panel) and movement times (right panel) in Experiment 8 (error bars represent standard errors).

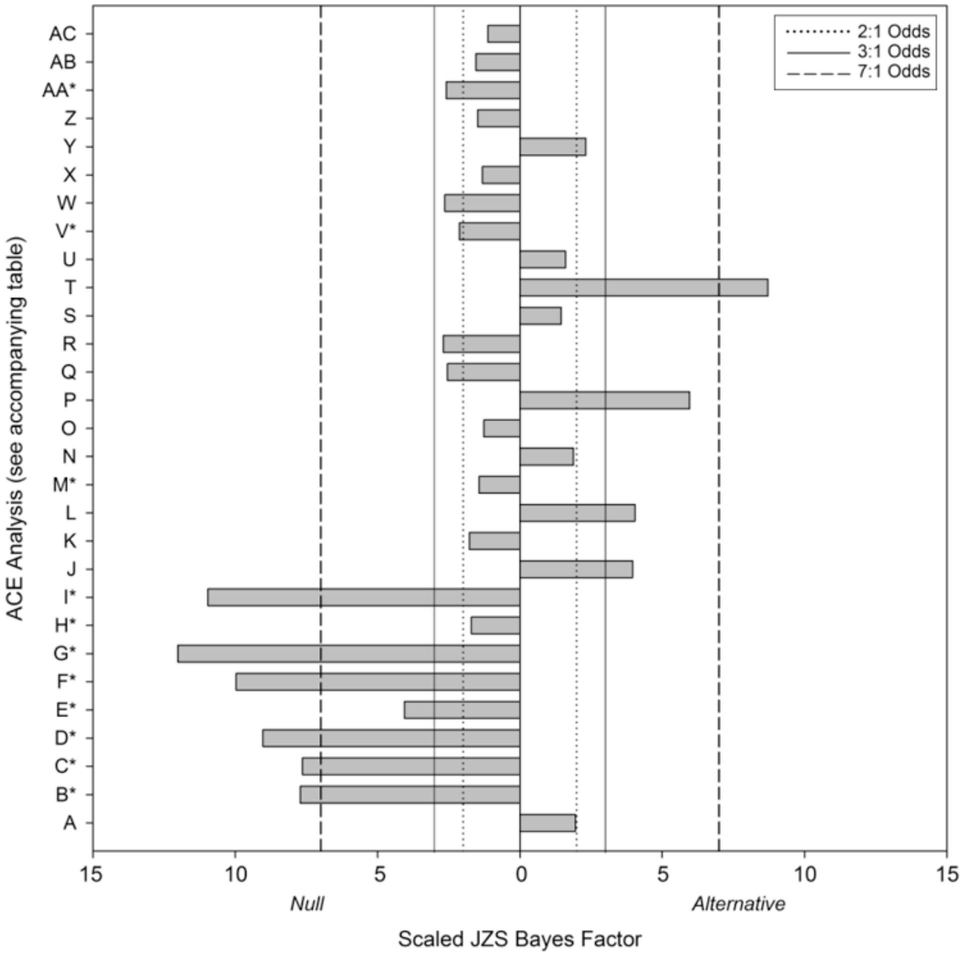

Figure 15.

Scaled JZS Bayes Factors for analyses reported in the present set of experiments (A through I) and the existing literature (J through AC). Left-going bars denote evidence in favor of the null hypothesis, and right-going bars denote evidence in favor of the null. Vertical lines represent cutoffs for the strength of the evidence. See Table 2 for additional statistics and citations. *Note: Analyses with * were originally reported as null results.*

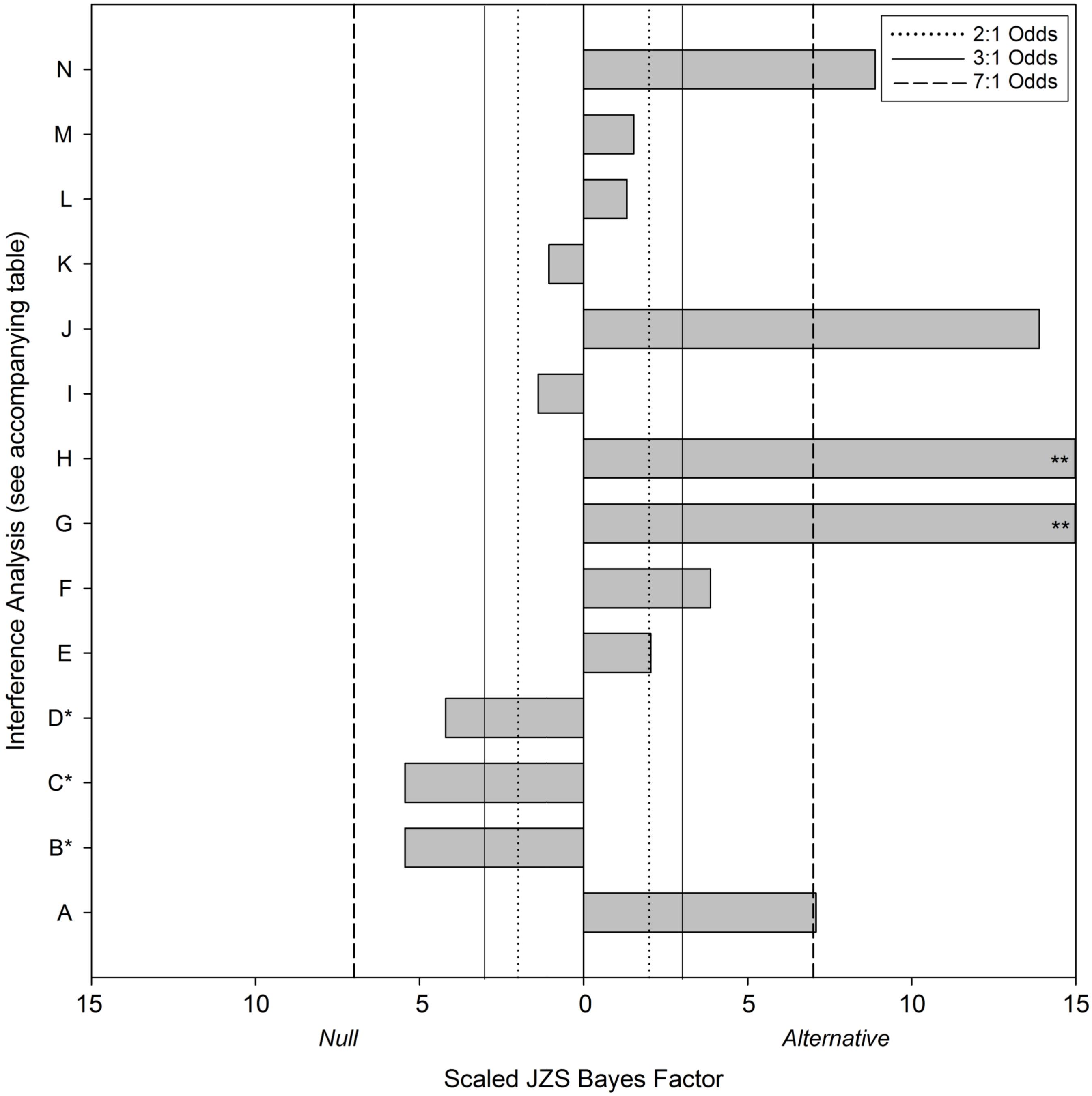

Figure 16.

Scaled JZS Bayes Factors for analyses reporting inhibitory ACEs. Left-going bars denote evidence in favor of the null hypothesis, and right-going bars denote evidence in favor of the null. Vertical lines represent cutoffs for the strength of the evidence. See Table 3 for additional statistics and citations. *Note: Analyses with * were originally reported as null results, and bars with ** exceed 100.*

Table 1
Participants per condition, Experiments 1 through 3

Condition	Exp 1	Exp 2	Exp 3
Sensible is near	62	24	19
Sensible is far	62	30	18

Table 2

ACE Facilitation Evidence

Figure	Article	Experiments	Analysis/Effect	Scaled JZS Bayes Factor	Odds Favor	2:1 Threshold	3:1 Threshold
A	Current	1	ACE interaction	1.96	Alternative	N	N
B	Current*	2	ACE interaction	7.72	Null	N	N
C	Current*	3	ACE interaction	7.64	Null	N	N
D	Current*	4A	ACE interaction	9.03	Null	N	N
E	Current*	4B	ACE interaction	4.06	Null	N	N
F	Current*	5	ACE interaction	9.97	Null	N	N
G	Current*	6	ACE interaction	12.02	Null	N	N
H	Current*	7	ACE interaction	1,71	Null	N	N
I	Current*	8	ACE interaction	10.96	Null	N	N
J	Glenberg & Kaschak (2002)	1	ACE interaction	3.97	Alternative	Y	Y
K	Glenbera & Kaschak (2002)	2A	ACE interaction	1.78	Null	N	N
L	Borreggine & Kaschak (2006)	1	ACE interaction	4.05	Alternative	Y	Y
M	Boulenger et al.(2006)*	2	RT for verbs vs nouns	1.44	Null	N	N
N	Boulenger et al. (2006)	2	Peak wrist acceleration for verbs vs nouns	1.89	Alternative	N	N
O	Zwaan & Taylor (2006)	2	Sentence rotation × Manual rotation	1.27	Null	N	N
P	Zwaan & Taylor (2006)	3	Observed rotaion × Sentence direction	5.97	Alternative	Y	Y
Q	Zwaan & Taylor (2006)	4	Verb region, congruence with rotation direction	2.55	Null	N	N
R	Zwaan & Taylor (2006)	5	Illusory rotation, verb region, congruence with sentence direction	2.70	Null	N	N
S	Glenberg et al. (2008)	1	ACE interaction	1.46	Alternative	N	N
T	Kaschak & Borreggine (2008)	1	500 ms condition	8.72	Alternative	Y	Y
U	Kaschak & Borreggine (2008)	1	toward-toward, 500 ms condition	1.61	Alternative	N	N
V	Kaschak & Borreggine (2008)*	1	away-away, 500 ms condition	2.13	Null	N	N
W	Taylor & Zwaan (2008)	1	Verb region, ACE interaction	2.64	Null	N	N
X	Taylor & Zwaan (2008)	1	Adverb region, ACE interaction	1.32	Null	N	N
Y	Taylor & Zwaan (2008)	2	Verb region, ACE interaction	2.32	Alternative	Y	N
Z	de Vega & Urrutia (2011)	1	Factual sentences, 200 ms delay	1.49	Null	N	N

Figure	Article	Experiments	Analysis/Effect	Scaled JZS Bayes Factor	Odds Favor	2:1 Threshold	3:1 Threshold
AA	de Vega et al. (2013)*	1	350 ms verb-cue SOA. motor response time	2.59	Null	N	N
AB	de Vega et al. (2013)	1	350 ms verb-cue SOA, semantic decision time	1.55	Null	N	N
AC	de Vega et al. (2013)	1	350 ms verb-cue SOA, total response time (motor & semantic)	1.13	Null	N	N

Note: Articles with asterisks in the first column originally reported a null effect for the listed analysis.

Table 3

ACE Interference Evidence

Figure	Article	Experiment	Analysis/Effect	Scaled JZS Bayes Factor	Odds Favor	2:1 Threshold	3:1 Threshold
A	Buccino et al. (2005)	2	Behavioral data, Effector interference	7.08	Alternative	Y	Y
B	Borreggine & Kaschak (2006)*	2	50 ms post-sentence cue. ACE interaction	5.45	Null	N	N
C	Borreggine & Kaschak (2006)*	3	500 ms post-sentence cue. ACE interaction	5.45	Null	N	N
D	Borreggine & Kaschak (2006)*	4	1000 ms post-sentence cue. ACE interaction	4.21	Null	N	N
E	Boulenger et al. (2006)	1	Peak movement latency, verbs vs nouns	2.05	Alternative	Y	N
F	Boulenger et al. (2006)	1	Peak movement amplitude, verbs vs nouns	3.87	Alternative	Y	N
G	Sato et al. (2008)	1	Effector interference	408.48	Alternative	Y	Y
H	Sato et al. (2008)	2	Effector × SOA	51070.94	Alternative	Y	Y
I	de Vega & Urrutia(2011)	1	Counterfactual, 100 ms SOA	1.39	Null	N	N
J	de Vega & Urrutia(2011)	2	Counterfactual, 100 ms SOA	13.89	Alternative	Y	Y
K	de Vega et al. (2013)	1	Motor response time, 100 and 200 ms SOAs	1.06	Null	N	N
L	de Vega et al. (2013)	1	Total response time, 100 and 200 ms SOAs	1.32	Alternative	N	N
M	de Vega et al. (2013)	2	Motor response time. 200 ms SOA	1.53	Alternative	N	N
N	de Vega et al. (2013)	2	Total response time, 200 ms SOA	8.89	Alternative	Y	Y

Note: Articles with asterisks in the first column originally reported a null effect for the listed analysis.