Distribution Agreement

Signature:

______________________________ _______________
Lauren McDonough Date

Grounded Congruency Effects: Automaticity and 'Strategery' in Cognition

By

Lauren A. McDonough
B.A. University of Notre Dame, 2008

Psychology

____
Lawrence W. Barsalou, Ph.D.
Advisor

____
Robert Hampton, Ph.D.
Committee Member

____
Stella Lourenco, Ph.D.
Committee Member

Accepted:

____
Lisa A. Tedesco, Ph.D.
Dean of the James T. Laney School of Graduate Studies

____
Date

Grounded Congruency Effects: Automaticity and ‘Strategery’ in Cognition

By

Lauren A. McDonough
B.A. University of Notre Dame, 2008

Advisor: Lawrence W. Barsalou, Ph.D.

An abstract of
A thesis submitted to the Faculty of the
James T. Laney School of Graduate Studies of Emory University
in partial fulfillment of the requirements for the degree of
Master of Arts
in Psychology
2010

Abstract

Grounded Congruency Effects: Automaticity and ‘Strategery’ in Cognition

By Lauren A. McDonough

According to theories of grounded cognition, words whose semantics are associated with a salient vertical position (e.g., CEILING vs. CARPET) should activate simulations of these positions in space. When responses are analogously made in the vertical dimension, grounded congruency effects should result (e.g., processing CEILING should be faster for an UP vs. DOWN response). Previous research obtained grounded congruency effects when participants used ink color (RED vs. BLUE) as a cue for response direction (Casasanto, 2008). Typically researchers assume that these effects are automatic, but they could possibly be strategic. In addition, we also explored a possible correlation between grounded congruency effects and empathic perception. Two experiments attempted to assess these issues with 6 groups of 24 participants each, but failed to replicate the original grounded congruency effect, leading us to question its reliability when ink color is used as a cue. We further discovered a motor facilitation effect for upward as opposed to downward responses not reported previously in this paradigm.

Grounded Congruency Effects: Automaticity and 'Strategery' in Cognition

By

Lauren A. McDonough
B.A. University of Notre Dame, 2008

Advisor: Lawrence W. Barsalou, Ph.D.

An abstract of
A thesis submitted to the Faculty of the
James T. Laney School of Graduate Studies of Emory University
in partial fulfillment of the requirements for the degree of
Master of Arts
in Psychology
2010

Table of Contents

List of Tables and Figures

Grounded cognition ties thoughts back into bodily states, motor events, and perceptual experiences on the basis of an interactive perceptual-motor simulation system. For example, when you read the word ‘bird,’ you internally simulate a high spatial location. In this way, the simulation system facilitates understanding. There is evidence for simulation in conceptual processing, action and language comprehension (Barsalou, 2007). But, its principal mechanisms have yet to be deeply explored. Is it an automatic or strategic process? Most experiments within the grounded cognition literature are not designed to discriminate between the two, and instead often assume that the underlying process is automatic. This assumption is unnecessary as several methods involving set manipulation exist that can discriminate between automatic and strategic processing. By coupling a behavioral paradigm in grounded cognition that involves spatial simulation with set manipulation, we attempted to identify whether simulation reported in many grounded cognition studies is automatic or strategic.

Grounded Cognition

The defining principle behind “grounded cognition” is that cognition is rooted in perceptual and motor experience, as opposed to a series of abstract symbol translations in the brain. Across the different modalities (vision, hearing, touch, taste etc.), simulations of external stimuli fuel our thought processes. For example, presented with the concept ‘pig,’ you simulate past experiences with pigs, what they smell like, feel like etc. Conversely, amodal cognitive views predict you will represent “pig” void of any sensory-motor experience.

At the neural level, simulation is a type of “reenactment” (Barsalou, Niedenthal, Barbey, & Ruppert, 2003). A stimulus is perceived and then encoded by conjunctive neurons in a feature map pattern. The pattern is later reactivated by associated neurons (through cognition, memory, language etc.), and the original perception is reenacted (Barsalou et al, 2003). Because it is not a full reenactment of perception, it can be somewhat imprecise or distorted. This distortion can occur because we have a “simulator” for each conceptual category that is an amalgamation of all our experiences with that category. Simulation takes place when a certain portion of the simulator is activated in a particular instance. There are three main types of simulation. A basic or simple simulation occurs on one modality. For example, you hear a song playing on the radio and later cognitively reenact that auditory pattern. In addition, simulation can also occur across the modalities. For example, you see a picture of flying squirrel and you simulate the smell of the animal, how soft the fur is, etc. Another type of simulation specifically involves the motor system and mirror neurons. The mirror neuron system, first located in monkeys, is considered one neural explanation of simulation (Arbib, 2005; Rizzolatti & Arbib, 1998). The cortical correlates in humans fire during self-generated action and observation of action executed by another (Decety, 2007; Gallese & Goldman, 1998; Jeannerod, 2006). Often, the observer will experience some resonating physical reactions (e.g. muscle activation) as a consequence of neurally mirroring the other (Avenanti, Bueti, Galati, & Aglioti, 2005; Jeannerod, 2006).

You do not even have to overtly witness an action for simulation and resonance to take place. Decety, Jeannerod, Germain, and Pastene (1991) measured respiration and heart rate while participants imagined themselves running and found that heart rate and respiration increased in tandem with the pace of imagined running.

In addition to action observation, language also activates modal simulations. While reading action related words, people simulate the motor movement, which in turn activates related muscle areas (Jeannerod, 2006). For example, reading verbs and adjectives describing happy and angry emotional states activated facial muscles associated with expressing those emotions. Participants were simulating the state described as if they were visually witnessing those emotions or expressing that affective state themselves (Foroni & Semin, 2009). Whether sparked linguistically or experientially, interaction between the motor and perceptual systems through simulation is the foundation of human cognition. For a comprehensive review of theory and research in grounded cognition, see Barsalou (1999, 2007).

Grounded congruency effects. Often the perceptual simulation or muscle activation that results from simulation can affect how we act, helping or hindering the fluidity of cognition. One such grounded effect, demonstrated repeatedly in the literature, is grounded congruency.[1] For example, grounded congruency effects can occur when a simulated action matches the action to be performed. Glenberg and Kaschak (2002) demonstrated a grounded congruency effect with action and language. Participants were required to read action-related sentences and judge whether the sentence made sense or not. To respond, participants pulled a lever toward themselves or pushed it away from themselves, depending on whether they wanted to answer 'yes, the sentence makes sense' or 'no, it does not.' All of the sentences had an implied direction. For example, "You hand the book to Maggie," implies a movement of your hand away from your body to give the book to Maggie. In judging sensibility of the sentences, participants were deeply processing the meaning of the sentence and simulating the actions they described. Depending on the implied direction in the sentence and the movement performed to respond 'yes' or 'no,' either the simulation would match or clash with the action. For example, if the 'yes, this sentence is sensible' response required participants to push the lever away from themselves as they read the sentence "You open the drawer," (simulation of movement toward self), the movement and simulation would be in opposite directions. When this incongruency

---

[1] Here we discuss grounded congruency effects in the realm of motor cognition, but the largest number of congruency effects reported has been perceptual: e.g., Zwaan and Madden (2005).

occurred, response times were relatively slow and often inaccurate. When lever movement matched the implied directions in sentences (i.e., simulation), participants were faster and more accurate (e.g., when the sentence was, "You open the drawer" and participants had to pull the lever toward themselves to respond). Congruent simulation and response even aided participant comprehension of the sentences (Glenberg & Kaschak, 2003).

Similar grounded congruency effects occur not only with words that describe action, but also with words that describe spatial location. When someone reads such words, associated spatial meanings are activated, which influences perception and action. For example, Casasanto (2008) demonstrated that when participants made an upward response to a word whose meaning was associated with up (e.g., 'bird'), responses were faster compared to conditions in which they made an upward response to a word whose meaning was associated with down (e.g., 'anchor'). Similarly, Ulrich and Maienborn (2008) found congruency effects for the abstract concept of time grounded in a mental timeline (oriented left to right). Participants judged whether sentences were related to the past (left on the mental timeline) or the future (right on the mental timeline). When future responses required a right-hand response and past judgments required a left-hand response (congruent conditions), participants were faster than when the response movements were reversed (incongruent conditions). But do these congruency effects reflect automatic or strategic activation of spatial meaning?

One reason to think that some of the reported grounded congruency effects, or some component of these effects, are not automatic is that often no fillers are used with the critical materials (e.g., if the critical materials are words whose meanings are associated with high and low spatial locations, then fillers would be words whose meanings are not associated with up or down). Perhaps such manipulations are so obvious that participants perceive them, thereby establishing a cognitive set. For example, participants become aware that there are high and low spatially located words within the experiment. Once this ‘high vs. low’ set is established, it strategically (as opposed to automatically) activates semantic information about up vs. down (or whatever is relevant) in the target words. Ulrich and Maienborn (2008) specifically addressed this issue in the aforementioned experiment using the coupling of time with the spatial representation of a mental timeline. When the task required participants to specifically judge the past or future qualities of the sentences, they demonstrated grounded congruency effects. However, when they simply had to judge the sensibility of the sentence (akin to the Glenberg and Kaschak studies), and not the relation to past and future, the congruency effects vanished. That is, when the spatial qualities of the sentences were not specifically necessary to complete the task judgment, they were not simulated, and did not affect the motor responses of the participants (causing grounded congruency effects). These results suggest that simulation of the spatial qualities associated with some words may be strategic, brought about as a result of the task at hand, and not an automatic process that happens regardless of context.

Automaticity and Set

What, indeed, does it mean for a process to be automatic, as opposed to strategic, in the first place? A controlled or strategic process depends on an individual’s attention and intention (Borgmann, Risko, Stolz, & Besner, 2007). This type of processing is more variable and less reliable than its automatic counterpart because an automatic cognitive process is a sequence of neural activation that, at the very least, is autonomous and does not require “active control or attention by the subject” (Moors & De Houwer, 2006; Shiffrin & Schneider, 1977; Posner & Snyder, 1975). ‘Active control’ implies conscious accessibility, another hallmark of strategic processing (Shiffrin & Schneider, 1977). In contrast, an autonomous act is self regulated and unaffected by external influences (Moors & De Houwer, 2006). In addition to autonomy, at least eight other features have been identified as possibly characteristic of an automatic process, indicating a certain level of variability in the conceptual preconditions of automaticity and feasibility of its measurement.

According to a Moors and De Houwer (2006), an automatic act is unintentional, goal independent, uncontrolled/able, autonomous, often purely stimulus driven, unconscious, efficient and fast. Each of these eight properties of an automatic process can be described further as follows:

An *intentional* act is ‘directed.’ One acts because of a goal in mind. The person exerts influence by initiating the act. On the contrary, an automatic process “begins regardless of intent” (Borgmann, Risko, Stolz, & Besner, 2007; Posner & Snyder, 1975), and is “not caused by a goal” (Moors & De Houwer, 2006).

A *controlled/able* act is much the same, except that a person may exert even more influence in a purposeful (non-random) way after the process begins and during the course of action. Conversely, an automatic act is carried out without active management. An uncontrolled act refers back to a specific processing goal, in that action x was not controlled by goal y.

Automatic processes are also *autonomous*. An autonomous act is very similar to an uncontrolled act, but autonomy is more inclusive. Instead of act x being uncontrolled by goal y, act x is uncontrolled by all potential goals. For Moors and De Houwer (2006) autonomy is the minimal requirement for an automatic process: the process “is not produced, altered, stopped, or avoided by a processing goal.”

A *goal dependent* event requires that a goal be the basis for partaking in the action. Importantly, according to Moors and De Houwer, a ‘goal’ can be either conscious or unconscious. An automatic process is goal independent in that a goal does not initiate the action. Automaticity can also be characterized by evidence that a process was *purely stimulus driven* in that it was “produced by the mere presence of a stimulus” (Moors & De Houwer, 2006).

Consciousness implies awareness derived from the senses and directed toward something outside of the self. Unconsciousness requires no attention or awareness, and is, to some degree, inaccessible. An automatic act is executed unconsciously.

Automatic processing feels unforced and easy because it is *efficient,* and expends minimal attentional and processing resources (Borgmann, Risko, Stolz, & Besner, 2007; Moors & De Houwer, 2006). It is also *fast* compared to strategic processing.

Though Moors and De Houwer separate these eight features, their definitions are closely intertwined, and the degree of discernibility between components is still debatable. Moors and De Houwer (2006) posit that to garner a true understanding of the degree of automaticity in a process, each of these eight features has to be measured, or at least considered, separately from the others. They emphasize separate analysis because each feature develops or is carried out to completion at a different rate. For example, a process may cross the threshold of being intentionally vs. unintentionally executed, but conscious awareness of such a change may not shift concurrently. Researchers might fail to identify a process as unconscious or unintentional because they are only taking one measurement at a certain time. If a second measurement of participant responses were taken, it would capture the shift in both the intentionality and consciousness features. This variability in the achievement of each of the eight features may be the source of the perception that automaticity is internally inconsistent, and may account for many conflicting results regarding whether a process is automatic or not.

In addition to disputes over conceptualization and measurement of automaticity, there appears to be an over-emphasis of its necessity in the field of grounded cognition. Because many processes from social perception to behavioral mimicry to affective judgments are automatic, lacking conscious volition or free will, cognitive science is, for the most part, based on deterministic principles (Bargh & Ferguson, 2000). Therefore it makes sense that researchers might assume simulatory processing and associated grounded congruency effects are included in the long list of 'automatic' processes. However, Bargh and Ferguson (2000) point out that the difference between automatic and strategic processing may not be so profound. The actual difference between an automatic and strategic or controlled process is not that strategic is inherently "caused," while automatic lacks a cause, but simply that researchers have yet to ascertain what controls automaticity. In the same line of thinking, "controlled processes must themselves be controlled" (Bargh & Ferguson, 2000). Being consciously aware of a procedure does not necessarily preclude it from being determined. Nevertheless, it would be beneficial to understand the nature of grounded congruency's underlying mechanism to either bolster support for existing theories based on assumptions of automaticity or relieve the field of unnecessary preoccupations with automatic processing if it turns out congruency effects result from strategic processing.

Evidence for two different stages in semantic processing, in particular, provides significant support for attenuating the 'need' for automatic processing in grounded cognition. The first stage of processing is an automatic, faster, more general, all-inclusive evaluation of a situation or logical matter (Smith, Shoben, & Rips, 1974). These characteristics often make an individual vulnerable to committing more errors in decision-making. Stage two of processing witnesses a dramatic decrease in errors because it is much more discriminatory during assessment, looking at defining characteristics as opposed to all the features of a category (Smith et al, 1974). However, strategic stage two processing is much slower than the automatic stage one. This may be why strategic processing occurs later or not at all in a situation that does not require this more effortful form of processing.

Studies illuminating the influence of language on activating motor representations support Smith et al.'s (1974) theory regarding a preliminary automatic stage of processing, followed by a strategically controlled stage two. Gentilucci and Gangitano (1998) measured how participants reached for a series of rods placed equidistantly from their body, but marked with the word "short" or "long." Participants were not given instructions about the words on the rods, but were simply told to reach and grasp the rod as quickly and accurately as possible. Researchers found that rods printed with the word "long" caused participants to initially make arm movements as if the rod were farther away (higher velocity and greater acceleration toward object). As the arm got closer to the rod, the movement was corrected, and they could still accurately pick it up, despite an initial error in reaching calculation. Simulation elicited by reading "long" on the rods may have activated the reach sequence for farther away, but the initial prompt can be overridden and corrected as the reach decelerated in time to pick up the rod without over-shooting the distance. Gentilucci and Gangitano (1998) demonstrated that language comprehension influenced motor response, but a second stage of more strategic processing took over to correctly adjust participant's movement when reaching for an object.

Because the two stages can influence cognition to varying degrees, a wide range of response times within and across experiments is measured, and perhaps a misidentification of processing type occurs (Smith et al., 1974). Therefore, paradigms within cognitive psychology identified as using strategic processes can still be grounded, and even perhaps automatic, as researchers are just measuring one stage of the process and not the other (e.g., stage two strategic, instead of stage one automatic or vice versa). Up to this point, researchers in grounded cognition may have been looking simply at stage one processing, and using its automaticity to define the field. Perhaps the more intriguing information lies in the deeper, controlled processing of stage two, still grounded in simulation and experience, but not as unconscious or automatic as stage one. With these points in mind, it appears the emphasis of automaticity in the grounded literature is not only untested, but also perhaps unnecessary and limiting. Because automaticity is not a prerequisite for the grounding of a process or conceptualization, a conflagration of this 'requirement' is potentially restrictive regarding experimental designs and the avenues of possible research in the field.

How do we begin to empirically distinguish between the strategic and automatic processing in a behavioral experiment in the first place? Within the attention literature, Posner, Snyder and Davidson (1980) identified a way to perform this differentiation using set manipulation. By introducing different ratios of neutral fillers (i.e., items unrelated to the critical stimuli), one can hinder strategic processing. It becomes harder to form a tactical plan in processing if fillers mask a pattern in the stimuli. For example, if the critical items in an experiment are words that have spatial qualities (high vs. low), and words with no spatial qualities like "but, and, the" are mixed in, it is harder for participants to identify the spatial set of critical items. If, however, processing for the task is not strategic but is instead automatic, then mixing fillers with the critical words will not affect processing. That is, congruency effects will still occur with or without fillers. Because the critical stimuli trigger processing automatically, it does not matter how many fillers are present. For example, participants will be faster to respond when the response movement is up and the word they are reading has high spatial qualities (e.g., 'bird') regardless of how many non-spatial words are present.

Several studies have used set manipulation to distinguish between automatic and strategic processing. In a series of detection and attention-orienting studies, Posner, Snyder and Davidson (1980) required participants to identify a change in luminescence at a particular location. Initially, eighty percent of the trials involved a cue that would correctly predict the location of the change in light, while in the remaining twenty percent of the trials the light was equally likely to appear in any of four set locations (i.e., it was not an accurate cue for light change). In other conditions, the number of filler, non-predictive cues was increased, such that the predictive cues were outnumbered by the non-predictive cues, 20% to 80%, respectively. When the ratio of predictive to unpredictive cues was 80:20, participants demonstrated an expectancy effect relative to when the ratio was 20:80. Response times were much faster with less fillers (non-predictive cues) because there was such a high proportion of correctly predictive cues. This facilitation effect became much smaller in other conditions that had considerably more filler items, because the cues did not typically forecast the critical item's location. Because the 20:80 ratio of critical items to fillers decreased the ability of the participant to adopt a cognitive strategy in making responses, there is strong evidence for a robust influence of set on response times in this particular task.

People often assume that the classic Stroop effect is completely automatic. Jacoby, Lindsay, and Hessels (2003), however, found similar effects as Posner et al. (1980) after manipulating the proportion of critical and filler items in Stroop experiments. Participants were shown a series of color-name words and asked to identify the font color. In some trials the word and font color were congruent, whereas in others they were incongruent. For example the word

BLUE in blue print would represent a congruent condition, and the word BLUE in red print would be incongruent. The conditions with 80% congruent (matching color), and 20% incongruent trials had an augmented Stroop effect. Participants were very slow to respond to incongruent trials and made many mistakes because they had come to rely and trust the semantic meaning (and simulation) of the word for their answers. When the ratio was reversed (high proportion of incongruent items), the Stroop effect was reduced because participants learned not to expect a correspondence between word and color. They were able to strategically overcome the interference between color word and font. Here again a manipulation of set drastically altered the results, nearly causing a complete reversal of findings. Such evidence indicates that the Stroop effect is not completely automatic as has been widely believed for decades.

The same fate has befallen the Simon effect, another phenomenon viewed widely as strictly automatic. The Simon Effect, faster response times when a response key location matches a target object position (Borgmann, Risko, Stolz, & Besner, 2007; Simon, 1990), is vulnerable to the same set manipulation as the Stroop effect. Borgmann et al. (2007) found that when there were more congruent trials (e.g., response and object location the same), response times were facilitated. When there were more incongruent trials (e.g., response location opposite object position), the Simon effect became inverted and unreliable.

In related work, Borgmann et al. (2007) presented ‘X’s and ‘O’s on different sides of a computer screen split down the middle. Subjects had to press a key either on their left or right to “identify” the target. For example, whenever they saw an X they had to press the left key regardless of the X’s position on the screen. Congruent conditions (e.g., when a left response key identified X, and X appeared on the left side of the screen) elicited faster response times. Borgmann and colleagues manipulated the ratio of congruent to incongruent trials. When incongruent trials (e.g., X appears left, but right key identifies X) dominated the set in a 75:25 ratio, the Simon effect was fallaciously reversed, suggesting, as with Jacoby, Lindsay, and Hessels (2003), that the Simon effect is a strategically based process as opposed to an automatic one.

All three of the aforementioned studies used proportions of fillers to critical items around an 80/20 ratio, which seems to be the optimal manipulation of set to demonstrate set effects. Some studies within the grounded and attention literature implement fillers, but only using around a 50/50 ratio or less (e.g., Jha, 2002; Pecher, Zeelenberg, & Barsalou, 2003; Wilson & Knoblich, 2005). Based on previous research, this proportion may not be drastic enough to remove a strategic process and/or illuminate an automatic one. Implementing set manipulation effectively in established grounded cognition experiments would not only help us understand what mechanism triggers grounded experience, but also correct a methodological problem in the field.

Assessing whether congruency effects are automatic. Three methods exist that are well suited for addressing whether processing is strategic or automatic (Posner, Snyder, & Davidson, 1980). Method 1 requires manipulation of the proportion of irrelevant fillers mixed with critical target words. For example, one can manipulate whether the ratio of critical (e.g., high vs. low words) to fillers (e.g., non-spatial words) is 80:20 or 20:80. High ratios of fillers to critical targets will obscure the critical manipulation (e.g., high vs. low). Because the fillers outnumber the critical targets, it becomes much more difficult if not impossible to detect the manipulation and establish a set. If the up-down congruency effect is automatic, it should remain regardless of the filler ratio used. If the effect results from participants deducing the manipulation and establishing a set because nothing obscures it, then the congruency effect should diminish or disappear when fillers outnumber critical items.

Another way to distinguish strategic vs. automatic processing is to manipulate set orthogonally to the semantic distinction of interest. For example, in an experiment where up vs. down spatial meaning is being manipulated, have participants focus on an orthogonal dimension associated with the words, such as whether they are open or closed class words. Manipulating awareness of an orthogonal set should obscure the critical manipulation, acting as a cover story. When the set is grammatical class, it will distract participants from the height manipulation in the words. Specifically, if participants are told that the words being processed are from open vs. closed class word sets, then they should be less likely to realize that high vs. low spatial location also varies. If the semantic distinction of interest (high vs. low) is automatic, then the associated congruency effects should remain unattenuated regardless of whether there is an orthogonal set or not (open vs. closed class). Grounded congruency effects for high-low should be just as strong when participants believe that the words vary in being open vs. closed class as when they are unaware of their class difference. If, however, the semantic distinction of interest is not automatic, then the associated congruency effects should be attenuated. Grounded congruency effects for high-low should diminish or disappear when participants are concentrating on the variance in open vs. closed class.

While methods 1 and 2 use set *interference* to elucidate whether strategic vs. automatic processes underlie grounded congruency effects, a third method could reveal a set *facilitation* process. Manipulating explicit awareness of the true set using no fillers or orthogonal sets makes this possible. In an experiment that varies whether words have high vs. low meaning, for example, manipulate whether participants are told explicitly whether the words have high vs. low associated locations or are told nothing about the manipulation. Essentially this technique explicitly creates a demand and assesses if that demand increases the size of the congruency effect. If participants in previous experiments did not induce the manipulation, such that the manipulation only produced congruency effects via automatic mechanisms, then the set congruency effect should increase when participants are told about the up-down manipulation, because this adds a set that is not normally present. Under these conditions, the set adds a second source of activation in addition to the original automatic activation. Conversely, if participants in previous experiments have induced the manipulation, then telling them about the manipulation should not increase the congruency effect, relative to not telling them. Because activation is strategic even when participants are not told about the manipulation, there should be no increase when they are told. Please see Figure 1 for a summary of automatic vs. strategic processing hypotheses.

The Role of Empathy in Set

In addition to examining the mechanism underlying grounded cognition, we also explore a possible correlation between susceptibility to grounded effects (e.g., congruency effects), emotional contagion, motor simulation and empathy. Shared action representations through simulation and the mirror neuron system may facilitate communication, social interaction, learning and understanding. Not only do we simulate the action and movement of others, but we also simulate their emotional state in our mind. It is not such a stretch to also apply grounded motor simulation to the representation of emotions. For example, witnessing a person smile stimulates the muscles and brain regions that would generate a smile for the observer, and this implicit activation generates happy emotions in the witness (Decety, 2007).

Evidence for grounded social embodiment appears at the most basic level through resonance behaviors such as yawning, contagious laughter, and emotional contagion, or an involuntary reproduction and "convergence" (Hatfield, 1994, p. 5) toward another's emotion (Hatfield, 1994; Jeannerod, 2006). Furthermore, such unconscious mimicry occurs with temperament, accents, intonation, speech rate, posture and gesticulations (Decety, 2007; Hatfield, 1994). This type of imitation (e.g. emotional contagion, yawning, and mimicry), however, is rudimentary, and does not allow for the imitator to fully grasp the circumstances surrounding the emotion (Jeannerod, 2006; Decety, 2007).

Theory of mind—the ability to understand that those outside the self have differing beliefs, thoughts, desires, and knowledge (i.e., 'minds') (Saxe & Powell, 2006)—develops in part from the information we receive through imitation. Grounded theory of mind based in modal, embodied simulation and the environment allows us to comprehend others by basing that understanding in our own experience and the intentions/actions associated with those states.

The actual perception of another's mental state within theory of mind is empathy. Empathy is a "sense of similarity between the feelings one person experiences and those expressed by others" (Preston & de Waal, 2002). The "sense of similarity" occurs because of a similar neural representation of a mental state between individuals (i.e., simulation) (Decety, 2007; Decety & Chaminade, 2003; Gallese, 2003). Empathy is "inner imitation" (Lipps, 1903 as cited in Gallese, 2003). Not only can mental states be represented, but affect is also shared. Mehrabian and Epstein (1972) refer to these different forms of empathy as "cognitive role taking and emotional responsiveness."

It is important to note the difference between empathy and sympathy. While empathy is the sharing of a state or emotion, sympathy is the "affective response" to that sharing (Decety &

Chaminade, 2003; Preston & de Waal, 2002). Sympathy, for example, is the “feelings of concern for a distressed or needy person” (Decety & Chaminade, 2003). In short, empathy often causes sympathy, but sympathy is not a necessary result of empathy. Decety (2007) deems this the “empathy paradox,” that is, empathy does not entail acting compassionately on shared feelings, and can be manipulated with amelioration or detriment in mind.

Regardless of how empathy is used, the underlying process is grounded, like theory of mind, using simulation and the environment to elicit another’s state in the self. Both emotional and sensorimotor representations of another’s state play a role in the self’s experience of other (Avenanti et al, 2005). Functional magnetic resonance imaging studies have uncovered that empathy for pain can be intricately tied to affect (Singer, Seymour, O’Doherty, Kaube, Dolan & Frith, 2004). Using Transcranial Magnetic Stimulation, Avenanti and colleagues (2005) discovered that empathy for pain can also include sensorimotor experience. When participants observed needles being inserted into another individual’s hand, electrical potential at that specific hand site in the participants were reduced. Inhibition of MEPs results when people are actually experiencing pain, implying that, in this study, a sensorimotor resonance for the other’s pain was occurring in the participant (separate from an emotion empathizing). It appears that both affective and sensorimotor information result in empathy between individuals (Aventis et al, 2005).

Whether it is a sensorimotor, affective or a combination of both, simulations occurring during empathic perception may be related to other simulatory processing. We hypothesize that empathy and perceptual-motor experience are closely linked through simulation. If action and empathic simulation are similar, our ability in one process may correlate with ability in the other. For example, a ‘high empathizer’ may also be especially adept at mentally representing an observed action. Of primary interest for the experiments here, participants high in empathy may demonstrate very large grounded congruency effects (i.e., they are very ‘good’ or very ‘high’ simulators) to the extent that these effects are automatic. Large congruency effects may also be seen for high empathizers even if grounded congruency effects reflect strategic processing. In a sense, high empathy participants are more “tuned in” to the demands of the experiment and simulate them internally, such that they affect their performance on the judgment task.

To assess this possibility, we administered the Emotional Contagion Scale, Balanced Emotional Empathy, Interpersonal Reactivity Index, the Hogan Empathy Scale, and Tellegen Absorption Scale at the end of the computer based behavioral experiment to obtain a measure within each participant of emotional contagion, empathic perception, and absorption. Simulation and mirroring in perceptual-motor cognition may be related to simulating the emotional states of others in several ways. It may be that they use the same underlying mechanism or network. It could also be that those who are more empathetic are highly oriented toward others and susceptible to their influence and instructions. The relationship may also follow in the opposite direction in that very developed simulation ability causes one to be ‘other oriented.’ Regardless of the cause for the correlation, we hypothesize that if someone is highly empathic, they will also demonstrate high grounded congruency effects, whereas low empathizers will show less or none.

Overview of the Experiments

Using a task developed by Casasanto (2008), we had participants identify the color of words displayed on a computer screen, with these responses requiring either an up or down movement of their arm. When a word appeared in blue ink, for example, participants had to move their arm up to press a button positioned up high, but when the word appeared in red ink, they had to move their arm down to press a low response button. The words presented on the screen were a mix of words with (critical) and without (filler) spatial qualities. For example, some of the words presented had a distinct spatial location like ‘ceiling’ and ‘floor,’ whereas other words did not have spatial qualities like ‘and’ and ‘the.’

Because words had high and low spatial qualities and the response movement was an up or down motion, a congruency effect should occur, previously reported in Casasanto (2008). When overt movement and simulation matched (e.g., ceiling in blue ink), response times should be faster when compared to trials in which simulation and movement were in opposite directions (e.g., ceiling in red ink).

Two experiments used three methods to assess whether grounded congruency effects are automatic: manipulating the ratio of targets to fillers, manipulating the presence vs. absence of irrelevant set, and manipulating awareness of true set. We also explored a possible correlation between grounded congruency effects, emotional contagion, and empathy.

In Experiment 1, we evaluated two of the three set manipulations on up-down congruency effects in an orthogonal between-subjects design: proportion of irrelevant fillers (20% vs. 80%), and presence vs. absence of orthogonal set (open vs. closed class) between participants. The dependent measure was a computer-based button press task in which participants had to make up vs. down responses based on the ink color of the word. The critical issue was whether the proportion of fillers and/or the presence of a false set (open vs. closed class) diminished/eliminated up-down congruency effects.

In Experiment 2, we assessed the third set manipulation, namely, whether awareness vs. no awareness of critical set (up vs. down words) affects up-down congruency effects (with no fillers). Again the aim was to discern whether this manipulation modulated up-down congruency effects.

Experiment 1

In this experiment, we examined whether the manipulation of filler word proportion and orthogonal set awareness interfered with grounded congruency effects. Evidence for interference would suggest evidence for processing with strategic, as opposed to automatic, components.

Method

Design and participants. Experiment 1 used a 2x2 design with two independent variables: ratio of fillers to critical words (20% vs. 80%) and whether a given orthogonal set was present or not (open vs. closed class words). Both variables were manipulated orthogonally across participants and assigned randomly to the four between-participant conditions. Response time and accuracy were measured for each participant. In addition, we also manipulated the color of the response keys between subjects. For half the participants the up key was red and the down key was blue, while the colors were switched for the other half of the sample.

Participants were drawn from a pool of undergraduates at a private university in the southeast. They received course credit for their involvement. Participants included 26 males and 70 females, whose age ranged from 17 to 28 with an average age of 19. The sample was predominately Caucasian (56%), with 13 percent African American, 6 percent Hispanic, 19 percent Asian, and 6 who identified themselves as other. Selection criteria included: fluency in English, normal, and corrected vision. We received informed consent and treated the participants in accordance with the ethical standards of the American Psychological Association.

There were 96 participants total in the experiment, with 24 assigned randomly to each of the 4 between-participant conditions created by crossing the 2 orthogonal factors. Thus, 25% of participants were given the false set (open vs. closed class words), with an 80:20 critical to filler word ratio, 25% were given the false set with a 20:80 critical to filler word ratio, 25% of participants were given no set with an 80:20 critical to filler word ratio, and 25% were given no set with a 20:80 critical to filler word ratio.

Materials. Critical words had associated spatial meanings (e.g., up or down), which we assume activate spatial simulations. 96 concrete open-class words were used as critical items. 48 words had literal (not metaphorical) 'up' spatial associations, and 48 had literal down associations (e.g., bird vs. anchor). All critical words were previously scaled for degree of up/down association (Casasanto, 2008). Ratings for high spatial words ranged from 2.5 to 8.5, and ratings for low spatial words ranged from 2 to 6.7[2]. The lists of high and low spatial words were controlled for number of letters, syllables, concreteness and Kucera-Francis written frequency based on data compiled by the MRC Psycholinguistic Database: Machine Usable Dictionary, version 2.00[3] (Coltheart, 1981). Each experimental condition included one quarter of the critical materials (24 words) in each semantic position (up vs. down) coupled with one of two possible response positions (up vs. down), as described in the procedure section. Pairings between word and response were counterbalanced between participants. An equal ratio of up and down words was present in each condition.

The filler words had no associated spatial meaning, therefore no spatial simulations were expected to facilitate or interfere with response times. Filler items were 384-closed class, non-spatial words. Closed class words are typically highly grammaticized words that are not productive in the sense that they do not continually change and accept the addition of new words, as do nouns, verbs, and modifiers. For example, closed class words include conjunctions like "and," determiners like "the," etc. In conditions with 80% fillers, all 384 items were used, while

---

[2] There was a statistically significant difference between the high vs. low spatial word lists for ratings of spatial location in that high words had more extreme ratings than low words.

[3] http://www.psy.uwa.edu.au/MRCDataBase/uwa_mrc.htm

in the 20% filler conditions, only 24 randomly selected fillers were present. See Appendix A for a complete list of all the critical and filler words.

Words appeared in the center of a computer screen in either red or blue font. Responses were made using a computer keyboard positioned vertically and directly to the right of the computer screen. The end of the keyboard with the letter “Q” was on top and the end with the letter “P” was on the bottom. To stabilize the keyboard, it was strapped to an apparatus consisting of three wooden boards: one resting flat on the table as the base with the long end coming toward the participant, one positioned erect and perpendicular to the base board at the end closest to the participant (creating a 90 degree angle in which the two axis extend up and away from the participant), and the last board positioned as a support between the first two to hold the erect board up in the air. To begin an individual trial, participants had to press the “H” key, which was always colored white for every participant. The “H” key was roughly aligned with the middle of the computer screen so that its position would line up with the position of the presented words. In order to make their word color judgments, participants had to press the “D” or “L” key. For half of the participants the “D” key was colored red and the “L” key was blue, while for the remaining participants the color designations were switched.

Individual difference measures of empathy. We combined three empathy measures, a measure of emotional contagion and a measure of absorption into a single computer-based questionnaire. Empathy scales included (a) the Hogan Empathy Scale (HES), (b) the Balanced Emotional Empathy Scale (BEES), and (c) the Interpersonal Reactivity Index (IRI). The HES focuses on the cognitive role-taking aspects of empathy, which includes the ability to attend to social situations and to be considerate of other’s affective states (Chlopan, McCain, Carbonell, & Hagen, 1985). In Hogan’s original definition of empathy, he discounted the vicarious feelings that may emerge while perceiving another, and specified that the representation of the affective state being observed need not be completely accurate (Hogan, 1969). A factor analysis of the HES uncovered four features: Social Self-Confidence, Even Temperedness, Sensitivity and Nonconformity (Chlopan et al., 1985). The HES uses 31 true-false questions from the California Psychological Inventory, 25 from the Minnesota Multiphasic Personality Inventory and 8 from a Q-sort task developed by the University of California's Institute of Personality Assessment and Research (IPAR). Hogan tested the validity of the HES against Q-sort empathy ratings ($r = .62$) and with a definition of social acuity ($r = .58$). Test-retest reliability was .84 in a sample of undergrads. Outside Hogan's own testing, other researchers have indicated lower reliability scores of .61 (Cross & Sharpley, 1982).

To account for the emotional aspects of empathy, we used the BEES, a revision of the original Emotional Empathic Tendency Scale (Mehrabian & Epstein, 1972). This scale includes 30 questions, answered using a Likert Scale ranging from very strong disagreement (-4) to very strong agreement (+4) (Mehrabian, 2000). Points are totaled, and the higher the score, the higher the individual's empathy. The mean score for males is 29 (*SD*=28), while a mean score for females is 60 (*SD*=21). Because the BEES is highly correlated with the Emotional Empathic Tendency Scale researchers similar assessments of validity are used for both, for example, by measuring the correlation between empathy, aggression and helping behavior. People high in empathy were more affected by the proximity of an individual they were instructed to act aggressively toward. For example, they were less likely to aggress if they could see the victim. They were also more likely to exhibit helping behaviors toward a distressed other. Internal consistency as measured by the coefficient alpha is .87, while the test-retest reliability is .79 (Mehrabian, 2000).

Davis's (1980) IRI is positively correlated with an earlier version of the BEES (Mehrabian, Young & Sato, 1988), but employs an intersection of both cognitive and emotional empathy. The 28 questions are answered on a five-point scale ranging from 0 (does not describe me well) to 4 (descries me very well). Through factor analysis four constructs appear to be measured, resulting in the four scales of the test: fantasy (FS), identification with made up characters/situations; perspective taking (PT), ability to espouse another's standpoint; empathic concern (EC), tendency to experience compassion for suffering others; and personal distress (PD), vicariously experience another's negative event (Davis, 1980). Internal reliability was nearly identical for both sexes on each scale: FS, .78 (males), .75 (female); PT, .75, .78; EC, .72, .70; PD, .78. Test-retest reliability was satisfactorily high for both sexes, ranging from .61 to .79 for males and .62 to .81 for females. Overall women displayed higher average scores on all of the scales when compared to men. In a test of validity, Davis found that the PT scale correlated highly with the HES ($r = .40$) and the FS and EC scales had greater associations with the Questionnaire Measure of Emotional Empathy (Mehrabian & Epstein, 1972) ($r = .52$, $r = .60$).

Combined with the empathy measures, we also included Doherty's (1997) Emotional Contagion Scale (EC). The scale itself includes 15 items that measure "mimetic tendency" (Doherty, 1997, p. 149) in relation to five different emotions: happiness, love, fear, anger, and sadness. The test can be further broken down into both a positive and negative subscale. Participants responded using a five-point Likert scale (never, rarely, usually, often, always) as to how fully they agreed with the test question. The EC is correlated with the vicarious emotional experience measures on Mehrabian and Epstein's (1972) Measure of Empathic Tendency ($r = .47$, $p <.05$) and the aforementioned IRI ($r = .37$, $p <.05$). Construct validity was also demonstrated with a strong correlation between emotional contagion, reactivity, "affective orientation, emotionality, sensitivity to others and self-esteem" (Doherty, 1997, p. 149). Doherty found a test re-test reliability of $r(41) = .84$, $p < .001$, and an overall internally consistent reliability rating of Cronbach's $\alpha = .90$.

The final questionnaire measure, the Tellegen Absorption Scale (TAS), is an appraisal of how prone someone is to switch to more self-privileged modes of experience, e.g. imagination, fantasizing etc. (Kremen & Block, 2002; Tellegen & Atkinson, 1974). It is also associated with "experiential involvement and alterations in attention" (Kremen & Block, 2002). The TAS includes 34 true/false questions. Reliability as measured by Kremen and Block (2002) is fairly high (coefficient alpha = .87 with population of 18year olds).

Procedure. All participants completed a consent form and then read written instructions for the experiment presented on the computer. Participants were required to explain the task back to the experimenter to ensure they completely understood the assignment.

The task was to indicate whether the font of words presented on the computer screen was either red or blue by pressing the corresponding red or blue key on the keyboard. Each trial of the experiment began with a black fixation cross (+) in the middle of the screen. To initiate the trial, participants had to press and hold down a white "home key" on the keyboard. Once they pressed (and held) the white key, there was a 500ms pause, and then the colored word would appear. Once the word appeared, participants released the white key, and pressed the corresponding red or blue key, depending on the color of the presented stimulus. Immediately after they made their color judgment by pressing the red or blue key on the keyboard, the screen returned to the fixation cross, and participants returned to holding down the white key. This procedure repeated itself for every trial until participants completed all of the trials. Participants were unaware that there were different spatial heights for the critical words or that the word's meaning could be congruent or incongruent with the height of the response key (i.e., nothing was mentioned explicitly). Brief 30-second breaks were built into the procedure (eight for high filler conditions and two for low filler conditions) to provide participants with an opportunity to rest their arm.

There was one important exception to the red vs. blue color judgment task. At random points during the experiment, the word "emory" appeared. Participants were instructed to always press the white key when they saw the word "emory," regardless of the font color. This was done in an effort to ensure participants were actually reading the words, and not just perceiving the font color and making a judgment without semantically processing the word meaning.

Trials were congruent, incongruent or neutral. All three trial types made it possible to assess grounded congruency effects. Congruent trials involved pressing the top keyboard button when the spatial height of the word was high (e.g., "sky") or the bottom button when the word had low spatial height (e.g., "basement"). Incongruent trials interfered with grounded congruency effects. Incongruent trials required participants to press the high button when the word implied a low spatial location (e.g., when the word presented was "basement"), or to press the low button when the word implied a high location (e.g., "sky"). Neutral trials involved filler words. Due to their lack of spatial qualities, no facilitation or interference should be elicited (e.g., responding with the top (or bottom) keyboard button when the word was "of").

There were two proportions of fillers used: a high filler proportion condition and a low filler condition. For half of the participants 80% (96) of the items were critical words, while 20% (24) were fillers (low filler condition), while the ratio was reversed for the other half of the participant pool, 96 critical words vs. 384 fillers (low filler condition). This manipulation was crossed with the instructions manipulation below. Nothing was said to participants about the proportion of filler to critical words in this experiment.

For consistency, all of the participants were told that the words they would see came from a large online database of millions of sentences. However, the main reason for this information was to account for why the closed class words would occur more often than the open class words in the high filler proportion condition of the experiment. For half of the participants, this was the only information they received about the words. For the remaining half of the participants, they were told that the presented words included both open and closed class words (crossed with the filler manipulation above). The definitions and examples of each type of word class were as follows: Examples of open class words include: "cat, to run, person, to Google, twitter, to tweet, swine-flue, to supersize, and mouse-potato." An open class word accepts new words into its category. For example, new nouns and verbs (like *"to Google"*) are constantly being incorporated into the English language. In contrast, a closed word class is resistant to the addition of new words. Examples of closed class words include: prepositions like "of, up, under, above, near," conjunctions like "and, but, or, no, so," and determiners like "a, the." It is very rare that new prepositions, conjunctions, or determiners are added to the English language; hence they are considered "closed" word classes.

Each trial was self-paced by the participant. Each participant completed either 124 (120 regular plus 4 "emory") trials (20% fillers) or 496 (480 regular plus 16 "emory") trials (80% fillers), including 41 (40 regular and 1 "emory") additional practice trials beforehand. Practice trials were the same as the experimental task, except they used different filler and critical spatial words than those presented in the actual task. The practice set also contained the same ratio of fillers to critical items that was present in the subsequent experimental manipulation.

After finishing the word-color judgment task, participants completed a "Personality Questionnaire," which included the three empathy measures, the emotional contagion scale, and the absorption scale. Participants were *not* told the questionnaire was a measure of empathy, emotional contagion, and absorption. Finally, participants were debriefed and given extra credit for their respective class requirements.

Results and Discussion

Congruency effect calculations. We calculated congruency effects based on both error rates and response time. To reiterate, a congruency effect is faster response times (and/or fewer errors) when mental simulation of spatial qualities and response direction match (congruent trials), and slower response times (and/or more errors) when simulated spatial location and response movement do not match (incongruent trials). In our experiment, congruent trials included instances where the critical word had an associated high spatial location and the response required an upward movement or when the critical word had a low spatial location and a downward response was required. Incongruent trials involved a high spatial word paired with a downward movement or a low spatial word paired with an upward response movement.

To calculate congruency effects we *initially* used the following formula: [(incongruent_high_words – congruent_high_words) + (incongruent _low_words – congruent_low_words)]/2. We subtracted congruent from incongruent in order to yield a positive number because we anticipated that incongruent trials would produce a longer response time and/or more errors. We separated calculations for high and low words in case there were any discrepancies between the two groups (high vs. low words). By separating the high and low words we essentially calculated a congruency effect for high words and a congruency effect for low words, combined them, and then divided by two to get the average of the two congruency effects (i.e., a complete congruency effect including both high and low words).

Response time calculations. Using the aforementioned congruency effect formula we calculated an average congruency effect for each of the four experimental conditions (filler proportion crossed with awareness of false set) for the full response time (word onset to button press). See Figure 2. Only one of the conditions produced a significant congruency effect, namely, the experimental condition with low percent fillers and no awareness of false set ($t(23) = 2.31$, $p = .015$). This condition yielded congruent response times that were significantly faster than incongruent response times. Furthermore, the patterns across the four conditions did not follow any of our predictions for the manipulations of filler proportion and awareness of false set manipulations, as described earlier. See Figure 1 for comparison to “ideal” results based on automatic and strategic processing.

Thoroughly puzzled, we separated our “complete” congruency effect into its component congruency effects: one for high spatial words and one for low spatial words, to see if one type of critical word might be demonstrating a significant congruency effect. See Figure 3 for congruency effects broken down by high and low spatial words. Here we discovered that for nearly all the conditions (except high proportion of fillers and no awareness of false set), high spatial words demonstrated a significant congruency effect. For high spatial words congruent trials were, on average, significantly faster than incongruent trials. Interestingly, low spatial words showed the opposite effect, an “incongruency effect,” if you will. For low spatial words incongruent trials were, on average, faster than congruent trials (the opposite of what one would expect). However, only one condition (low proportion of fillers and awareness of false set) had an “incongruency effect” significant from zero.

Based on these results, it looked as if there might be a “motor effect”: upward responses appeared to always be faster than downward responses, regardless of whether the trial was congruent or incongruent. To check this hypothesis we computed a contrast that directly assessed whether there was bias in response direction, specifically: [(down_response_high_words – up_response_high_words) + (down_response_ low_words – up_response_low_words)]/2. As can be seen, this contrast assesses whether, for the same words (high or low), down responses are slower than fast responses. As Figure 4 illustrates, this direct assessment of a motor response bias yielded the same results as our informal observations of this bias across component congruency effects in Figure 4.

Thus, one thing that we uncovered in this experiment, not reported in previous publications, is that there is a motor bias in this kind of congruency task. Specifically, downward movements are significantly slower than upward movements. It may be that something about the human musculature makes it easier to make these upward movements.

To control for these motor effects, we needed to use an alternate formula for computing congruency effects that takes into account the different speeds for making up vs. down movements. To control for movement direction, we used the following “alternate congruency effect” formula for the remainder of our calculations: [(Up_response_low_words – Up_response_high_words) + (Down_response_high_words – Down_response_low_words)]/2 or written with respect to congruency like the initial formula: [(incongruent_low_words – congruent_high_words) + (incongruent_high_words – congruent_low_words)]/2. As can be seen from the first version of this formula, each half controls for a movement direction while seeing if words incongruent with this motion are slower than words congruent with it. As can further be seen, these contrasts no longer control for words, given that different words are now being compared within each half of the formula (e.g., high words vs. down words for up responses). Nevertheless, it was acceptable to use this formula because the two groups of words (high and low) were well controlled, with no evidence for a significant difference between the two based on the number of letters, syllables, concreteness and written frequency. Thus, there should not be any bias associated with processing different groups of words in assessing the comparisons critical to this contrast.

Even when using the new alternate congruency formula, however, only two conditions (awareness of false set paired with high percent fillers, and no awareness of false set paired with low percent fillers) demonstrated congruency effects that were significant from zero (Figure 5) with $t(23) = 2.31$, $p = 0.015$, $d = .47$ and $t(23) = 2.08$, $p = .025$, $d = .42$ respectively. Only in these two experimental groups did participants demonstrate significantly faster response times for congruent trials (e.g. high spatial word and upward response direction) and slower times for incongruent trials (e.g. high spatial word paired with downward response). See Table 1 for a summary of average response times for congruent and incongruent trials across experimental condition. While significant, differences in milliseconds were quite small, and the effect size was fairly moderate. Further, none of the results logically followed the predictions we had made for the manipulations of filler proportion and awareness of false set based the presence of strategic vs. automatic processing. If processing was automatic, there should have been no evidence for a difference between any of the conditions. However, if the mechanism was strategic, the conditions with high proportions of fillers and/or an orthogonal set should have had an attenuated or no evidence for a congruency effect (Figure 1). In reality, we could not fully evaluate these hypotheses because we did not find significant congruency effects in all four conditions.

In further searching for congruency effects across all four conditions, and of larger magnitude, we broke down each full response time into its component response times: stimulus onset to button release, and button release to button press. We then conducted the same "alternate congruency effect" calculation as we performed on the 'full response times.' As Figure 6 illustrates, neither of the component response times had significant congruency effects in any condition, and all of the differences between conditions were very small in magnitude. See Table 2 and 3 for a summary of average response times across condition for congruent and incongruent trials.

Even when we broke the congruency effect into its constituent congruency effects (one for upward responses and one for downward responses) only a few conditions had significant effects (Figure 7). In the first component response time, stimulus onset to button release, the condition with awareness of false set and a high proportion of fillers was the only significant congruency effect, and it was only significant for downward responses: $t(23) = 1.74$, $p = .048$. For our second component response time, button release to button press, only the condition with no awareness of false set and a low proportion of fillers had a significant congruency effect, and again, it was only significant for downward responses: $t(23) = 1.87$, $p = .037$. Again, we found these results quite puzzling with no apparent conclusions to be drawn regarding our hypotheses because we did find significant congruency effects across all four conditions. It is also worth noting again that all significant effects that did occur were very small, always being less than 10 milliseconds.

Across all three response time calculations (full and two component), only a handful of experimental conditions demonstrated significant, very small, congruency effects. The majority of the time participants were not significantly faster on congruency trials when compared to incongruent trials. No clear picture can be drawn regarding whether these congruency effects are based in strategic or automatic processing because we did not consistently elicit a congruency effect. This may have to do with the fact that the task (color judgment) did not force the participants to reliably process the meaning of the words. Deep, simulatory processing may be necessary to demonstrate congruency effects. We explore these conjectures further in the general discussion.

Error rate calculations. We calculated the percent error for each participant and replaced participants who exhibited greater than 15% error. Though somewhat higher than typical, we used 15% as our cutoff because of the nature of the experimental task. Often participants moved in the correct response direction, but hit an adjacent, 'incorrect' button on the keyboard. To account for higher error rates due to these responses we relaxed our cutoff criteria (to 15%), and also calculated "true errors." True errors only included inaccuracies in response direction. Responses in the correct direction, but pressing an adjacent button were coded as 'correct.' See Figure 8 for a summary of percent error and percent true error across experimental conditions.

We also calculated a congruency effect based on true error rates for each experimental condition (see Table 4 for a summary of average error rates for incongruent and congruent trials across experimental condition). We used the same alternate congruency effect formula (controls for response direction) with our true error rate that we used for the second version of the response time congruency calculations: [(Down_response_high word – Down_response_low word) + (Up response_low_word – Down_response_high_word)]/ 2. An independent sample t-test revealed none of the conditions had significant congruency effects. Error rates for congruent trials were not significantly less than those for incongruent trials. Please see Figure 9 for a summary of results.

Critical word item analyses. Still on the hunt for significant congruency effects we decided to search at the word item level. We compiled an average response time for up responses and an average response time for down responses for each critical word across all participants. We subtracted average up response time from average down response time to obtain an average congruency effect for each word. All of the words were previously normed for spatial height (Casasanto, 2008). Theoretically, words rated especially high or especially low (i.e., that is having the most extreme spatial locations) should affect response times more than words that have less extreme spatial locations. Words with the highest and lowest spatial locations should cause the largest congruency effects. We conducted a Pearson correlation test between the normed spatial word heights for high and low critical words and the congruency effect corresponding to each individual word. Correlations between congruency effects and normed height for low spatial words were analyzed separately from high spatial words so that we would not have to account for motor effects from response direction (which affect high and low spatial words conversely). There were no significant correlations between word height and congruency effect, that is, words rated as spatially highest or lowest did not correspond with more robust congruency effects (high words: $r(46) = -.01$, $p = .961$; low words: $r(46) = -.15$, $p = .315$). Again, this may be because participants were not processing the stimuli deeply and not simulating the spatial locations associated with each word.

Correlation between empathy scales and congruency effects. Once we calculated the congruency effects, we analyzed their correlation with the five empathy, emotional contagion and absorption scales. We used full response time because it had the most significant congruency effects. A Pearson correlation test revealed that there was no significant correlation between any of the empathy scales (BEES, EC, HES, IRI, TAS) and congruency effects for full response time (*r*(94) = .03, *p* = .521). Participants high in empathy did not exhibit high congruency effects or vice versa. We did not examine correlations between congruency effects for error rate and the empathy scales because error rates did not have any significant congruency effects. Even though we did not uncover a significant correlation between the five scales and spatial simulation, we feel this may have more to do with lack of congruency effects rather than lack of a correlation between the underlying mechanisms of the two processes. Participants may not have reliably used deep, simulation based processing for the task, and we cannot definitively rule out a possible correlation between empathic ability and tendency to simulate.

Emory error rate calculations. On “emory trials,” participants were instructed to always press the white key regardless of the font color of “emory.” We examined the error rate on trials to see if participants were actually reading the words, not just perceiving the font color and making a judgment without semantically processing the word meaning. All four conditions, high proportion of fillers and no false set, low proportion of fillers and no false set, high proportion of fillers and false set (open vs. closed class words), low proportion of fillers and false set, had error rates that were less than 5% (2.3%, 4.2%, 3.1%, 3.1%, respectively). This suggests participants were, to some extent, reading the words.

Experiment 2

In this experiment, we again attempted to replicate previous studies that demonstrate evidence for congruency effects with color judgment (e.g. Casasanto, 2008). In addition, we also examined whether explicitly drawing participant’s attention to the spatial qualities of the words might facilitate congruency effects, as described in the introduction.

Method

Design and participants. Experiment 2 used a randomly assigned independent groups design with the following independent variables: word height (high vs. low), response height (up vs. down) and awareness (vs. no awareness) of the critical set in the experiment. On half the trials, word height and response height matched, while on the remaining trials, they were incongruent. Awareness of the critical set in the experiment was manipulated across participants. Response time and accuracy measurements were recorded for each trial. In addition, we counterbalanced the color of the response keys (red vs. blue in top vs. bottom location) between subjects.

Participants were 48 undergraduates at a private university in the southeast, participating for course credit. There were 12 males and 36 females, ages 17 to 34, with an average age of 19. The sample was predominately Caucasian (59%), 8 percent African American, 6 percent Hispanic, and 27 percent Asian. Fluency in English and normal or corrected vision were the only selection criteria invoked. We received informed consent and treated the participants in accordance with the ethical standards of the American Psychological Association. There were 24 participants in each of the two conditions (awareness vs. no awareness of set).

Materials. The same materials and response apparatus were used as in Experiment 1, with the exception of the filler materials (i.e., filler words, open and closed class word instructions). No fillers were used in this experiment, only critical words with up vs. down spatial qualities. The same 96 concrete, open classed, critical words from Experiment 1 were used in Experiment 2. There were 48 with literal up associations, and 48 with literal down associations. All critical materials were used in each condition. An equal ratio of up to down words was always present. The same empathy, emotional contagion and absorption scales were used as in Experiment 1.

Procedure. Participants followed nearly the same procedure as Experiment 1. All participants completed a consent form and then read written instructions for the experiment presented on the computer. Participants were required to explain the task back to the experimenter to ensure they completely understood the assignment.

Participants simply had to indicate the color of the font for words presented on the computer screen (red or blue) by pressing the corresponding keyboard key. Each trial of the experiment began with a black fixation cross (+) in the middle of the screen. To initiate the trial, participants had to press and hold down the white “home key” on the keyboard. Once they pressed (and held) the white key, there was a 500ms pause, and then the colored word would appear. Once the word appeared, participants released the white key, and pressed the corresponding red or blue key, depending on the color of the presented stimulus. Immediately after they made their color judgment by pressing the red or blue key on the keyboard, the screen returned to the fixation cross, and participants returned to holding down the white key. This procedure repeated itself for every trial until participants completed all of the trials. Trials were either congruent or incongruent. There was no “neutral” trial condition because this experiment did not include any filler non-spatial words. Two 30-second breaks were built into the procedure so participants could rest their arm.

Three “emory” manipulation check trials were also randomly interspersed in each condition. On these trials, regardless of the color of the word “emory,” participants had to press the white key. These trials were included to make sure participants were semantically processing the words.

All of the participants were told that the words they would see came from a large online database of millions of sentences. However, for half of the participants, this was the only information they received about the words (replication of previous studies). For the remaining participants, they were specifically told about the high and low spatial qualities of the words as follows: "The words you will see all have spatial qualities. For example, some words will imply a low spatial location like "tunnel, culvert, dunk, grub, short," while others will imply a higher location like "falcon, above, balcony, soar, top.""

Each trial was self-paced by the participant. Each participant completed 99 trials (96 regular plus 3 "emory"), including 41 (40 regular and 1 "emory") additional practice trials beforehand. Practice trials were the same as the experimental task, except they used different critical spatial words than those presented in the actual task.

After finishing the colored word task, participants completed the same "Personality Questionnaire" from Experiment 1. Again, participants were *not* told that the questionnaire was a measure of empathy, emotional contagion, and absorption. Finally, participants were debriefed and given extra credit for their respective class requirements.

Results and Discussion

Congruency effect calculations. We calculated congruency effects based on both error rates and response time in the same manner as for Experiment 1. Again, a congruency effect involves faster response times (and/or fewer errors) when mental simulation of spatial qualities and response direction match up (congruent trials), and slower response times (and/or more errors) when simulated spatial location and response movement do not match (incongruent trials).

Response time calculations. Using our preliminary congruency effect formula, we calculated an average congruency effect for the two experimental conditions (awareness or no awareness of the true set: high and low spatial words) for the full response time (stimulus onset to button press). See Figure 2. Again, there were no fillers in this experiment in an effort to replicate previous incidents of congruency effects with this paradigm and to test for a possible facilitation effect when participants are aware of the spatial qualities of the words. None of the conditions produced a significant congruency effect, that is, participants were not significantly faster for congruent trials when compared to incongruent trials.

When we separated our "complete" congruency effect into its component congruency effects: one for high spatial words and one for low spatial words, we found that high spatial words did demonstrate significant congruency effects both when participants were aware of the true set ($t(23) = 2.88$, $p = .004$), and when they were not ($t(23) = 3.04$, $p = .003$). Conversely, low spatial words caused an "incongruency effect" (incongruent trials were significantly faster than congruent trials) in both conditions ($t(23) = -3.37$, $p = .001$ and $t(23) = -2.65$, $p = .007$, respectively). See Figure 3 for congruency effects broken down by high and low spatial words.

As in Experiment 1, these results offer further evidence for a motor effect in which upward responses appeared to always be faster than downward responses, regardless of whether the trial was congruent or incongruent. Again, we collapsed the component congruency effects (one for high spatial words and one for low) across condition and assessed the potential bias of movement direction directly). The results obtained for collapsed congruency effects and independent calculations based on response direction (Figure 4) were identical. Once again, we discovered evidence for a previously unreported motor effect in this paradigm in which downward movements were always slower than upward movements.

To control for these motor effects, we used the same alternate congruency effect formula from experiment 1. Again, this formula controls for the influence of response direction (i.e., up responses are faster than down responses), but does not explicitly control for word spatial height. Because the two groups of words (high vs. low spatial words) were not significantly different in their number of letters, syllables, concreteness or written frequency, it was acceptable to construct the alternate formula in this way.

None of the calculations using the alternate congruency formula yielded significant results (Figure 5 - 7). Neither of the component congruency effects (one for up responses and one for down responses) were significant. Further examination of the component response times (stimulus onset to button release and button release to button press) produced the same null results. See Tables 1-3 for a summary of average response times (full and component) for congruent and incongruent trials across condition. Thus, we were unable to replicate previous demonstrations of congruency effects (e.g., Casasanto, 2008) in two experimental conditions that theoretically should have yielded the most robust effects of any of the conditions in experiment 1 or 2. Given this non-replication and lack of significant congruency effects, we were unable to evaluate whether automatic or strategic processing underlies grounded congruency effects.

Error rate calculations. We used the same 15% error cutoff in experiment 1 as experiment 2, replacing participants with higher error rates. Again we calculated "true errors" namely, responses in the wrong direction, while coding instances of hitting an adjacent button on the keyboard, but in the correct direction as correct. See Figure 8 for a summary of percent error and percent true error across the two experimental conditions.

We also calculated a congruency effect based on true error rates for each experimental condition, using the same alternate congruency effect formula for true error rate that we used for response time congruency calculations. See Table 4 for a summary of average error rate for congruent and incongruent trials across experimental condition. Similarly to experiment 1, none of the conditions had significant congruency effects: awareness of true set: $t(23) = 0.003$, $p = 0.50$, and no awareness of set: $t(23) = 0.001$, $p = 0.50$. Error rates for congruent trials were not significantly less than those for incongruent trials (Figure 9).

Critical word item analyses. As in experiment 1, we conducted the same congruency effect analysis at the critical word item level in experiment 2, namely, correlating the normed spatial word heights for high and low critical words with the congruency effect corresponding to each individual word. Correlations between congruency effects and normed height for low spatial words were analyzed separately from high spatial words so that we would not have to account for motor effects from response direction (which affect high and low spatial words conversely). There were no significant correlations between word height and congruency effect, that is, words rated as spatially highest or lowest did not correspond with more robust congruency effects (high words: $r(46) = -.01$, $p = .961$; low words: $r(46) = -.15$, $p = .315$). Again, this may be because participants were not processing the stimuli deeply and not simulating the spatial locations associated with each word.

Correlation between empathy scales and congruency effects. Because we did not find any significant congruency effects for response time or error rates we did not conduct a correlation analysis between the congruency effects and the five empathy related scales.

Emory error rate calculations. “Emory trials” were examined in the same manner as experiment 1 to assess whether participants were semantically processing the words. Both conditions, awareness of the critical manipulation (high vs. low words), and no awareness of verticality, had error rates that were less than 5% (4.2%, 1.4%, respectively). This suggests participants were, to some extent, reading the words.

General Discussion

Across six different conditions, 144 participants in two experiments, we were unable to consistently replicate grounded congruency effects demonstrated previously in the literature (e.g. Casasanto, 2008). In all but two conditions (no awareness of false set paired with low percent fillers, and awareness of false set paired with high percent fillers) participants were not significantly faster for congruent trials compared to incongruent trials. In addition, participants did not make fewer errors on congruent trials compared to incongruent trials.

Based on our original hypotheses, if grounded congruency effects involve automatic processes, participants should have been faster in congruent trials compared to incongruent trials in every condition in both experiments. The congruency effect may have even been more robust in experiment 2 when participants were aware of the spatial qualities of the words because this knowledge may facilitate simulation of the spatial location of the critical words. However, if grounded congruency effects emerge from strategic processing, the results would have been more varied (Figure 1). In conditions with low percentages of fillers (20% to zero), the congruency effect should have been present. A high percentage of fillers would have masked the critical word set (and their spatial qualities); at the very least attenuating, if not completely wiping out the congruency effect. Awareness of false set (open and closed class words) should have influenced the congruency effect similarly to a high percentage of fillers. When participants were given a false set to attend to it should have distracted them from the true set in the experiment (high and low words), either attenuating or removing the grounded congruency effect. Because we did not consistently find a congruency effect in either response time (full and component) or error rate we cannot evaluate whether these effects in relation to our strategic and automatic mechanism hypotheses. In the same way, we cannot assess whether a correlation exists between empathic perception (simulation of other minds) and congruency effects (simulation of semantic spatial properties).

Why were we unable to replicate these grounded congruency effects? In our paradigm participants controlled when each trial began, while in Casasanto's (2008) study, words appeared at a fixed interval of 2000ms. This difference may have played a role in our non-replication, however, we did use the exact same stimuli and color-judgment task as Casasanto (2008).

It may be that we did not replicate Casasanto's (2008) findings because of the nature of the color-judgment experimental task itself. Identifying the font color of a word does not necessarily require one to read and understand the word being presented. One could just perceive the color, press the corresponding button on the keyboard, and not process the word meaning at all. In order for the spatial qualities of words to be activated through simulation, much deeper processing of the word may be necessary. Color-judgment may not require this deep processing, and consequently, this task will not reliably elicit situated simulation or the grounded congruency. Casasanto's (2008) findings may reflect a false positive for congruency effect, owing to an auspicious sample.

The conditions necessary for people to use deep, simulatory processing are explicated in the Language and Situated Simulation (LASS) Theory proposed by Barsalou, Santos, Simmons, and Wilson (2008). In the LASS Theory, both the linguistic and simulatory system are heavily engaged in conceptual processing. Linguistic processing, for the most part, involves just language. For example, if you see the word "apple," you might initially think "red," "fruit," "tart." Situated simulation denotes processing of objects and actions invoked in context. If you observe someone kicking a ball you not only simulate part of a kicking motion, but you also most likely construct a situation in which the ball is kicked, for example your last soccer match. Likewise, if you are simulating the concept of a 'giraffe' you do not imagine it floating in space; instead it is placed in a particular scene based on your own experience with giraffes, perhaps involving how the fur feels to the touch, and the smell of the feed at the zoo.

Linguistic processing tends to be activated at a much faster rate than other forms of processing, and can be classified as 'superficial' compared to the much deeper processing that occurs when accessing conceptual information through situated simulation (Barsalou et al, 2008). The two systems work in tandem, often influencing responses from the other. Using the above example, if someone says the word apple you may generate a few quick linguistically prompted thoughts like 'fruit' and 'red,' but then as you begin to process 'apple' deeper, you perhaps simulate the last time you ate an apple. Thoughts of how sweet it was may trigger a few linguistically generated concepts like 'sugary' and 'crunchy,' and so on as the interaction between the two systems continues in a ripple effect.

It is important to note, in some situations, these knowledge systems may be operating discretely, but in others, the situated simulation and linguistic systems may be in use simultaneously (Simmons, Hamann, Harenski, Hu, & Barsalou, 2008). It can take up to 100ms longer for simulations to reach consciousness (Solomon and Barsalou, 2004). Perhaps both linguistic and simulation cognitions arise concurrently, but the executive system chooses based on the necessary depth of processing for completion of the task at hand to attend to one or the other. As a result, whichever system has the allocated attention controls the responses to the task (Simmons et al, 2008). When time is limited and the task is relatively easy, the linguistic system will dictate processing, whereas more difficult tasks, requiring longer processing time and the involvement of personal experience, are completed using the simulation system (Barsalou et al, 2008; Simmons et al, 2008). Considerable brain imaging evidence supports the LASS Theory. In one exemplary fMRI study, Simmons and colleagues (2008) found no overlap in neural activity for a task that involved superficial word association task (language processing), and generating a situation (situated simulation processing).

It is possible that our paradigm using a color judgment task may not have required enough deep processing to involve the situated simulation system, and consequently we did not observe any grounded congruency effects. The spatial numerical association of response codes (SNARC) effect literature demonstrates similar variability demonstrating congruency effects with a color-judgment task. The basic SNARC effect paradigm requires participants to make an irrelevant judgment about a number stimulus (parity, orientation etc.) (Dehaene, Bossini, & Giraux, 1993). Response times for small numbers are faster with the left hand, while response times for larger numbers are faster when the task is completed with the right hand (Dehaene, Bossini, & Giraux, 1993; Keus & Schwarz, 2005). It appears participants simulate a mental number line, and are faster to respond when their movement matches (is congruent) with their simulation (e.g. a large number is presented and they must respond with their right hand). When their movement is incongruent with their simulation (e.g. large number present and left hand response required) response times are slower.

While Keus and Schwarz (2005) were able to elicit the SNARC effect when participants simply judged the color of the presented number, Fias, Lauwereyns, and Lammertyn (2001) failed to find the effect with a color judgment task. Fias, Lauwereyns and Lammertyn (2001) hypothesize that there is a lack of overlap between the processing areas for magnitude (parietal) and color (ventral stream), and consequently, little processing interference occurs. Without processing interference the SNARC effect vanishes. The depth of processing may also play a part in their failure to elicit the SNARC effect with the color judgment. The ease of judging color makes it a superficial task involving shallow processing. Participants may not be simulating the number line because the task is so easy, and deeper, simulatory processing is unnecessary to complete the assignment.

It is necessary that we address the limitations of our study (i.e., depth of processing) with a series of follow up experiments that force participants to deeply process the semantic qualities of the words. Demonstrating a lack of congruency effect with shallow processing and the existence of the effect with deeper processing will provide further support for the LASS Theory, and prevent a serious misconception in the literature if previous paradigms involve chance occurrences of congruency effects (e.g. Casasanto, 2008). For example, if we require participants to indicate whether a word is concrete or abstract before they make a color-judgment this may spur deeper, simulatory processing. Making an abstractness judgment will force the participant to think not only about what the word means, but also the nature of its instantiation in the world (concrete object vs. abstract concept). Simulation (which may include the spatial location of the word) may be necessary to answer the abstractness question. Participants would make the font color judgment of the word directly after judging its abstractness. In this way, they would still be deeply processing the word, and its associated spatial qualities might affect how fast or how many errors they make when responding (congruency effects).

If judging how abstract a word is does not elicit congruency effects, it may be necessary to explicitly draw attention to the spatial qualities of the words for those qualities to influence response times and error rates. Instead of judging abstractness, we could have participants judge whether the words have spatial qualities. After indicating whether the word stimulus has spatial qualities, then participants would make the color-judgment. Theoretically, the associated spatial locations of the words would still be active during the color-judgment task, and would influence response times and possibly error rates.

If we find congruency effects using tasks that require deeper processing of the words, we can then confidently assess the degree of automaticity involved by adding a high ratio filler words to the task as per our initial experiment. These subsequent studies will also allow us to further explore the motor effects (i.e., upward movements are always faster than downward responses), results unreported in any other experiments.

Tables

Table 1.

Congruent and Incongruent Trial Averages by high and low spatial words for full Response Time (ms)

Expt	Instruction	Filler	Congruent High	Incongruent Low	Incongruent High	Congruent Low
1	No	High	682.48	681.21	697.86	695.41
1	No	Low	700.54	705.85	734.85	719.34
1	O/C	High	686.65	690.26	704.14	690.69
1	O/C	Low	728.00	722.38	752.07	748.54
2	H/L		682.79	683.75	718.42	718.72
2	No		736.63	743.05	774.08	774.64

Table 2.

Congruent and Incongruent Trial Averages by high and low spatial words for Stimulus Onset to Button Release (ms)

Expt	Instruction	Filler	Congruent High	Incongruent Low	Incongruent High	Congruent Low
1	No	High	311.00	313.22	304.70	306.50
1	No	Low	373.78	379.73	375.75	369.46
1	O/C	High	343.23	342.72	340.69	333.01
1	O/C	Low	331.92	330.71	328.54	328.52
2	H/L		335.34	340.96	333.64	329.89
2	No		350.20	347.07	343.47	342.70

Table 3.

Congruent and Incongruent Trial Averages by high and low spatial words for button release to button press (ms)

Expt	Instruction	Filler	Congruent High	Incongruent Low	Incongruent High	Congruent Low
1	No	High	371.83	369.11	393.16	388.93
1	No	Low	326.77	326.12	359.10	349.88
1	O/C	High	343.16	347.54	363.45	357.81
1	O/C	Low	396.08	391.70	423.53	420.02
2	H/L		347.45	342.79	384.78	388.83
2	No		386.42	395.98	430.61	431.93

Table 4.

Congruent and Incongruent Trial Averages by high and low spatial words for True Error Rate.

Expt	Instruction	Filler	Congruent High	Incongruent Low	Incongruent High	Congruent Low
1	No	High	.00	.01	.02	.01
1	No	Low	.00	.01	.01	.02
1	O/C	High	.01	.01	.03	.03
1	O/C	Low	.01	.00	.01	.01
2	H/L		.01	.00	.03	.01
2	No		.01	.00	.01	.01

Figures

Figure 1. 'Ideal' hypothesized results for automatic vs. strategic processing

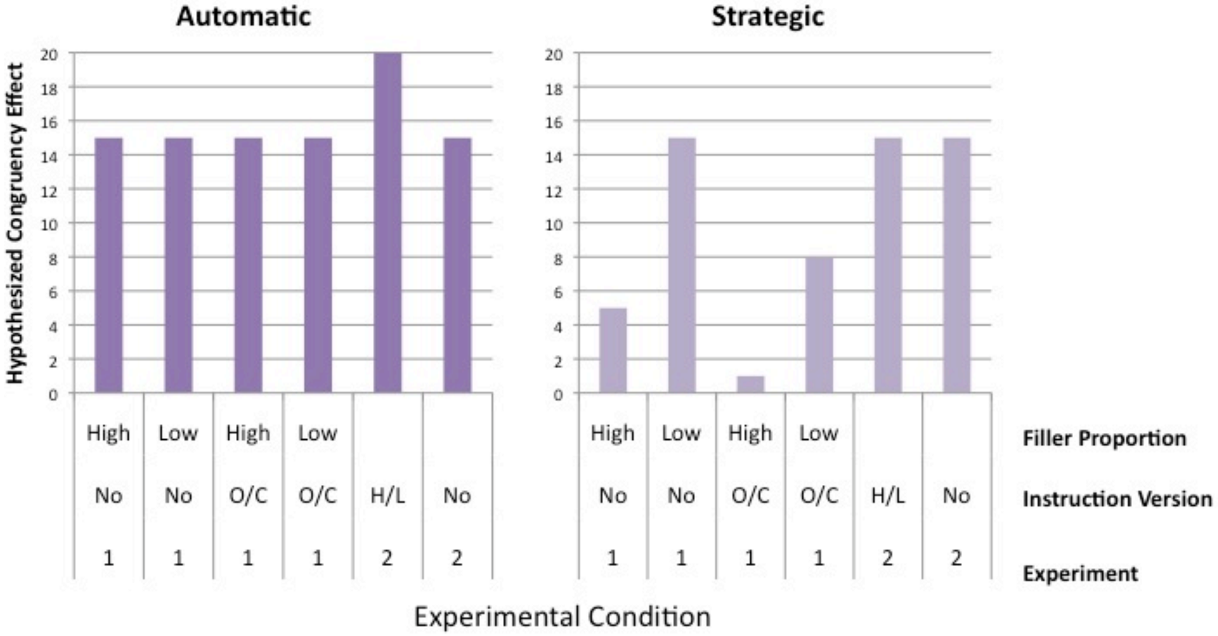

Figure 2. Original congruency effect for full response time.

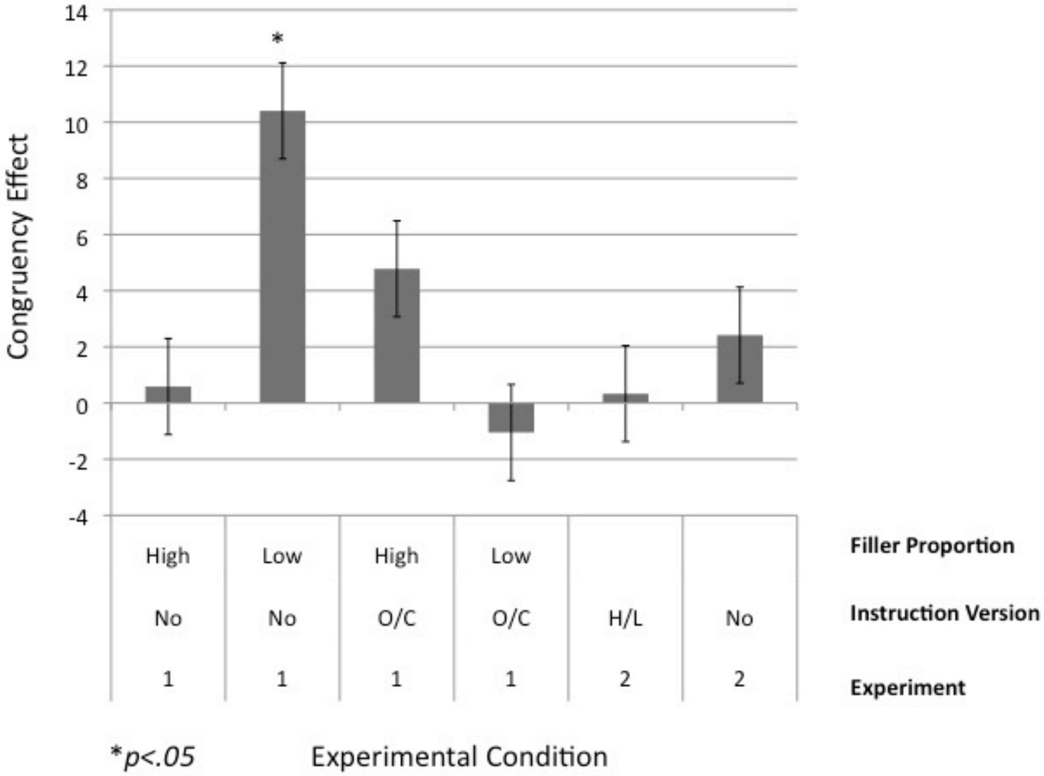

Figure 3. Original congruency effects broken into high and low spatial words

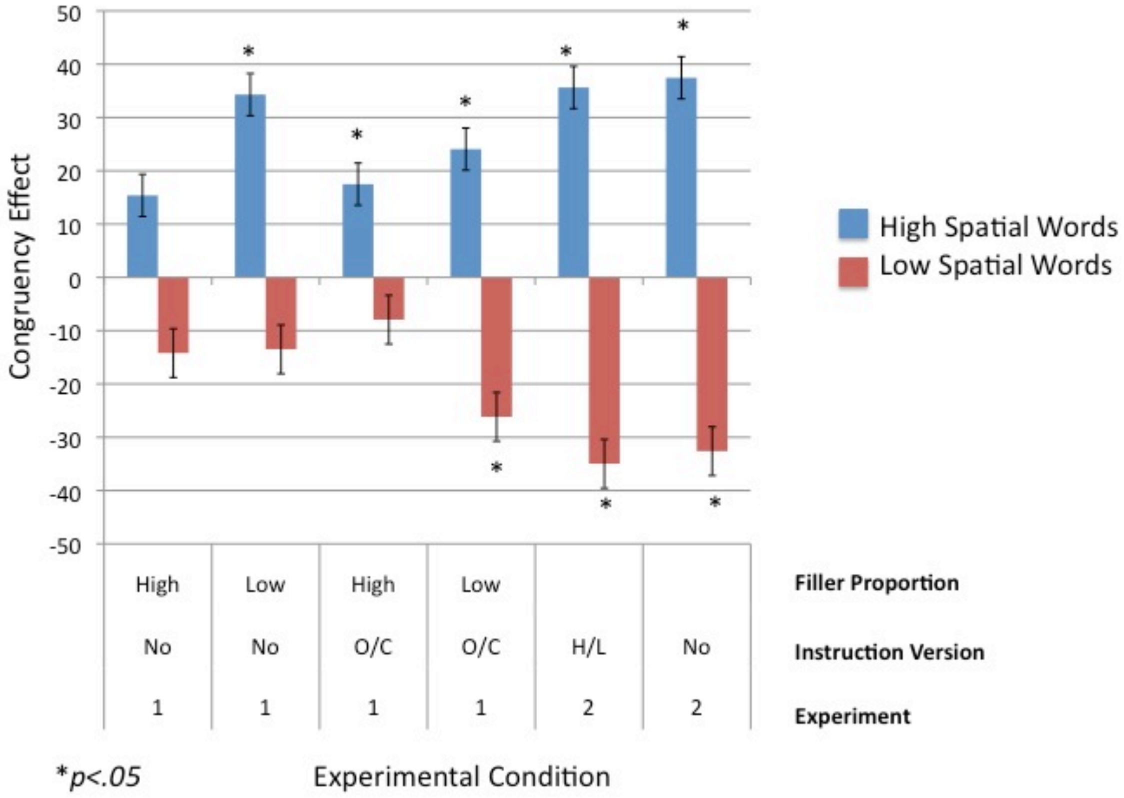

Figure 4. Demonstration of motor effects

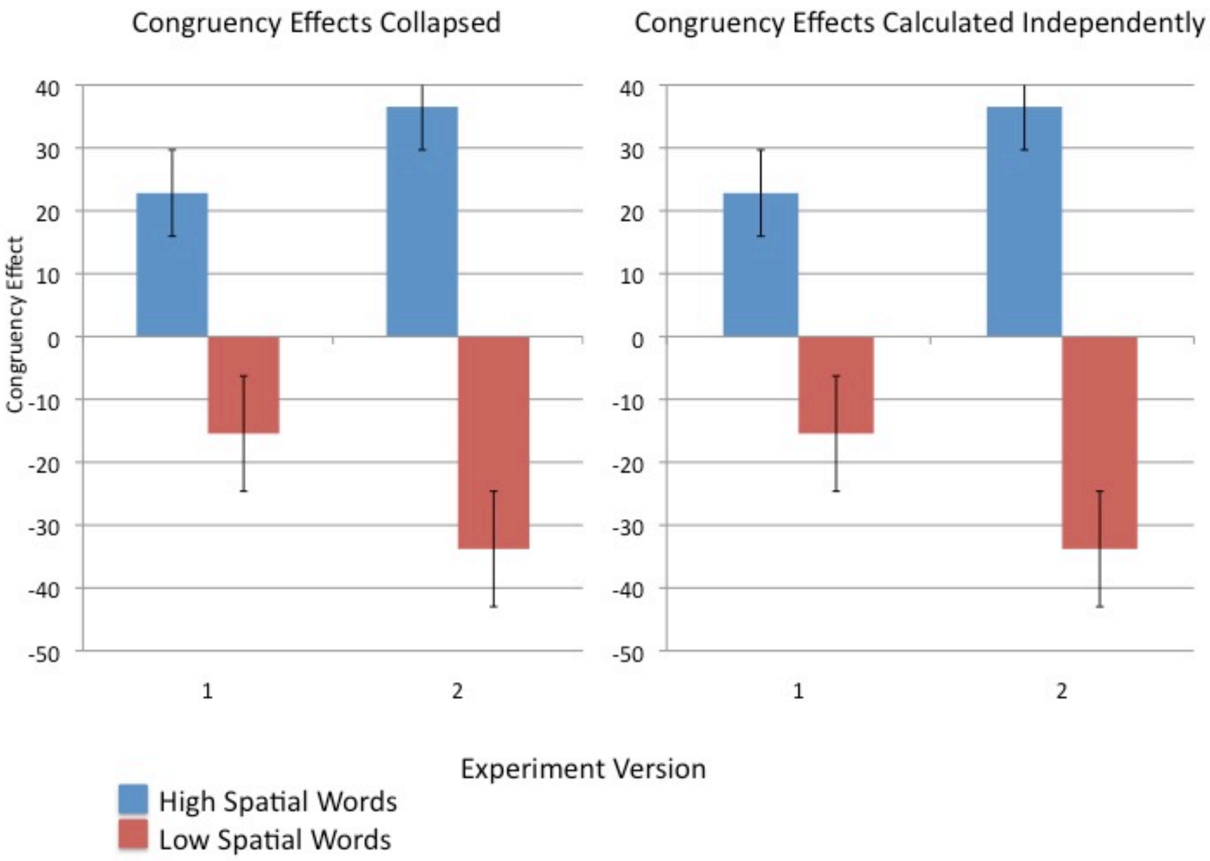

Figure 5. Alternate congruency effect for full response time

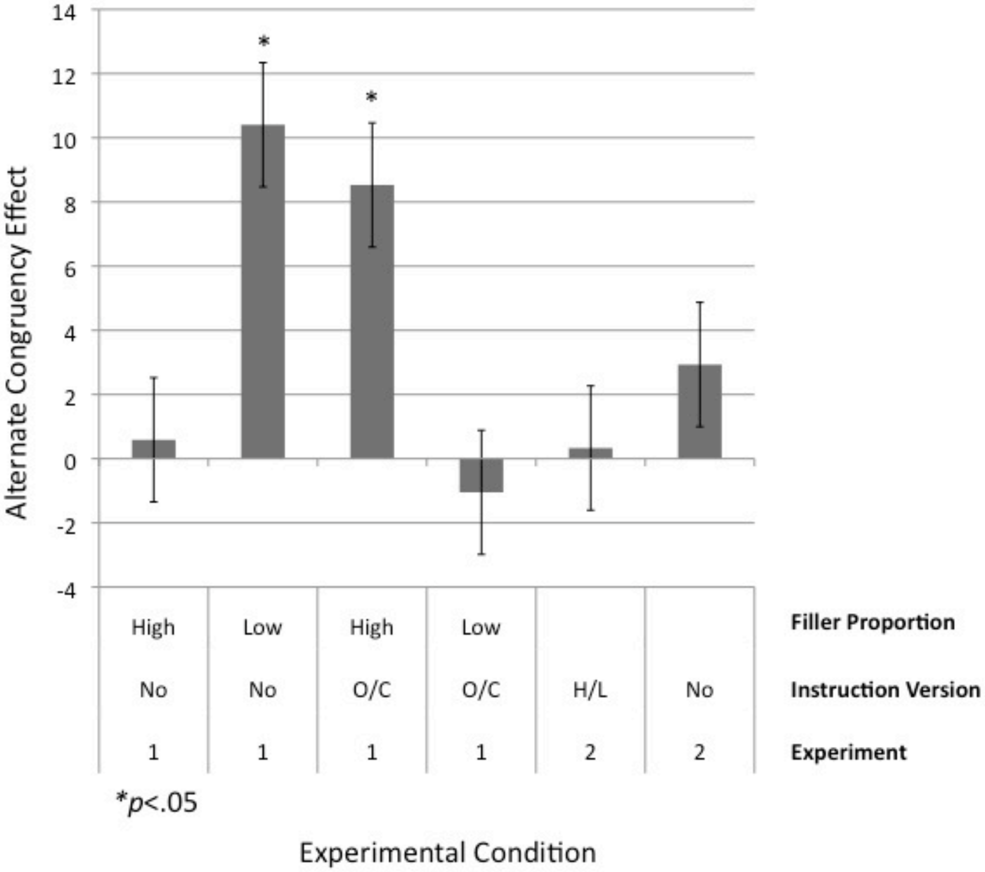

Figure 6. Alternate congruency effect for component response times

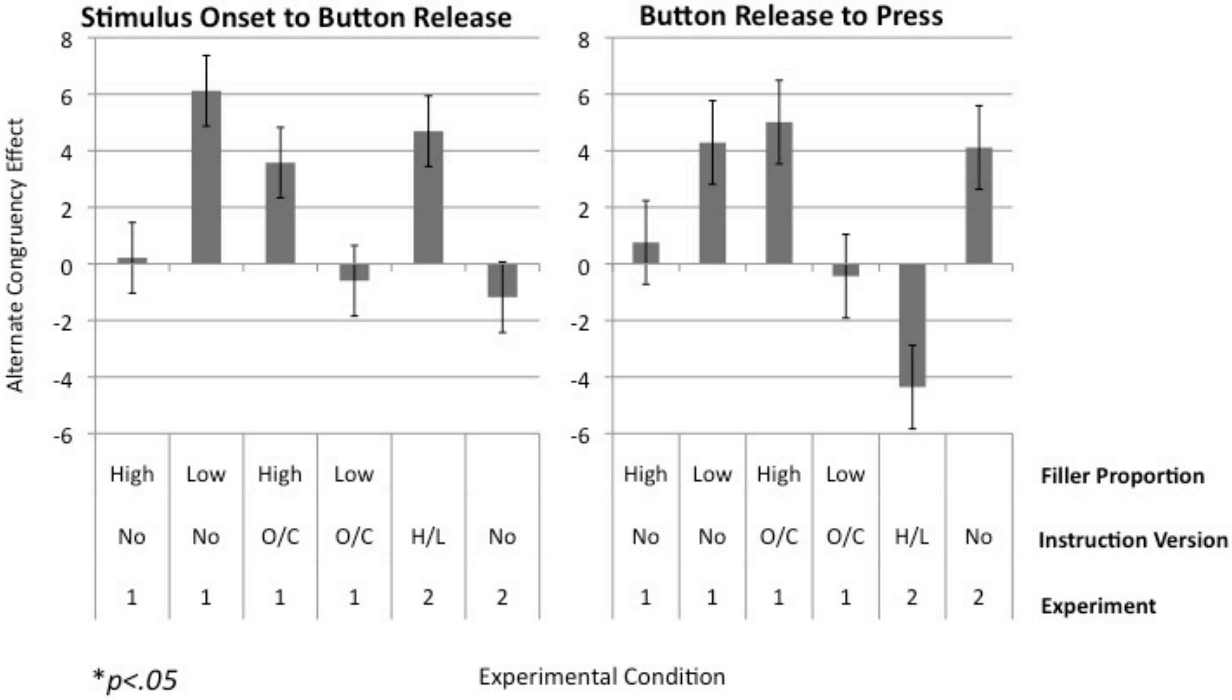

Figure 7. Alternate congruency effect for component response times broken into high and low spatial words

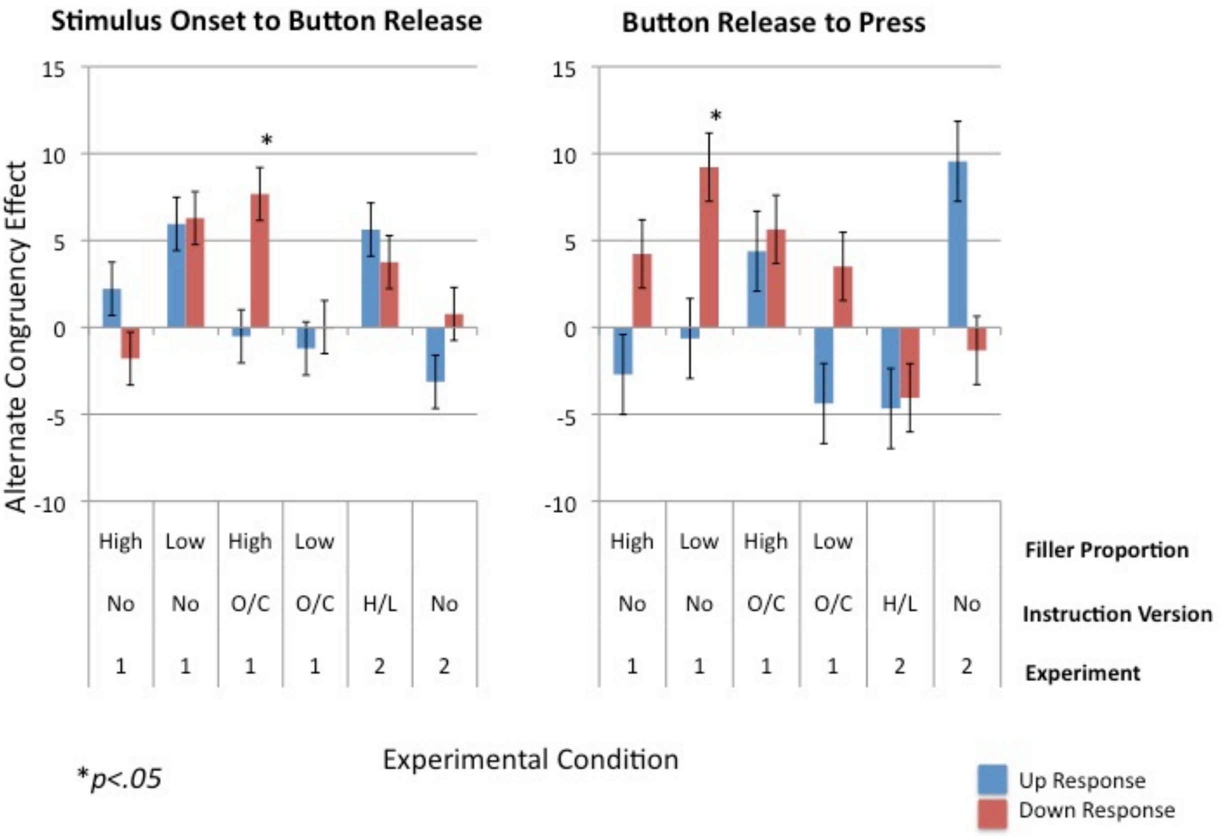

Figure 8. Percent error and true percent error

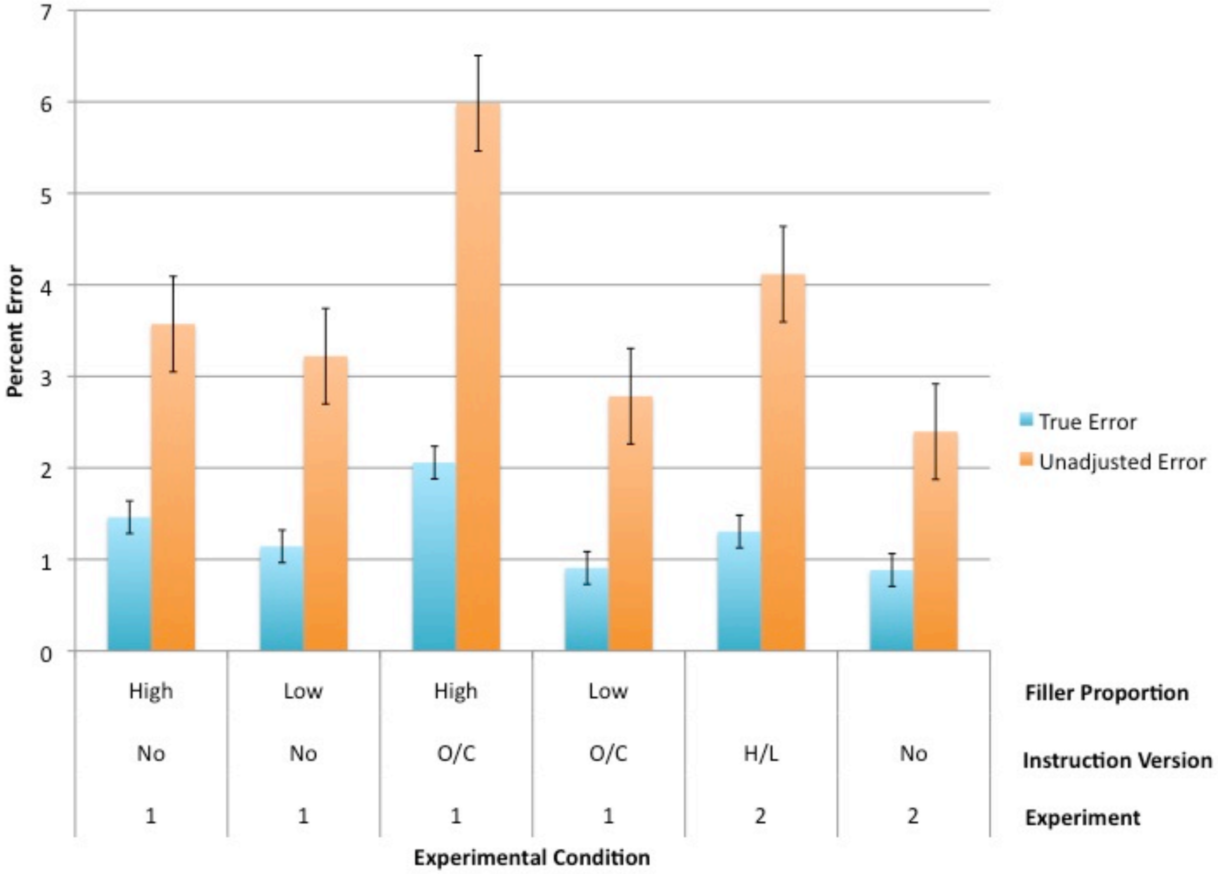

Figure 9. Alternate congruency effect for true percent error

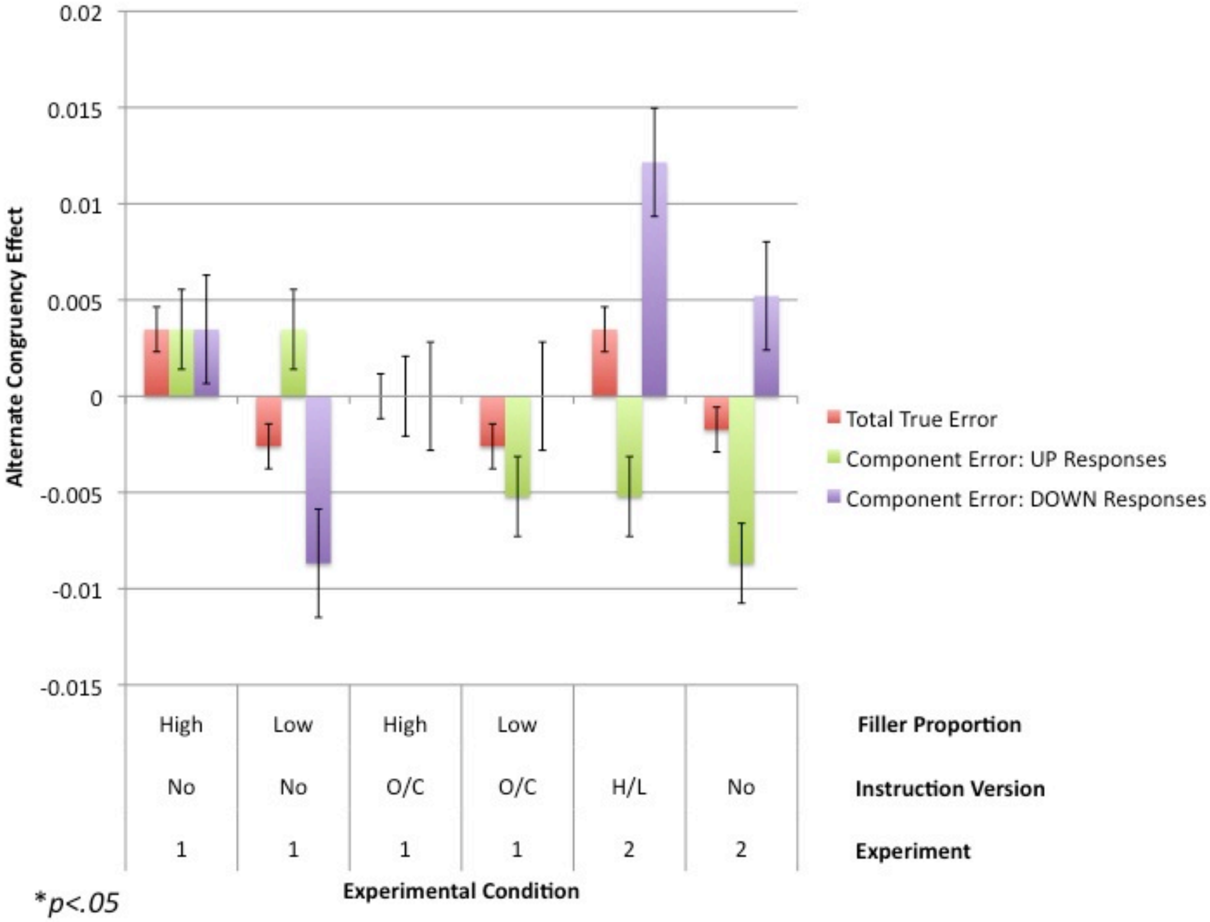

Figure Captions

Figure 1. Hypothesized results for experiment 1 and 2 based on automatic vs. strategic processing. The scale is arbitrary. The relationship between experimental conditions is more important than the exact numbers for congruency effects.

Figure 2. Congruency effects for full response time (stimulus onset to button press) using original formula. Includes all six experimental conditions across both experiments. Experiment 1: high proportion of fillers and no false set (no); low proportion of fillers and no false set, high proportion of fillers and false set (open vs. closed class words); low proportion of fillers and false set. Experiment 2: awareness of critical word set (H/L); no awareness of critical word set.

Figure 3. Original congruency effect for full response time broken down into component congruency effects: one for high spatial words and one for low spatial words. Includes all six experimental conditions across both experiments.

Figure 4. Experiment 1 and 2 motor effects demonstration. The original congruency effect for full response time was broken down into high and low spatial word congruency effects and collapsed across the four experimental conditions (left). The collapsed congruency effect graph was compared to an independent calculation of congruency effect based solely on response direction not congruency of trial (right). The two graphs are identical.

Figure 5. Alternate congruency effects for full response time controlling for motor effects. Includes all six experimental conditions across both conditions. Experiment 1: high proportion of fillers and no false set (no); low proportion of fillers and no false set, high proportion of fillers and false set (open vs. closed class words); low proportion of fillers and false set. Experiment 2: awareness of critical word set (H/L); no awareness of critical word set.

Figure 6. Alternate congruency effect for component response times. Left: stimulus onset to button release. Right: button release to button press. None of the experimental conditions had significant congruency effects.

Figure 7. Alternate congruency effect for component response times broken down into component congruency effects: one for high spatial words and one for low spatial words. Left: stimulus onset to button release. Right: button release to button press.

Figure 8. Percent error and true percent error for all six experimental conditions across both experiments. Experiment 1: high proportion of fillers and no false set (no); low proportion of fillers and no false set, high proportion of fillers and false set (open vs. closed class words); low proportion of fillers and false set. Experiment 2: awareness of critical word set (H/L); no awareness of critical word set. For “true percent error,” responses that were made in the correct direction, but not hitting the exact key were counted as accurate. None of the experimental conditions had significant congruency effects.

Figure 9. Alternate congruency effect for true percent error across all six experimental conditions. None of the experimental conditions had significant congruency effects.

References

Avenanti, A., Bueti, D., Galati, G., & Aglioti, S.M. (2005). Transcranial magnetic stimulation highlights the sensorimotor side of empathy for pain, *Nature Neuroscience, 8(7),* 955-960.

Arbib, M.A., (2005). From monkey-like action recognition to human language: An evolutionary framework for neurolinguistics. *Behavioral and Brain Sciences, 28,* 105-167.

Bargh, J.A. and Ferguson, M.J. (2000). Beyond behaviorism: On the automaticity of higher mental processes. *Psychological Bulletin, 126(6),* 925-945.

Barsalou, L. W. (2007). Grounded cognition. *Annual Review of Psychology, 59,* 617-645.

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

Barsalou, L.W., Niedenthal, P.M., Barbey, A.K., & Ruppert, J.A. (2003). Social embodiment. *Psychology of Learning and Motivation, 43,* 43 – 92.

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

Beilock, S. L., and Holt, L. E. (2007). Embodied preference judgments. *Psychological Science. 18(1)*, 51-57.

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

Bub, D.N., Masson, M.E.J., and Cree, GS (in press). Evocation of functional and volumetric gestural knowledge by objects and words. *Cognition.*

Calder, A. J., Keane, J., Manes, F., Antoun, N. & Young, A. W. (2000) Impaired recognition and experience of disgust following brain injury. *Nature Neuroscien*ce, *3*, 1077–1078.

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

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

Decety, J. (2007). A social cognitive neuroscience model of human empathy. In E. Harmon-Jones (Ed), *Social neuroscience: Integrating biological and psychological explanations of social behavior*. New York, NY: Guilford Press.

Decety, J., Jeannerod, M., Germain, M. & Pastene, L. (1991). Vegetative response during imagined movement is proportional to mental effort. *Behavioural Brain Research, 42*, 1-5.

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

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

Foroni, F. & Semin, G.R. (2009). Language that puts you in touch with your bodily feelings: The multimodal responsiveness of affective expressions. *Psychological Science, 20*, 974-980.

Gallese, V. (2002). The roots of empathy: the shared manifold hypothesis and the neural basis of intersubjectivity, *Psychopathology, 36,* 171-180.

Gallese, V. (2003). The manifold nature of interpersonal relations: the quest for a common mechanism, *The Royal Society*, *538,* 517-528.

Gallese, V., and Goldman, A. (1998). Mirror neurons and the simulation theory of mind-reading. *Trends in Cognitive Sciences*. *2(12)*, 493- 501.

Gentilucci, M. & Gangitano, M. (1998). Influence of automatic word reading on motor control. *European Journal of Neuroscience, 10,* 752 -756.

Glover, S., Rosenbaum, D.A., Graham, J., & Dixon, P. (2004). Grasping the meaning of words. *Experimental Brain Research*, *154*, 103-108.

Glenberg, A. M., & Kaschak, M.P. (2003). The body's contribution to language. In *The Psychology of Learning and Motivation*, Vol. 43, ed. B Ross, pp. 93–126. New York: Academic.

Glenberg, A. M., & Kaschak, M. P. (2002). Grounding language in action. *Psychonomic Bulletin and Review*. *9(3)*, 558-565.

Hatfield, E., Cacioppo, J., & Rapson, R. (1994). *Emotional contagion*. New York: Camridge University Press.

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

Jeannerod, M. (2006). *Motor cognition: What actions tell the self.* New York: Oxford University Press.

Jha, A.P. (2002). Tracking the time-course of attentional involvement in spatial working memory: An event-related potential investigation. *Cognitive Brain Research, 15,* 61-69.

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

Kremen, A. and Block, J. (2002). Absorption: Construction explication by q-sort assessments of personality. *Journal of Research in Personality, 36,* 252-259.

Mehrabian, A. (2000). Manual for the Balanced Emotional Empathy Scale (BEES). (Available from Albert Mehrabian, 1130 Alta Mesa Road, Monterey, CA 93940.

Mehrabian, A., & Epstein, N. (1972). A measure of emotional empathy. *Journal of Personality*. *40(4)*, 525-543.

Mehrabian, A., Young, A.L., and Sato, S. (1988). Emotional empathy and associated individual differences. *Current Psychology: Research and Reviews, 7,* 221-240.

Moors, A. & De Houwer, J. (2006). Automaticity: A theoretical and conceptual analysis. *Psychological Bulletin, 132(2),* 297-326.

Pecher, D., Zeelenberg, R., & Barsalou, L.W. (2003). Verifying different –modality properties for concepts produces switching costs. *Psychological Science*, *14(2),* 118 – 123.

Posner, M. I., & Snyder, C. R. R. (1975). Attention and cognitive control. In R. L. Solso (Ed.), *Information processing and cognition: The Loyola Symposium* (pp. 55-82). Hillsdale, NJ: Erlbaum.

Posner, M.I., Snyder, C.R.R., & Davidson, B.J. (1980). Attention and the detection of signals. *Journal of Experimental Psychology: General, 109(2),* 160-174.

Preston, S.P., & de Waal, F. (2002). Empathy: Its ultimate and proximate bases. *Behavioral and Brain Sciences*. *25*, 1-72.

Proffitt, D.R. (2006). Embodied perception and the economy of action. *Perspectives on Psychological Science, 1,* 110-122.

Proffitt, D.R., Bhalla, M., Gossweiler, R., & Midgett, J. (1995). Perceiving geographical slant. *Psychonomic Bulletin and Review, 2*, 409–428.

Rizzolatati, G., & Arbib, M. A. (1998). Language within our grasp. *Trends in Neuroscience,* *21*, 188-194)

Saxe, R., & Powell, L.J. (2006). It's the thought that counts: Specific brain regions for one component of theory of mind, *Psychological Science, 17(8),* 692-699.

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

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

Simmons, W.K., Hamann, S.B., Harenski, C.N., Hu, X.P., & Barsalou, L.W. (2008). fMRI evidence for word association and situated simulation in conceptual processing. *Journal of Physiology – Paris, 102,* 106-119.

Singer, Seymour, O'Doherty, Kaube, Dolan & Frith (2004). Empathy for pain involves the affective but not sensory components of pain, *Science 303,* 1157–1162.

Smith, E.E., Shoben, E.J., & Rips, L.J. (1974). Structure and process in semantic memory: A featural model for semantic decisions. *Psychological Review, 81 (3),* 214-241.

Solomon, K.O., & Barsalou, L.W. (2004). Perceptual simulation in property verification. *Memory & Cognition, 32,* 244-259.

Tellegen, A. and Atkinson, G. (1974). Openness to absorbing and self-altering experiences ("absorption"), a trait related to hypnotic susceptibility. *Journal of Abnormal Psychology, 83,* 268-277.

Tucker, M., & Ellis, R. (2001). The potentiation of grasp types during visual object categorization. *Visual Cognition, 8,* 769-800.

Ulrich, R. and Maienborn, C. (2008). Left-right coding of past and future in language: The mental timeline during sentence processing. Under revision.

Van den Bergh, O., Vrana, S., & Eelen, P. (1990). Letters from the heart: Affective categorization of letter combinations in typists and nontypists. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 16*, 1153–1161.

Wilson, M. & Knoblich, G. (2005). The case for motor involvement in perceiving conspecifics. *Psychological Bulletin, 131 (3),* 460-473.

Ziv, M. & Frye, D. (2003). The relation between desire and false belief in children’s theory of mind. *Developmental Psychology, 39(5),* 859 – 876.

Zwaan, R.A. & Madden, C.J. (2005). Embodied sentence comprehension. In *Grounding Cognition: The Role of Perception and Action in Memory, Language, and Thought*, ed. D Pecher, R Zwaan, pp. 224–45. New York: Cambridge Univ. Press.

Appendix A

Critical Words

High spatial words. airship, apex, arch, balloon, belfry, branch, ceiling, cloud, collar, crown, eagle, float, pulpit, fly, forehead, fountain, hair, hat, hawk, head, helmet, hill, hood, kite, lark, lift, lightning, missile, monument, moon, mountain, plane, platform, rocket, roof, sky, skyscraper, smoke, spire, star, steeple, sun, tip, tornado, tower, umbrella, volcano, wig

Low spatial words. anchor, ankle, asphalt, base, basement, basin, boot, burrow, carpet, casket, cave, cellar, chasm, dip, dirt, ditch, dive, drain, dungeon, earth, fall, feet, floor, foundation, grass, ground, heel, hole, land, mat, pit, puddle, root, rug, sand, sewer, shoe, sink, slipper, sock, soil, submarine, sunset, toe, tomb, tumble, valley, worm

Filler Words

a, about, according, again, all, although, an, and, another, any, anybody, anyone, as, at, because, both, but, circa, despite, during, each, either, else, ever, everybody, everyone, everything, except, extra, few, for, he, her, hers, herself, him, himself, his, i, if, it, its, itself, likewise, me, mine, my, myself, neither, nevertheless, nobody, nor, notwithstanding, now, of, once, one, oneself, only, opposite, or, other, otherwise, ought, our, ours, ourself, ourselves, own, per, same, she, so, some, somebody, someone, something, still, such, than, that, the, their, theirs, them, themselves, then, these, they, this, those, though, to, together, unless, unlike, unto, us, versus, we, what, whatever, whatsoever, when, whenever, where, whereas, wherever, whether, which, whichever, while, who, whoever, whom, whomever, whose, why, with, without, yet, yonder, you, your, yours, yourself, yourselves