Literally and figuratively speaking: How concepts and perception influence each other using Stroop paradigms

A Thesis submitted to the College of Graduate Studies and Research in partial fulfillment of the requirements for the degree of Master of Arts in Psychology
Cognition and Neuroscience Program
Department of Psychology
University of Saskatchewan
Saskatoon, Saskatchewan

by
Eric Lorentz

Permission to Use

In presenting this thesis in partial fulfillment of the requirements for a Master's Degree from the University of Saskatchewan, I agree that the Libraries of this University may make it freely available for inspection. I further agree that permission for copying of this thesis in any manner, in whole or in part, for scholarly purposes may be granted by the professor or professors who supervised my thesis work or, in their absence, by the Head of the Department or the Dean of the College in which my thesis work was done. It is understood that any copying, publication, or use of this thesis or parts thereof for financial gain shall not be allowed without my written permission. It is also understood that due recognition shall be given to me and to the University of Saskatchewan in any scholarly use which may be made of any material in my thesis.

Requests for permission to copy or to make other uses of materials in this thesis in whole or part should be addressed to:

Head of the Department of Psychology
University of Saskatchewan
Saskatoon, Saskatchewan, S7N 5A5
Canada

OR

Dean
College of Graduate Studies and Research
University of Saskatchewan
107 Administration Place
Saskatoon, Saskatchewan, S7N 5A2
Canada

Abstract

Metaphorical phrases like ‘seeing red’ in anger or ‘red-hot’ for temperature take abstract ideas and ground them more tangible, physical phenomena, suggesting a relationship between semantic meaning and visual attributes. We tested the dominant direction of influence between abstract semantic processing and visual perception by pairing words and visual attributes in the same stimulus. Semantic categorization was used to test if visual attributes moderate semantic decisions by including both congruent and incongruent pairings with visual attributes. For instance, categorizing ‘scald’ in red font colour as hot (congruent) can be compared to ‘scald’ in blue font colour (incongruent), or ‘freeze’ in red font colour (incongruent) to determine if visual attributes (e.g., colour) automatically affect semantic decisions. Using the same stimuli, visual attribute categorization (e.g., categorizing ‘scald’ as red) was used to test if word meaning automatically affects perceptual decisions. Experiment 1 included mad (red congruent) and sad (blue congruent) emotion words, whereby semantic categorization revealed consistent congruency effects (i.e., shorter RTs and fewer errors with congruent trials than incongruent trials), but not with colour categorization. Experiment 2 extended these effects to the domain of temperature, including hot (red congruent) and cold (blue congruent) temperature words. Semantic categorization revealed consistent congruency effects on RT and errors, but not with colour naming. Experiment 3 extended Experiment 2 by including the neutral colour green. In semantic categorization, congruent pairs showed facilitation relative to neutral, and incongruent pairs showed interference relative to neutral, whereas only facilitation occurred with colour categorization for red-hot pairs. These results support the obligatory processing of visual attributes in semantic tasks, grounding abstract semantic meaning in colour processing. In the reverse direction, colour categorization tasks also showed semantic influences, although smaller and less consistently. Experiment 4 tested the generality of these effects in the visual domain of time processing. Congruent and incongruent pairs were generated by combining short durations and long durations with temporally associated words (e.g., ‘brief,’ ‘eternal’). Congruency effects occurred consistently on RTs and error rates for the duration categorization (revealing semantic influences), but only consistently on errors with semantic categorization. Thus, word meaning serves as the dominant attribute in the domain of time, indicating varying strengths of automaticity between visual attributes. These experiments explore the generality and boundary conditions of how visual attributes, like colour and time perception, and word meaning share representations, whereby asymmetries provide new evidence regarding the automatic direction of processing influences in these domains.

Acknowledgements

Though only my name is reserved for the cover of this thesis, I could not have accomplished all this without the dedication and support of those around me.

It has been my pleasure to have Ron Borowsky as a supervisor, for his unwavering support and guidance while giving me the freedom to pursue my own interests. His patience and wisdom has helped me overcome the obstacles in finishing this thesis and program, and all with a healthy sense of fun and humor.

I am grateful to have stayed in a fantastic lab with dear friends and wonderful colleagues. Marla Mickleborough has showed me to be passionate about the work I do, and how rewarding that work can be. Jim Cheesman lit the research flame, and has inspired me in my creativity and diligence. And of course, my dear friends in the lab that have held with me through the ups and downs, the challenges and successes. This would not have been an experience, but a job, without Chelsea, Layla, Tessa, Kathryn, Sarah, and Conley. There is no reward like sharing the learning and teaching of frontier research, and the many adventures we have spent together.

For always trying my experiments, no matter how demanding, I want to make special mention of my girlfriend Megan. Her suggestions, insights, and celebrations in the victories, as well as support in the setbacks have helped me believe in myself and persevere.

I want to thank my family for their constant support, concern, and love. If you are wondering ‘how much longer until you’re done?’- now you know!

Finally, I appreciate the financial support I have received to pursue this research.

Thank you all.

Table of Contents

List of Tables

List of Figures

Chapter 1: Introduction and Review of Literature

How do we describe our subjective experiences to others when they are unique and personal? Philosophers describe *qualia* as 'what it is like' to experience such mental states as seeing red, feeling heat, smelling cinnamon, or hearing a burst of thunder. Dennett (1988) argues that such qualia are ineffable, that is, they cannot be effectively communicated by means other than direct experience. And yet, we meaningfully describe our experiences with each other quite regularly. Moreover, the language we use is often more abstract than basic qualia, for instance in describing a *warm* exchange between *close* friends. Here, warm and close are not literal descriptors, but they capture an abstract, figurative quality of friendship that two people share. The way we use language to communicate these experiences reveals a relationship between experience and semantic comprehension. Although the common use of these metaphors suggests a unique relationship between semantic meaning and perceptual representations, it remains unclear how such physical attributes may relate to semantic knowledge. To examine the cognitive architecture of this relationship, and its possible mechanism(s), four studies are presented describing the interactions that occur in making perceptual or semantic decisions while simultaneously processing related perceptual and semantic cues.

1.1 Semantic Organization and Language Comprehension

Semantic memory is broadly defined as a part of long-term memory involving the meaning of a word or phrase, and concepts that describe worldly knowledge (Tulving, 1972; Martin, 2001). Semantic knowledge is thought to be represented across a network (Thompson-Schill, Kan & Oliver, 2006), where accessing or engaging semantic memory involves spreading activation (Collins & Loftus, 1975). For instance, a search through the semantic network may begin with activating a conceptual node (e.g., 'hot'), which then spreads activation to related, or ‘nearby’ concepts (e.g., ‘fire,’ ‘red’) until a goal is satisfied like finding a related item, or verifying some relationship between facts. Before discussing theories of semantic organization further, it should be noted the terms *semantic* and *conceptual* are used interchangeably. This is because authors tend to describe ‘conceptual representations’ in regards to theories of semantic organization (e.g., Barsalou, 1999), while using ‘semantic representations’ in experimental studies of semantic organization (e.g., Martin, 2001), while both describe the meaning of words and their referents.

The organization of semantic memory is mainly described in two ways. In one prominent model, concepts are organized by logical, hierarchical relationships (Quillian, 1968), whereby propositions that are true of all members of a concept (e.g., living things) are stored at the top at a superordinate category level. Propositions true of some members (e.g., animals have muscle, but not plants) are stored at lower levels, until facts that are true of only one member are stored within the individual concept (e.g., stripes of a zebra). While this model predicts faster verification of facts about individual concepts (e.g., zebras have stripes versus animals have muscles), the evidence does not consistently favor this model (e.g., McCloskey & Glucksberg, 1979). Alternatively, the semantic network may be organized based on the related features of concepts (Rips, Shoben & Smith, 1973), where the number of shared features between concepts (or typicality of a member) reflects category membership. This *graded* category membership finds stronger support, as more typical members of categories (e.g., a robin is a bird) are verified faster than less typical members (e.g., a chicken is a bird). Their model posits two stages, the first assesses very high or low overall familiarity of characteristic and defining features together for fast responses (e.g., a salmon is a fish: ‘yes,’ a salmon is a bird: ‘no’). The second stage processes cases requiring specifically defining features (e.g., a whale is a fish: ‘no’), taking more time to respond. Support for this model indicates responding is based on similarity of features between members, rather than a hierarchical model.

The notions of typicality and similarity between member features within categories still raise questions about the format of semantic representation, and it has been debated whether these representations are amodal (e.g., propositional or symbolic; e.g., Caramazza & Shelton, 1998; Plaut, 1999), or involve modal representations with motor, visual, or tactile processes (e.g., Allport, 1985; Pulvermuller, 1999). Standard theories of cognition describe the semantic network as amodal, where network activation and concepts are represented symbolically in an abstract, non-modular system, that is, with little or no relation to the experience of those concepts in perception, action, or introspection (e.g., Pylyshyn, 1981, 1984; Zwaan, 1999; see Barsalou & Hale, 1993, for a review). For instance, Tulving (1972) argues that input to semantic memory begins perceptually in experience, although the representation of perceptual properties themselves are not stored in semantic memory. Rather, new abstract representations are formed with no relation to the original experience (Pylyshyn, 1984). As such, no simulation in perceptual or sensorimotor processes would occur when retrieving a concept (Mahon & Caramazza, 2008).

On the other hand, some theorists argue that semantic knowledge is shared with perceptual processes (e.g., Barsalou, 1999; Kan, Barsalou, Solomon, Minor, & Thompson-Schill, 2003). As early as Locke (1690), philosophers have hypothesized a modal relationship between knowledge and experience: “Whence comes [the mind] by that vast store, which the busy and boundless fancy of man has painted on it with an almost endless variety? … To this I answer, in one word, From experience.” Such musings have inspired psychologists to map the interface between semantics and experience. Under this view, different attribute domains in semantic knowledge are supported by separate brain regions, from colour and motion, to action and abstract representations (Allport, 1985, Gibbs, 2005). In describing the relationship between experience and semantic memory, I will use the word *perception* in the broad sense to capture perceptual representation in memory. In this way, perception encompasses early perceptual input (e.g., differential activation of cones in the retina due to different wavelengths of light), its simulation (e.g., imagine the colour red), and perceptual judgments (e.g., responding ‘red’ as your judgment of the colour you perceived).

Grounded cognition rejects the traditional amodal view, positing the central role of experiential simulation in cognition (Barsalou, 2008). According to grounded cognition theory, simulations reenact perceptual, motor, and introspective states that occurred during the original experience within the environment, supporting action, perception, and semantic representations. For instance, Allport (1985) describes the distributed, sensorimotor representation of a telephone, involving action (e.g., grasping), tactile (e.g., hard, smooth), spatiomotor (e.g., near the ear, lightweight), visual (e.g., dark, reflective), and auditory (e.g., voice, dial tone) elements. Prominent evidence of grounded cognition can be seen across different modalities, broadly categorized into visual or functional domains. For instance, in action-perception, Masson, Bub, and Breuer (2011) showed that seeing an object (e.g., mug, pan) automatically triggered participants to simulate grasping and functional actions, but only when the handle was positioned for its functional action. In vision, Hansen, Olkkonen, Walter, and Gegenfurtner (2006) found evidence that participants simulate an object’s natural colour (e.g., a yellow banana), when viewing achromatic objects (e.g., a gray banana), as their perceived colour shifted towards the opponent colour (i.e., a bluish banana). Indeed, after tiring from a run, a hill may look steeper, and after wearing a heavy pack, a path may look longer, arguably because the simulated effort to traverse these obstacles is greater (Proffitt, 2006). Thus, grounded cognition and distributed semantic-sensorimotor representations provide a model of embodied cognition, linking perceptual and semantic domains in simulation.

Specific to language, a model of semantic memory must also include more abstract concepts like *emotion* or *time*. Metaphors, in which a word or phrase is applied to an object or action that is not literally applicable, exemplify how abstract semantic meaning and domains of perception are often related. Lakoff and Johnson (1999) argue that metaphors recruit mental imagery from sensorimotor domains to support semantic representations. An example of a primary metaphor they use is ‘understanding as grasping.’ Here, ‘grasping’ an idea, or a concept going ‘over one’s head’ uses physical object manipulation (or lack there of) as a proxy for the abstract idea of comprehension. Primary metaphors have a minimal structure of cross-domain attributes, relating abstract concepts (e.g., ‘understanding’) and more concrete, often sensorimotor attributes (.e.g., ‘grasping’; Grady, 1997). Other primary metaphors include ‘intimacy as closeness,’ or ‘time as motion’ (Lakoff & Johnson, 1999). According to metaphor representation theory, the example that nice people are ‘sweet’ is not because they are literally sweet (and would not be if eaten!), but because interacting with nice people and eating sweet foods are similarly pleasant (Meier, Moeller, Reimer-Peltz, & Robinson, 2012).

In the embodied theory of language comprehension, Zwaan (2004) further maintains that concepts employ simulation in supporting semantic representations. For instance, Zwaan and Yaxley (2003) showed participants were faster to identify that words were semantically related (e.g., attic, basement), when spatially oriented as expected (i.e., attic on top, basement on bottom) compared to the reverse. Similarly, read words like ‘kick’ are shown to activate overlapping regions of the brain that are also active during the word’s referenced perception or action (Kan et al. 2003; Pulvermüller, 1999). In addition, when comprehending language, hand and eye movements are shown to be consistent with the described situation (Glenberg & Gallese, 2012). Overall, such findings bring strong evidence of simulation in semantic comprehension, and demonstrate how metaphors and simulation are not simply by-products of language, but are central to supporting semantic representation and cognition.

How might these metaphorical relationships and the embodiment of semantics develop? Johnson's (1996) theory of conflation proposes that, in the course of learning, young children do not distinguish between the two meanings in commonplace metaphors. The idea that 'affection is warmth', or 'intimacy is closeness', regularly co-occur in language. Young children identify this covariation and, at first, understand these concepts as the same, only separating them in a later period of differentiation. Through this mechanism of conflation, these associations become automatic, and persist even after they are distinguished. Similarly, Piaget and Inhelder (1969) take the developmental perspective that cognition first operates on sensorimotor representations before more abstract concepts. Because abstract thought develops later, it recruits and is built on sensorimotor representations, fundamentally linking perceptual-conceptual pairings into adulthood. A compelling demonstration of primary metaphor operating at the perceptual level involves the 'good is bright, bad is dark' metaphor (Meier, Robinson, & Clore, 2004). Participants were faster to categorize words as positively valenced when presented in white, compared to black, and faster to categorize words as negatively valenced when presented in black, compared to white. However, identifying the presented colour as black or white with the same stimuli did not show this congruency effect, revealing the automatic perceptual-conceptual relationship with this metaphor.

1.2 Perception and Semantic Processing Conflict

A related literature describes the reverse direction, whereby semantic processing automatically modulates performance on a perceptual task. A classic example is the Stroop task (Stroop, 1935), which was originally used to compare colour naming solid squares to colour words written in mismatched coloured ink (e.g., ‘red’ in blue ink, or ‘blue’ in green ink). Participants were significantly delayed in colour naming words written incongruently compared to colour naming squares, indicating interference from reading words.

Stroop effects occur when a natural relationship occurs between the carrier word’s referent and its physical attribute, such that congruent and incongruent pairs occur (Algom, Chajut, & Lev, 2004). As MacLeod (1992) describes, Stroop effects are a critical measure of attentional capacity. Cattell (1886) first reported that word-reading was faster than naming objects and their properties (like colour) aloud. When Stroop (1935) demonstrated that words produce interference when they co-occurred with colour, it was further evidence that word-reading was more automatic, demanding less attention resources, and that colour-naming was more controlled, or voluntary.

Since this landmark study, evidence supports the locus of this interaction at both semantic and response stages of processing (Zhang & Kornblum, 1998). For instance, Klein (1964) found that participants named the colour of rare words faster than common words, and common words faster than colour-related words like ‘grass’ in red ink and ‘lemon’ in blue ink, supporting the semantic relatedness of colour words with colour representations. Similarly, Lorentz, McKibben, Ekstrand, Gould, Anton, and Borowsky (2016) found a semantically based Stroop effect with colour-associated words after controlling for word-reading processes (e.g., pseudohomophones that sound like words versus actual words) and contingency effects (differences in repeating congruent conditions more often than incongruent conditions). By mapping multiple colour responses to the same button, Zhang and Kornblum (1998) also showed that participants were slower to identify the colour of incongruent words (e.g., ‘red’ in green ink), than congruent words (e.g., ‘red’ in red ink), even when both red and green responses mapped to the same button. This ‘stimulus-stimulus’ conflict between word and colour supports a semantic source of conflict during a perceptual task.

Interestingly, word-reading such incongruent stimuli (e.g., ‘red’ in green font) compared to neutral (e.g., ‘red’ in black font) does not produce the same conflict compared to colour naming (Stroop, 1935; see MacLeod, 1991, for a review). These results further support the dominance of semantic processing, leading theorists to posit that semantic access in word-reading with skilled readers is automatic (e.g., Dagenbach, Carr, & Wilhelmsen, 1989; LaBerge & Samuels, 1974; Neely; 1977; see Carr, 1992 for a review). Still, some argue word-reading is not purely automatic, because reducing the scope of attention (e.g., only colouring one letter; Besner, Stolz, & Boutilier; 1997, or presenting distractor words; Kahneman & Chajczyk, 1983) can mitigate some of the Stroop effect. Even so, the robustness of Stroop effects among its many variants supports rapid and minimal effort in accessing semantic representations when reading, providing evidence of semantic dominance in these perceptual tasks.

Although many studies suggest asymmetries between perceptual and semantic processing, Richter and Zwaan (2009) provided evidence that colour perception and semantic processing may share underlying representations in a bidirectional relationship using two tasks. First, they demonstrated that lexical decisions on colour words were faster when preceded by a congruent colour than an incongruent colour, indicating a link from perceptual processing to semantic representation. Second, they found that a colour discrimination task was also speeded when a congruent colour word (e.g., ‘red’) was presented between the two similarly shaded colours. These results indicated that colour representations are activated when processing colour words, supporting the simulation account of language comprehension, in addition to colours priming semantic activation.

Overall, evidence for influences between perceptual and semantic representations comes from two perspectives. During semantic tasks, grounded cognition supports the dominance of perceptual processes supporting semantic representation through simulation and embodiment (Barsalou, 2008; Zwaan, 2004) as well as basic metaphors (Lakoff & Johnson, 1999). During perceptual categorization tasks, in the reverse direction, evidence also supports shared representations through the automatic processing of semantic cues (Lorentz et al. 2016; Klein, 1964; Zhang & Kornblum, 1998). While most studies find dominance of one representation (i.e., perceptual or conceptual), with bidirectional effects shown occasionally (e.g., Richter & Zwaan, 2009), much of the relative dominance (and degree of directional influence) remains to be explored in terms of when one modality will more strongly influence another.

Grounded cognition suggests that the more abstract a concept is, the more it may recruit perceptual or bodily processes, and thus the easier a semantic task may be completed when related attributes are present (e.g., faster; more accurately). In the reverse direction, the automaticity of word reading suggests rapid access to semantic knowledge, which may subsequently activate perceptual or bodily simulation. Although these effects are relatively well explored with literal Stroop items like colour and colour-associated words (e.g., ‘red,’ or ‘lemon,’; Anton et al. 2014; Klein, 1964; Lorentz et al. 2016; Richter & Zwaan, 2009) as well as actions and objects (e.g., ‘mug,’ or ‘attic’; Masson et al. 2011; Zwaan & Yaxley, 2003), these effects are less understood with more abstract concepts. In particular, the evidence is still lacking as to when one representation (perceptual or semantic) may be more dominant with metaphor-based concepts related to perceptual processes.

In order to test the source dominance of perceptual and conceptual representations in more abstract associations, the following four studies combine related semantic and perceptual attributes while manipulating the task-relevant attribute. The logic of these four experiments involves an exploration of three semantic domains. In Experiment 1, the concept of emotion, involving the ‘red is mad’ and ‘blue is sad’ metaphor will be tested using mad and sad associated words presented in red or blue. Experiment 2 explores the abstract association of temperature and colour, involving the ‘hot is red’ and ‘cold is blue’ metaphor, extending the effects found in the domain of emotion. Next, Experiment 3 employs the colour green as a neutral baseline to test the effects of facilitation and interference with temperature and colour associations, separately. Finally, using the ‘time is space’ and ‘time is motion’ metaphor, Experiment 4 tested the relationship of time-associated concepts with time-perception to examine the generality of the perceptual-conceptual link. The results of these four studies are discussed in relation to the above literature, demonstrating both generality and boundary conditions to the dominant processing of perceptual and conceptual information sources.

Chapter 2: Emotion Stroop Experiment

This chapter is based on the following manuscript in revision with *Cognition and Emotion*. The chapter has been edited to ensure consistency with the thesis.

Lorentz, E., Gould., L., Ekstrand., E., Mickleborough. M., & Borowsky. R. (2016). Feeling blue’ and ‘seeing red’: Evidence grounding discrete emotions in colour perception from a Stroop paradigm. Under review at *Cognition and Emotion.*

When describing the subjective experience of emotion, we often borrow language from the perceptual domain, including ‘feeling blue’ for sadness and ‘seeing red’ for anger, in an attempt to connect these experiences to the real world (Lambie & Marcel, 2002). Such expressions suggest that these emotion words in the semantic network are grounded in more concrete perceptual experiences like colour processing (Lakoff & Johnson, 1999; Williams, Huang, & Bargh, 2009). For example, simulating the concept of anger may draw on observed or produced aggressive actions (Berkowitz, 1993), as well as perceived flush expression that occur when someone is angry (Changizi et al. 2006). Along similar lines in this perspective, the spreading activation of a concept like ‘sadness’ in the semantic network may not be limited to other semantic concepts like *cry* or *alone*, but also perceptual representations like looking down (Meier & Robinson, 2006), or conjuring relatedly dark colours like blue or black (Nelson, McEvoy, & Schreiber, 2004).

To test the experiential account of emotional grounding in colour processing, Fetterman, Robinson, and Meier (2012) examined the ease with which participants could categorize emotion words depending on the colour the word was presented. Specifically, they examined the hypothesis that anger is grounded in the colour red. In their first experiment, they had participants categorize words as either fear, anger, or neutral meaning when presented in either gray or red colours, and found that red colour facilitated anger categorizations compared to gray, but did not facilitate fear categorizations compared to gray. In a second experiment, they extended these findings by testing sad and anger emotion words in the colours red and blue. They found significantly faster anger categorizations in the colour red than the colour blue, while categorizing sad emotions in blue did produce this effect. When examining colour categorizations (i.e., *rage* is ‘blue’), they found no evidence that emotion meaning modulated colour categorization speed. They concluded that the asymmetry between tasks provides unambiguous support for the grounding of anger emotion in red colour perception, and thus support for their theory that metaphors reflect the grounding of emotion in perceptual processes.

Contrary to the findings of Fetterman et al. (2012), Sutton and Altarriba (2008) found that colour categorization of emotion words did reveal congruency effects in terms of faster reaction times (RTs) with congruent combinations compared to incongruent combinations. Namely, the words 'angry' or 'rage' presented in red, 'sad' or 'depression' presented in blue, 'scared' or 'coward' presented in yellow, and 'greed' or 'envy' presented in green showed faster RTs than incongruent combinations. Thus, they argued that the representation of emotion words in memory can be studied in the same manner as other colour associates like '*fire*' in red font (Klein, 1964). Under this framework, colour-related emotion words in the semantic network activate automatic spreading activation from the word's meaning to the colour.

The results of Sutton and Altarriba (2008) extend the processing of emotion words to a larger body of literature with colour associates (see also Anton et al. 2014; Klein, 1964), whereby spreading activation from a word's meaning shows a congruency effect with colour-naming. Fetterman et al. (2012) argued that this direction of effect of words acting on colour categorization was absent in their own results, and that colour processing is the dominant process that acts on emotion-word meaning, whereas Sutton & Altarriba found evidence for emotion word meaning generating congruency effects during colour categorization. Unfortunately, Fetterman et al. (2012) used 37.5% less trials with colour categorization than word categorization, and Sutton & Altarriba (2008) did not include a word categorization task. Thus, claims about asymmetries in congruency effects remains on uncertain grounds.

Differences between these two studies make drawing strong conclusions about the relationship between colour and emotion words even more difficult. First, while Fetterman et al. (2012) examined individual colour and emotion combinations, Sutton and Altarriba (2008) analyzed congruency effects as an average across colours and emotion types, which makes it difficult to determine if some conditions were driving their effects (e.g., anger-red associations). Second, the difference in the number of colours (four with Sutton & Altarriba, 2008; two with Fetterman et al. 2012) and the number of words per emotion category (two with Sutton & Altarriba, 2008; ten in Expt 1, and six in Expt 2 with Fetterman et al. 2012) suggest potential sources for these contradictory results. For instance, colour-emotion associations found in the colour categorization task may be peculiar to the limited word set employed by Sutton and Altarriba (2008) and not generalize to a larger set of words.

Overall, there is evidence that supports both emotion words acting on colour-categorization, and colour acting on emotion categorization of words, although the mechanism of these effects remains unclear. Given the incompatibility of these two accounts, we sought to address three key findings that have shown inconsistencies. First, we will examine whether reliable Stroop effects will occur with colour associated emotion words when employing a larger set of words for each category. Second, we will examine whether congruency effects in the word categorization task only occur for anger-meaning words, as Fetterman et al, (2012) found. Third, we will examine if congruency effects occur for both colour and word categorization tasks, or if an asymmetry exists, suggesting a dominance of colour or word processing with emotion words.

2.1 Hypotheses

We tested the association of emotion-related words and colours, evaluating the hypotheses that: (1) emotion-related words are grounded in colour simulation, reflected by a colour congruency effect on semantic categorization, and (2) colour identification is influenced by the semantic processing of emotion words, reflected by a semantic congruency effect on colour categorization. Given the results of Fetterman et al. (2012), we hypothesized dominant congruency effects with semantic categorization. To the extent that a direct association occurs between emotion and colour (e.g., Sutton & Altarriba, 2008), we hypothesized smaller congruency effects on colour categorization.

2.2 Method

2.2.1 Participants

Thirty-three undergraduates were recruited from the University of Saskatchewan psychology participant pool and given course credit for their participation. Participants had a mean age of 19.5 years (*SD* = 2.00, range of 18-26 years), normal or corrected-to-normal vision, and English as their first language. Participants were tested individually after giving informed consent. The experiment was approved by the University of Saskatchewan Research Ethics Board.

2.2.2 Stimuli

The stimuli included 48 words, half of which had a meaning related to sad emotion (sad meaning words), and the other half of which were related to the emotion of anger (‘mad’ meaning words; see Appendix A). Words were selected through a systematic search of synonyms for ‘sad’ and ‘anger,’ as well as their synonyms, and verified for clear meaning with fellow research assistants. Sad and mad word lists were matched for length, $t(46) = .965$, $p = .339$ and word frequency, $t(46) = .982$, $p = .331$ (Balota et al. 2007). Each word was presented once in the colour red (RGB = 255, 0, 0; HSB = 0°, 100, 100) and once in the colour blue (RGB = 0, 0, 255; HSB = 240°, 100, 100) in each task block.

2.2.3 Apparatus

The experiment was programmed and run using E-Prime (www.pstnet.com) on a Lenovo 6075-DPU computer with a 60Hz Compaq 7500 colour monitor. An Audio-Technica ATR1200 microphone was used to activate the voice key in a PST serial response box.

2.2.4 Design and Procedure

On each trial, a white fixation-cross appeared in the center of a black screen until the participant pressed the spacebar. The fixation cross remained for 500 ms and was replaced by the stimulus on a black background until the participant vocally categorized either colour ('red' or 'blue) or semantic category ('mad' or sad'), triggering the microphone. The experimenter then coded the response as mad/sad meaning with semantic categorization, red/blue colour with colour categorization, or indicated a spoiled trial (e.g., smacking lips; vocalization failed to trigger the microphone). Ten randomly chosen practice trials drawn from the 96 test stimuli appeared prior to each task block, which led directly into test trials. The 96 test trials were chosen randomly, and task order (colour categorization; semantic categorization) was counterbalanced between participants. Half the stimuli were congruent (red-mad or blue-sad) and the remaining half was incongruent. After completing experimental tasks, participants rated each word (presented in white) on the strength of its emotion association using a five point scale (1 being moderate and 5 being extreme, given that each word was chosen to have at least a moderate association to emotion).

2.3 Results

After removing spoiled trials, median RTs for correct trials and errors were analyzed by-subjects (see Figures 2.1 and 2.2) and by-items. Medians were chosen as they are more protected from outliers. For each analysis, a 2 x 2 x 2 analysis of variance (ANOVA) included Colour (red, blue), Emotion (mad, sad), and Task (colour categorization, semantic categorization). Each factor was repeated by-subjects, whereas by-items, Emotion was between items (see Table 2.1 for ANOVA results). Overall, the main effect of Task was consistent by-subjects and by-items, indicating faster RTs on colour categorization than semantic categorization. The interaction between Colour and Emotion was also consistent, supporting congruency effects on RTs. The effects shown on RTs were not mitigated by errors, as there were no results with errors in the opposite direction to support a speed accuracy trade-off. Finally, the three-way interaction on RTs, including Task, indicates that congruency effects were significantly different between tasks. That is, Figures 2.1 and 2.2 show the predicted congruency effects for semantic categorization, but not for colour categorization.

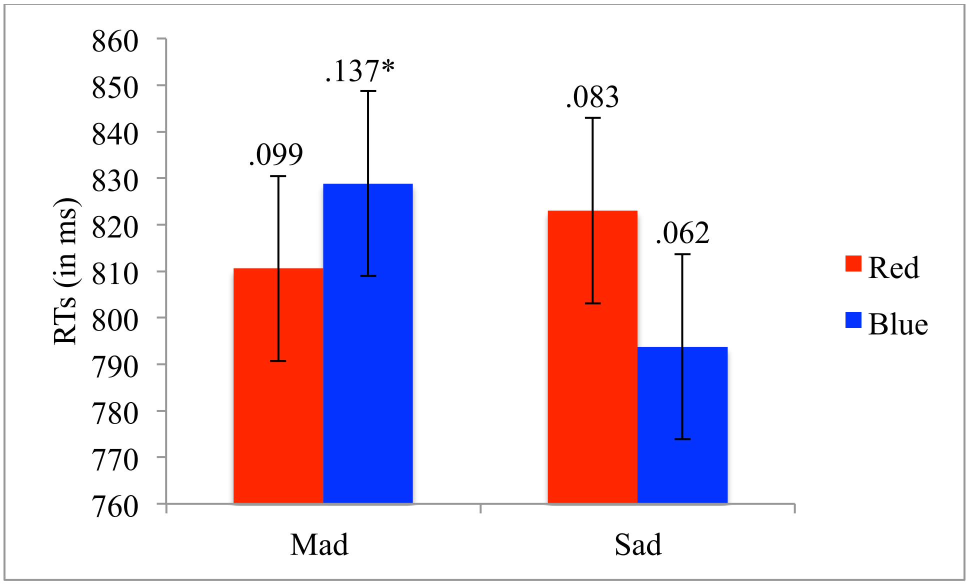

Figure 2.1. Semantic emotion categorization RTs as a function of Emotion and Colour. Error rates are displayed above as a proportion out of 1 (with a 95% confidence interval of .022; asterisks for all figures represent significant differences by errors to the adjacent bar). For all figures, intervals 95% confidence intervals (CIs) represent Loftus and Masson’s (1994) 95% repeated measures CIs using the within-subjects error term (95% CI = 19.2 ms on RTs).

Table 2.1

ANOVA Results for Experiment 1 by-subjects (s) and by-items (i) for Median RTs and Errors

Experiment 1	F_s	F_i	$(df_1, df_2)_s$	$(df_1, df_2)_i$	MSE_s	MSE_i	p_s	p_i
Latency								
Colour	1.84	4.20	(1, 32)	(1, 46)	1473.73	1211.81	.185	*.046
Emotion	2.06	1.16	(1, 32)	(1, 46)	2767.09	5190.77	.161	.287
Task	324.51	781.19	(1, 32)	(1, 46)	324.51	6121.44	*< .001	*< .001
Colour x Emotion	7.66	16.01	(1, 32)	(1, 46)	1376.97	1211.81	*.009	*< .001
Colour x Task	8.02	4.14	(1, 32)	(1, 46)	1162.71	1269.13	*.008	*.048
Emotion x Task	.09	.19	(1, 32)	(1, 46)	3030.82	6121.44	.767	.668
Colour x Emotion x Task	6.85	8.57	(1, 32)	(1, 46)	1186.54	1269.13	*.013	*.005
Errors								
Colour	.52	.41	(1, 32)	(1, 46)	.003	.002	.478	.526
Emotion	16.60	4.33	(1, 32)	(1, 46)	.003	.009	*< .001	*.043
Task	16.45	8.53	(1, 32)	(1, 46)	.006	.008	*< .001	*.005
Colour x Emotion	13.60	16.58	(1, 32)	(1, 46)	.003	.002	*.001	*< .001
Colour x Task	.25	.14	(1, 32)	(1, 46)	.003	.002	.621	.715
Emotion x Task	3.89	1.53	(1, 32)	(1, 46)	.005	.008	.057	.222
Colour x Emotion x Task	.28	.32	(1, 32)	(1, 46)	.004	.002	.598	.573

Note: MSE = mean square error; df = degrees of freedom; * = significant at $p < .05$.

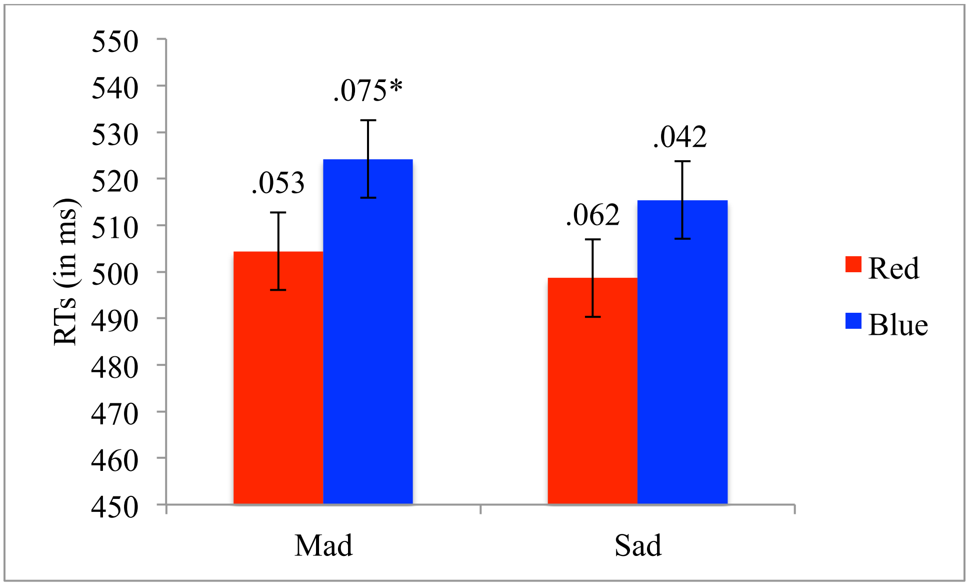

Figure 2.2. Colour categorization RTs as a function of Colour and Emotion. Error rates are displayed above as a proportion out of 1 (95% CI = .019). RTs have a 95% CI of 8.3 ms.

We also conducted Bayesian analyses on the average congruency effect with RTs in each task, which provide a better estimate of the probability for a hypothesis than traditional *p*-value testing, given the obtained data (e.g., Masson, 2011). With colour categorization, the average congruency effect was calculated for each colour by subtracting the congruent RTs from the incongruent RTs. Then, participants' red and blue average congruency effects were summed and divided by two, producing an average colour categorization congruency effect of 1.55 ms. The colour categorization congruency effect of 1.55 ms was compared to 0 in a one-sample *t*-test, *t*(32) = .438, *p* = .664, where the Bayes factor in support of the null hypothesis was 6.746, and thus a posterior probability of .871, indicating positive evidence for the null hypothesis (i.e., no congruency effect).

In the semantic categorization task, the average congruency effect was calculated separately for mad and sad meaning words by subtracting congruent RTs from incongruent RTs. Mad and sad congruency effects were then summed and divided by two for each participant, yielding an average congruency effect for semantic categorization of 23.74 ms. The average congruency effect of 23.74 ms was then compared to 0 in a one-sample *t*-test, *t*(32) = 2.939, *p* = .006, where the Bayes factor in support of the alternative hypothesis was 5.625, and thus a posterior probability of .849, indicating positive evidence for the alternative hypothesis (i.e., congruency effect).

Next, we compared mad and sad word lists on ratings of their emotional association (see Appendix A). Using a scale of 1 – 5 (representing 'moderate' to 'extreme'), mad words had a mean rating of 3.99, and sad words had a mean rating of 3.30. The mad word list had a larger mean rating than the sad word list (after controlling for inequality of variances, Levene's test *F*(46) = 6.20, *p* = .016) when submitted an independent samples *t*-test, *t*(40.41) = 5.675, *p* < .001.

To test if effects in the item-analysis of correct median RTs were due to these rating differences, we repeated the ANOVA with Colour, Emotion and Task, with Rating as a continuous variable. In this analysis, the effect of Rating and its interaction with other variables is tested on RTs to examine if differences between list ratings may have produced the effects. If

Rating is a significant source of these effects, then adjustments on RTs to account for differences in list ratings will prevent this ANOVA from replicating the previous ANOVA without Rating. The main effect of Task was still significant, $F(1,45) = 18.36$, $MSE = 6080.49$, $p < .001$. In addition, the interaction between Colour and Emotion was still significant, $F(1,45) = 8.90$, $MSE = 1238.56$, $p = .005$. Finally, the three-way interaction was still significant, $F(1,45) = 5.49$, $MSE = 1296.29$, $p = .024$. The remaining main effects of Colour, Emotion, and Rating, and their two way interactions failed to reach significance, Fs(1,45) < 2.17, ps > .147. These results largely replicated the by-subjects and by-items analyses without Rating, with the exception of the significant interaction between Colour and Task in the by-subjects analysis, and the significant main effect of Colour and significant interaction between Colour and Task found in the by-items analysis. Thus, differences in congruency effects should not be attributed to differences in rating across word lists.

Importantly, when Task Order was included as a between subjects variable with RTs, the main effect of Task Order and its interactions did not reach significance, Fs(1, 31) < .817, ps > .372. By errors, the only the Task x Emotion x Task Order interaction approached significance, $F(1, 31) = 3.801$, $MSE = .004$, $p = .060$. As the remaining main effect of Task Order and its interactions did not reach significance, Fs(1, 31) < 1.732, ps > .197, the interpretation of congruency effects should not be attributed to carry-over effects between tasks.

2.4 Discussion

We investigated the direction of association between emotion and colour perception using both colour and semantic categorization tasks. Specifically, we investigated if the association between the colour blue and ‘sad’ emotion and the colour red and ‘anger’ emotion would be stronger during colour categorization or semantic categorization, revealing a dominant influence of word or colour processes in the domain of emotion.

In the colour categorization task, RTs did not indicate congruency effects typically found with colour associates. In the present experiment, red categorizations were consistently faster than blue categorizations. Although this pattern produced the expected congruency effect on RT with mad words, it also produced the reverse pattern of congruency effects on RT for sad words, producing evidence for no congruency effects according to the Bayesian analysis. Thus, there are inconsistent congruency effects in colour categorization. The simplest interpretation is that red colour categorization was faster than blue colour categorization. In the semantic categorization task, the results indicated the expected association between red with anger emotion and blue with sad emotion. Both the RTs and error rates in the by-subject and by-item analyzes indicated faster RTs and fewer errors with congruent colour and emotion combinations, a pattern that persisted when controlling for emotion ratings (see Figure 2.3).

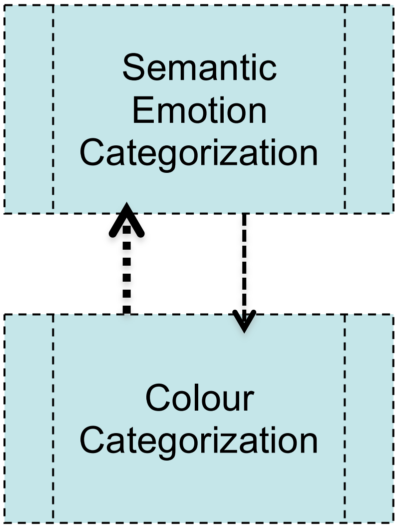

Figure 2.3. Relationship between colour and semantic emotion categorization. Evidence for congruency effects occurs from colour to semantic categorization for mad and sad words, while evidence for automatic semantic processing is inconsistent, occurring only for red-mad colour categorizations.

Our results align more closely with the results of Fetterman et al. (2012) than of Sutton and Altarriba (2008), albeit with important distinctions. In semantic categorization, we demonstrate new experimental support for an association between the colour blue and sad emotion, previously demonstrated only with self-report methods (Nelson et al. 2004). In addition, our results provide support for an association between the colour red and anger, which may have a strong physiological basis (e.g., becoming flush when becoming mad; Changizi et al. 2006). According to Boyatzis and Varghese (1994), these colour-emotion schemas form early in childhood, and become progressively more complex into adulthood (Hemphill, 1996). For instance, Hemphill found in prior self-report studies that blue is also related to feelings of calmness and positive affect, indicating a multifaceted relationship between the colour blue and emotion that merits further study. In addition, one limitation of the current study, and similar studies is the categorical view of emotion, although much research supports emotion as a dimensional construct (Niemic & Warren, 2002). Thus, a more nuanced approach to the relationship between emotion and colour should be the next step for future research in this area.

Although our results showed a congruency effect in colour categorization for mad emotion words, they did not support Sutton and Altarriba's (2008) finding that colours and emotions have distinct congruencies in the colour categorization task. Our task employed two emotion types (mad and sad), whereas their study employed a larger set of four emotion types (anger, sadness, envy, and cowardice), which they averaged together when examining congruency effects. In particular, Sutton and Altarriba used words that were shown to generate colour words in a free association task (Nelson et al. 2004), supporting the conclusion that their effect was generated by the proximity with which specific emotion and colour words exist in the semantic network, supporting a direct association. Our own word lists were generated on the principle of deeply probing two emotion concepts (i.e., sad and mad emotion) within the semantic network, testing how general the association between colour and mad and sad emotion words can be generalized to words. Certain words in our lists support the conclusion of Sutton and Altarriba (2008), where ‘angry’, an item they included, resulted in a Stroop effect of 130.5 ms in our colour categorization task (see Appendix B). However, our Bayesian analyses show that the colour categorization congruency effect would have a posterior probability of only .129 (i.e., 1 - .871). In this way, we suggest specific words in an emotion category (e.g., ‘angry’) may be directly related to colour in the semantic network, despite having a similar emotion association rating as words less related to colour (e.g., *vile,* Stroop effect 12 ms). Thus, an important distinction may be drawn that some mad or sad emotion words are colour-related, but word membership in a mad or sad category does not categorically show a direct association with colour.

It remains, however, if this pattern of congruency effects is peculiar to the domain of emotion and colour processing. In a similar metaphorical relationship, the temperature meanings of hot and cold are also related to the colours red and blue, respectively (e.g., Ho,Van Doorn, Kawabe, Watanabe, and Spence, 2014). To examine if the dominant effects of colour processing on semantic categorization generalizes beyond the semantic domain of emotion, we tested this relationship with temperature in a similar manner to Experiment 1.

Chapter 3: Temperature Stroop Experiments

Experiments 2 and 3 are in press with *Visual Cognition.* The manuscript has been adapted for this chapter to ensure consistency with the broader thesis.

Lorentz, E., Ekstrand, C., Gould, L., & Borowsky, R. (in press). Red-hot! How colour and semantic temperature processing interact in a Stroop-like paradigm. *Visual Cognition*.
http://dx.doi.org/10.1080/13506285.2016.1183742

Temperature and colour associations have largely been studied in rating the thermal properties of physical colours, or describing the colours evoked by different temperatures, leaving the mechanism of this association largely unexplored. For instance, projected and surface red colours are rated warmer than blue colours (Lewinski, 1938; Newhall; 1941). This 'cold-blue' and 'red-hot' association develops gradually, solidifying around age 18 (Morgan, Goodson, & Jones, 1975). These associations also occur with real objects. Osgood, Suci, & Tannenbaum (1957) had participants rate five objects presented in six colours on a scale of warmth, and found red objects were consistently rated warmer than objects in all other colours.

In the above experiments, the purpose may have been obvious to participants, where expectancy of an association may have influenced the results. To investigate an unbiased association, Ho et al. (2014) employed two experiments examining a correspondence between colour, physical temperature and temperature words. In the first experiment, participants categorized colour patches (red, blue) and thermal words ('warm', 'cold') via a two alternative, forced choice button-press task. Participants were faster when congruent red-warm and blue-cold were paired (i.e., same button for hot/red) than with incongruent pairings (i.e., same button for red/cold), indicating the colour and temperature correspondence. Their second experiment involved priming physical temperatures and temperature words with different colour patches, as well as priming colour patches with physical temperatures or temperature words. Participants were instructed to categorize either warm or cold physical temperature, the words 'warm' or 'cold', or the colour patches red or blue in separate blocks using verbal responses (i.e., say 'red'/'blue', or 'hot'/'cold'). The critical finding in these priming tasks showed that colour primes produced faster responses with physical temperature categorization in congruent pairings than incongruent pairings (i.e., a congruency effect), although physical temperatures failed to produce congruency effects with colour categorization. Additionally, words failed to produce congruency effects with colour categorization, and colours failed to produce congruency effects with word categorization.

Two major methodological differences between Ho et al. (2014) and prior Stroop and semantic categorization tasks suggest how a relationship between colour and temperature words has yet to be found. First, while prior studies combined colour and words in the same stimulus (e.g., 'fire' in red font, 'angry' in blue font), Ho and colleagues separated primes from targets by 2000 ms, a time determined by the apparatus used to generate thermal primes. As stimulus onset asynchronies are typically much shorter (e.g., 200 ms; Neely, 1991), when separated at all, a 2000 ms SOA may allow prime processing to resolve before the target is presented. Secondly, Ho et al. employed a limited, two word set ('warm' and 'cold'). Presented as targets, these words may have simply been read aloud instead of semantically categorized, greatly speeding responding. In contrast, similar studies like Anton, Gould, and Borowsky, (2014), Fetterman et al. (2012) and Sutton and Altarriba (2008) employed between 6 and 10 words, which serve to test a larger portion of the semantic network, minimize repetition, and preclude reading the target (e.g., 'anger', 'fire') as the response (i.e., 'red,' 'blue') .

To test the semantic temperature and colour association similarly to Experiment 1, and prior Stroop and semantic categorization tasks, we employed a larger set of 24 hot and 24 cold meaning words. Each word was presented in both red and blue to create an equal number of congruent and incongruent pairs. Each pair was presented once in two blocks, one block for colour categorization (red, blue), and one for semantic categorization (hot, cold). A set of 24 words in each list was used to generalize to items in the semantic network, in addition to generalizing to subjects.

3.1 Hypotheses

Similar to Experiment 1, we tested the association of temperature-related words and colours, evaluating the hypotheses that: (1) temperature-related words are grounded in colour simulation, reflected by a colour congruency effect on semantic categorization, and (2) colour identification is influenced by the semantic processing of temperature words, reflected by a semantic congruency effect on colour categorization. In line with prior colour and semantic categorization studies, we removed target words from the response set, combined colours and words in the same target, and employed a larger set of words for each category to create optimal semantic categorization. To the extent that the temperature meaning is grounded in colour through red-hot and cold-blue associations, in line with the results of Experiment 1, we predicted a dominant influence of colour processing on semantic categorization, shown by colour congruency effects in semantic categorization. Given a natural association between hot-red and cold-blue, suggested by prior self-report studies (e.g., Morgan et al. 1975), we predicted an influence of temperature words on colour processing, producing Stroop effects in the colour categorization task.

3.2 Method

3.2.1 Participants

Twenty-four undergraduates (M_{age} = 20.2 years; *SD* = 1.55, range 18- 24) were recruited from the University of Saskatchewan psychology participant pool and given course credit for their participation. All participants had normal or corrected-to-normal vision, and spoke English as their first language. The experiment was approved by the University of Saskatchewan Research Ethics Board.

3.2.2 Stimuli

Twenty-four cold meaning words and 24 hot meaning words (see Appendix B) were selected through a systematic search of synonyms for ‘hot’ and ‘cold,’ as well as their synonyms, and verified for clear meaning with fellow research assistants. Lists were matched for length, $t(46) = .763$, $p = .449$, and word frequency, $t(46) = .052$, $p = .958$, (Balota et al. 2007). Each word was presented in size 18 Arial font once in the colour red and once in the colour blue in each task block.

3.2.3 Apparatus

The apparatus was identical to Experiment 1.

3.2.4 Design and Procedure

The design and procedure were identical to Experiment 1.

3.3 Results

After removing spoiled trials, median RTs for correct trials and errors were analyzed by-subjects (see Figures 3.1 and 3.2) and by-items. For each analysis, a 2 x 2 x 2 analysis of variance (ANOVA) included Colour (red, blue), Temperature (hot, cold), and Task (colour categorization, semantic categorization). Each factor was repeated by-subjects, whereas by-items, Temperature was between items (see Table 3.1 for ANOVA results). Overall, the main effect of Task was consistent across subjects and items, indicating larger RTs for semantic categorization than colour categorization. The main effect of Colour was not consistent between subjects and items. Consistent interactions between Colour and Temperature by subjects and items also support congruency effects, with faster RTs and less errors with congruent red-hot and cold-blue pairs than the remaining incongruent pairs. However, including Task in the three-way interaction indicates that congruency effects were significantly different between tasks. Although congruency effects are shown with semantic temperature categorization (see Figure 3.1), these effects were not consistent across hot and cold meanings with colour categorization (see Figure 3.2).

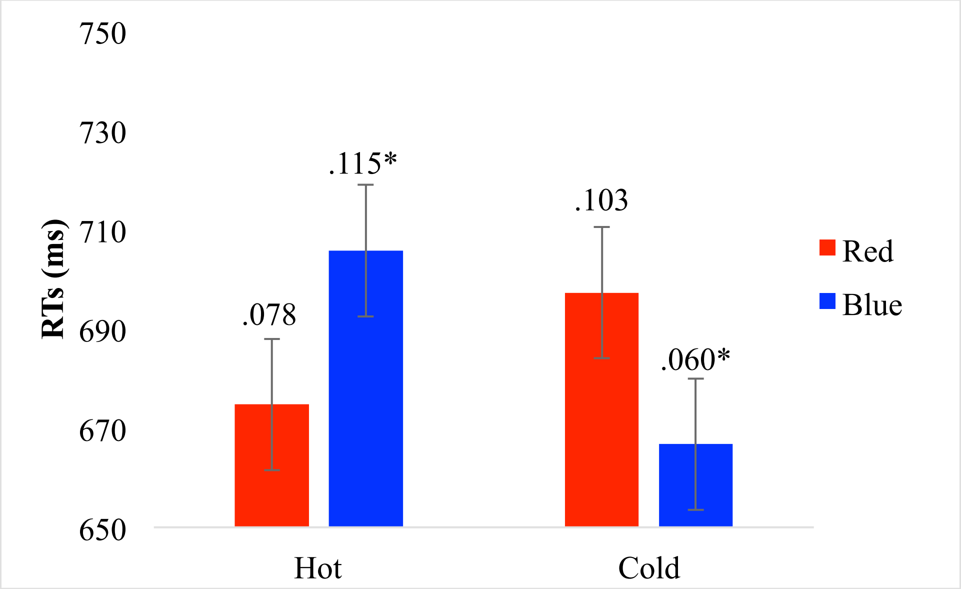

Figure 3.1. Semantic temperature categorization RTs as a function of Colour and Temperature. Errors are displayed above as a proportion out of 1 (95% CI = .022). RTs have a 95% CI of 13.2 ms.

Table 3.1

ANOVA Results for Experiment 2 by-subjects (s) and by-items (i) for Median RTs and Errors

Experiment 2	F_s	F_i	$(df_1, df_2)_s$	$(df_1, df_2)_i$	MSE_s	MSE_i	p_s	p_i
Latency								
Colour	5.71	10.62	(1, 23)	(1, 46)	449.51	994.24	*.025	*.002
Temperature	.21	< .01	(1, 23)	(1, 46)	1084.00	1190.83	.653	.971
Task	186.04	1634.02	(1, 23)	(1, 46)	10661.62	1266.13	*< .001	*< .001
Colour x Temperature	20.45	30.77	(1, 23)	(1, 46)	807.85	994.24	*< .001	*< .001
Colour x Task	2.03	.14	(1, 23)	(1, 46)	1186.39	1002.72	.168	.711
Temperature x Task	1.83	3.00	(1, 23)	(1, 46)	966.50	1266.13	.190	.090
Colour x Temperature x Task	10.78	10.33	(1, 23)	(1, 46)	662.54	1002.72	*.003	*.002
Errors								
Colour	.39	.21	(1, 23)	(1, 46)	.002	.003	.538	.652
Temperature	.53	.41	(1, 23)	(1, 46)	.003	.004	.475	.524
Task	12.16	12.48	(1, 23)	(1, 46)	.003	.004	*.002	*.001
Colour x Temperature	6.07	5.27	(1, 23)	(1, 46)	.003	.003	*.022	*.026
Colour x Task	1.62	.54	(1, 23)	(1, 46)	.001	.004	.216	.465
Temperature x Task	2.17	1.33	(1, 23)	(1, 46)	.002	.004	.154	.255
Colour x Temperature x Task	8.97	5.57	(1, 23)	(1, 46)	.003	.004	*.006	*.023

Note: MSE = mean square error; df = degrees of freedom; * = significant at $p < .05$.

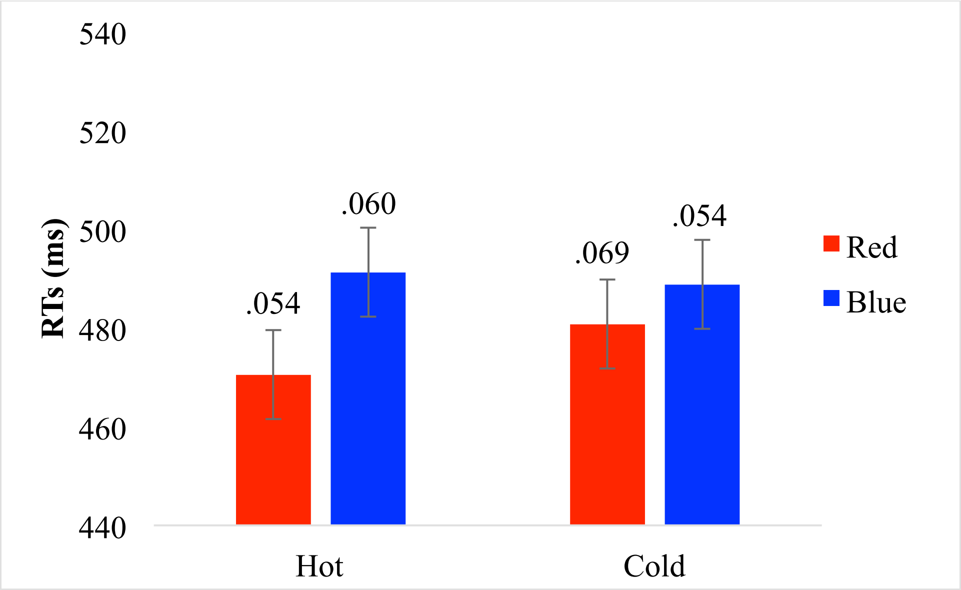

Figure 3.2. Colour categorization RTs as a function of Colour and Temperature. Errors are displayed above as a proportion out of 1 (95% CI = .0182). RT have a 95% CI of 9.0ms.

We also conducted Bayesian analyses on the average congruency effect on RTs in each task, where prior odds for both the null and alternative hypotheses were set to one. In the colour categorization task, the average congruency effect of 6.35 ms was compared to 0 in a one-sample *t*-test, *t*(23) = 1.361, *p* = .187. The Bayes factor in support of the null hypothesis was 2.688, and thus a posterior probability of .729, indicating weak evidence for the null hypothesis (i.e., no congruency effect), or a posterior probability of .271 for the alternative hypothesis. In the semantic categorization task, the average congruency effect of 30.75 ms was compared to 0 in a one-sample *t*-test, *t*(23) = 4.894, *p* < .001, where the Bayes factor in support of the alternative hypothesis was 434.33, and thus a posterior probability of .998, indicating very strong evidence for the alternative hypothesis (i.e., congruency effects).

When comparing hot and cold word lists on ratings of temperature association (see Appendix), hot words (*M* = 4.21) had significantly higher ratings than cold words (*M* = 3.44), *t*(46) = 4.77, *p* < .001. We follow-up this difference in our replication with Experiment 3.

Importantly, when Task Order was included as a between subjects variable with RTs, the main effect of Task was marginally significant, *F*(1, 22) = 3.25, *MSE* = 31078.82, *p* = .085, and the Task by Task Order interaction was significant, *F*(1, 22) = 6.93, *MSE* = 8476.69, *p* = .015. These results indicated faster RTs on the second task, especially for colour identification appearing second. As the remaining interactions were not significant for either RTs or errors, *F*s(1, 26) < 1.92, *p*s > .17, the interpretation of congruency effects should not be attributed to carry-over effects between tasks.

3.4 Discussion

We tested the dominant mechanism of association between colour and temperature by employing both semantic and colour categorization tasks. Semantic categorization revealed significant and consistent congruency effects, with speeded responding and less errors for congruent pairings (i.e., red-hot and blue-cold) compared to incongruent pairings (i.e., red-cold and blue-hot). In addition, a congruency effect appeared for hot meaning words with colour categorization, but not cold meaning words on RTs, with no significant effects of errors. These results provide evidence of a dominant influence of colour processing over semantic categorization compared to the reverse direction (see Figure 3.3).

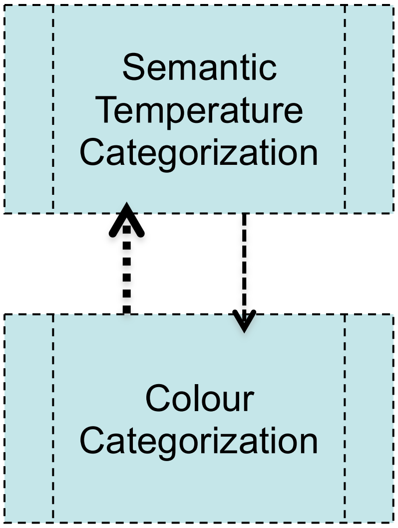

Figure 3.3. Relationship between colour and semantic temperature categorization. Evidence for congruency effects occurs from colour to semantic categorization for hot and cold words, while evidence for automatic semantic processing is inconsistent, occurring only for red-hot colour categorizations.

Consistent with the metaphor that hot is red and cold is blue, these results more broadly support the theory of grounded cognition (e.g., Barsalou, 2008), as decisions about abstract temperature concepts were systematically influenced by related, but task-irrelevant colour cues. Although the use of metaphor may be used as a helpful analogy, our results point to the obligatory association of colour with temperature. Similar to Meier et al. 2004, the stimulus colour appears to automatically suggest its temperature association. These results are also consistent with the theory of metaphor representation (Lakoff & Johnson, 1999), whereby concept knowledge builds on physical feature representations.

In the reverse direction, some evidence supports the influence of semantic processing on colour categorization, although a dissociation between red-hot and cold-blue conditions suggests this association is not based on metaphor. Rather, as Algom et al. (2004) indicates, a natural relationship may exist between hot meaning words and the colour red. One plausible explanation is that hot words show a direct association with naturally red objects, where hot associated words like ‘lava,’ ‘magma,’ or ‘fever’ may also be colour-associates, similar to words like ‘lemon’ with yellow or ‘grass’ with green (Klein, 1964). Another origin of these direct associations may stem from physiological effects of temperature on the human body. Generating excess internal body heat leads to dilation of the blood vessels and a flush, red facial expression (Shearn, Bergman, Hill, Abel, & Hinds, 1990), which can also be associated with increased activity during emotions of anger and sexual arousal (Changizi, Zhang, & Shimojo, 2006).

In supporting the dominance of colour processing over semantic processing in the domain of temperature, our results generalize the grounding of semantic concepts in perception, as suggested through primary metaphor (Lakoff & Johnson, 1999). The dominance of perceptual encoding with abstract concepts like emotion and temperature bring strong support for the theory of grounded cognition, as our effects suggest an obligatory association between perception and concept knowledge.

After completing the prior experiment, feedback from reviewers from *Visual Cognition* requested that we conduct an additional experiment, first to show a replication, and second to evaluate whether the congruency effect for semantic categorization was driven by facilitation, interference, or both. In addition, we took the opportunity to determine if colour associations were driving Stroop effect in colour categorization, and not differences in word list ratings between hot and cold word lists. In order to address these questions, we conducted a second experiment using a neutral baseline, which would assess the contribution of facilitation or interference in each task. Secondly, the use of a neutral colour green can address the mechanism of congruency effects in the colour categorization task, as green is neutral with respect to temperature associated words, but incongruent with respect to red or blue colour associated words.

3.5 Hypotheses

If hot words are also red colour associate words, then presenting words like 'lava' in red will be congruent, and 'lava' in any other colour, like green or blue, would be incongruent, suggesting only facilitation will be present comparing red-hot to red-green combinations, and no difference between hot-green and hot-blue combinations. If temperature association is driving congruency effects with hot words in colour categorization, then presenting words like 'lava' in green should be neutral, and 'lava' in blue should be incongruent, generating both facilitation and interference. Thus, Experiment 2 was replicated with the addition of a neutral colour (i.e., green) to allow for the assessment of facilitation and inhibition relative to this neutral baseline for each task.

3.6 Method

Experiment 3 was identical to Experiment 2 with the inclusion of green colour (RGB = 0, 0, 255; HSB = 120°, 100, 100). Colour categorization included the response 'green.' Twenty-four additional participants (ages 18-34, *M* = 21.5, *SD* = 4.0) were recruited who spoke English as their first language and had normal or corrected-to-normal vision. Experiment 3 also included a test of colour vision using Ishihara plates (colourvisiontesting.com/ishihara.htm); none of the participants had colour vision deficits.

3.7 Results

After removing spoiled trials, median RTs for correct trials and errors were analyzed by-subjects and by-items (see Figures 3.4 and 3.5). For each analysis, a 3 x 2 x 2 ANOVA included Colour (red, green, blue), Temperature (hot, cold), and Task (colour categorization, semantic categorization). Each factor was repeated by-subjects, whereas by-items, Temperature was between items (see Table 3.2 for ANOVA results). Again, the main effect of Task was consistent across subjects and items, which indicated faster responding with colour categorization than semantic categorization. Consistent interactions between Colour and Temperature, both by subjects and by items, supported congruency effects. However, the three-way interaction indicated congruency effects were significantly different between task, with a pattern consistent with congruency effects for semantic categorization (see Figure 3.4), but less so for colour categorization (see Figure 3.5).

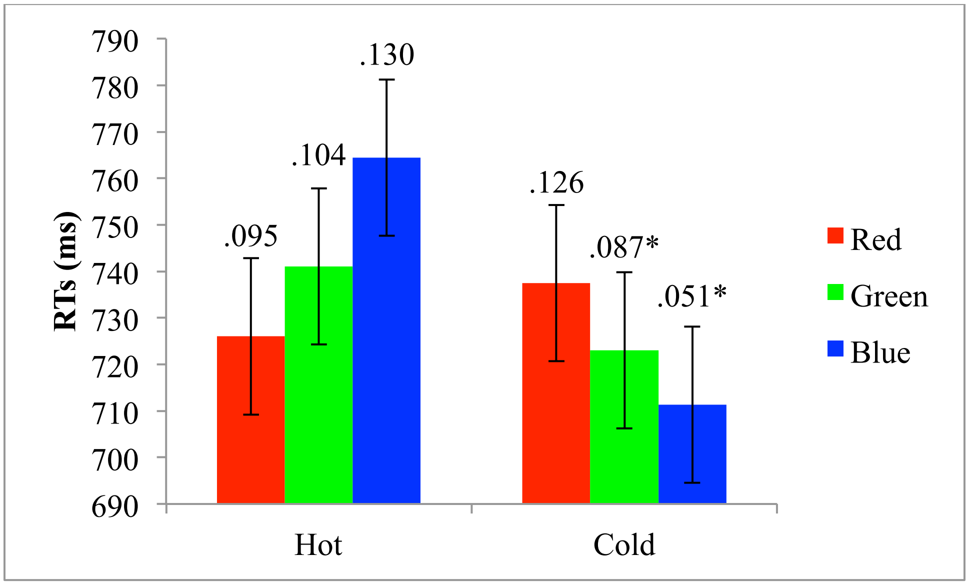

Figure 3.4. Semantic temperature categorization RTs as a function of Colour and Temperature. Errors are displayed above as a proportion out of 1 (95% CI = .031). RTs have a 95% CI of 16.8 ms.

Table 3.2

ANOVA Results for Experiment 3 by-subjects (s) and by-items (i) for Median RTs and Errors

Experiment 3	F_s	F_i	$(df_1, df_2)_s$	$(df_1, df_2)_i$	MSE_s	MSE_i	p_s	p_i
Latency								
Colour	1.73	2.85	2, 46	2, 92	1297.60	1135.17	.190	.063
Temperature	5.32	.55	1, 23	1, 46	2082.51	2627.30	*.030	.465
Task	194.65	849.23	1, 23	1, 46	11737.75	2845.76	*< .001	*< .001
Colour x Temperature	11.58	7.79	2, 46	2, 92	918.10	1135.17	*< .001	*.001
Colour x Task	.52	1.77	2, 46	2, 92	887.31	992.22	.596	.177
Temperature x Task	2.88	.14	1, 23	1, 46	1398.22	2845.76	.103	.711
Colour x Temperature x Task	2.39	4.74	2, 46	2, 92	1362.36	992.22	.103	*.011
Errors								
Colour	.04	.02	2, 46	2, 92	.002	.002	.962	.977
Temperature	.59	1.07	1, 23	1, 46	.007	.004	.450	.305
Task	24.52	60.41	1, 23	1, 46	.009	.004	*< .001	*< .001
Colour x Temperature	14.04	15.66	2, 46	2, 92	.003	.002	*< .001	*< .001
Colour x Task	3.34	4.28	2, 46	2, 92	.003	.002	*.044	*.017
Temperature x Task	1.84	3.36	1, 23	1, 92	.007	.004	.189	.073
Colour x Temperature x Task	3.28	3.78	2, 46	2, 92	.002	.002	*.047	*.026

Note: MSE = mean square error; df = degrees of freedom; * = significant at a $p < .05$ level.

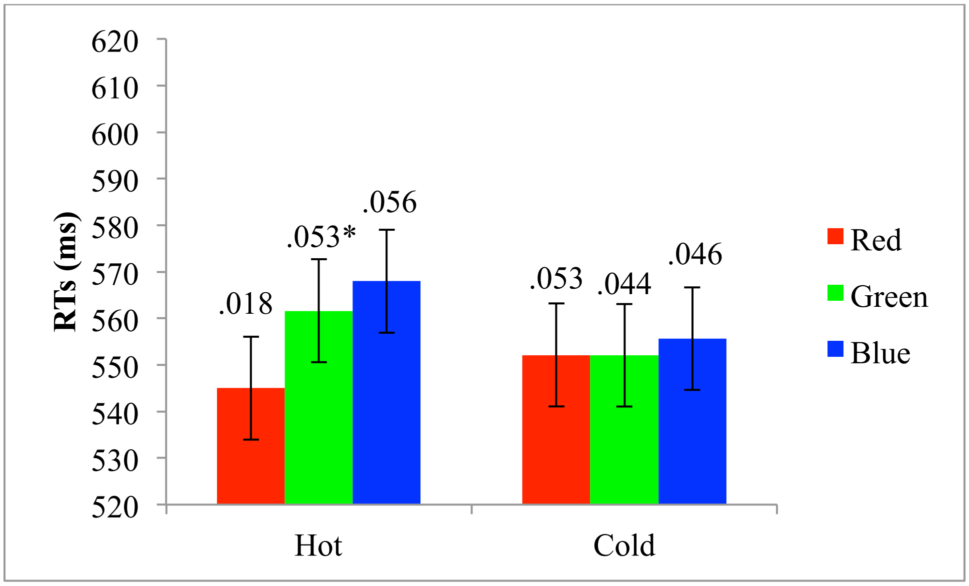

Figure 3.5. Colour categorization RTs as a function of Colour and Temperature. Errors are displayed above as a proportion out of 1 (95% CI = .0128). RT have a 95% CI of 11.0 ms.

Again, Bayesian analyses (e.g., Masson, 2011) were conducted on the average congruency effects separately in each task for median RTs. In the colour categorization task, the average congruency effect of 9.71 ms was compared to 0 in a one-sample *t*-test, $t(23) = 2.546$, $p = .018$, where the Bayes factor in support of the null hypothesis was 2.961, and thus a posterior probability of .747, indicating weak evidence for the congruency effects. In the word categorization task, the average congruency effect of 32.31 ms was compared to 0 in a one-sample *t*-test, *t*(23) = 3.305, *p* = .003, where the Bayes factor in support of the alternative hypothesis was 13.15, and thus a posterior probability of .929, indicating strong evidence for the alternative hypothesis (i.e., congruency effects).

We also tested facilitation and interference separately relative to neutral in both tasks with RTs. To test facilitation in Semantic Categorization (see Figure 3.4), a 2 Temperature (hot, cold) x 2 Congruency (congruent; neutral) ANOVA revealed the effect of Congruency ($M_{\text{congruent}}$ = 718.65, M_{neutral} = 732.03, difference = 13.4 ms) was significant, *F*(1, 23) = 4.388, *p* = .047. For interference, the 2 Temperature (hot, cold) x 2 Congruency (neutral; incongruent) ANOVA showed the effect of Congruency (M_{neutral}= 732.03, $M_{\text{incongruent}}$ = 750.96, difference = -18.9 ms) was significant, *F*(1, 23) = 6.383, *p* = .019. The Temperature x Congruency interactions were not significant, *F's*(1, 23) < .402, *p*'s > .532. With Colour Categorization, the 2 Temperature (hot, cold) x 2 Congruency (congruent; neutral) ANOVA revealed no main effect of congruency, *F*(1, 23) = 1.105, *p* = .304, although the interaction was significant, *F*(1, 23) = 4.514, *p* = .045, suggesting the facilitation effect was different for hot and cold words (see Figure 3.5). Finally, the 2 Temperature (hot, cold) x 2 Congruency (incongruent; neutral) ANOVA revealed no main effect of Congruency, *F*(1, 23) = 1.105, *p* = .304, nor an interaction, *F*(1, 23) = .244, *p* = .626, indicating no effect of interference.

When comparing hot and cold word lists on ratings of temperature association, hot words (*M* = 4.14) had significantly higher ratings than cold words (*M* = 3.78), *t*(46) = 2.61, *p* = .012. To examine whether differences in item ratings could account for our ANOVA results, we re-ran the three-way ANOVA by-items after eliminating items with a rating lower than 3.5 (removing eight items in the cold list, one item in the hot list (see Appendix B). The lists were no longer significantly different by rating ($M_{\text{cold}} = 4.12$, $M_{\text{hot}} = 4.18$; $t(37) = -.573$, $p = .57$). Again, the main effect of Task ($F(1, 37) = 840$, $p < .001$), Colour by Temperature interaction ($F(2, 74) = 6.63$, $p = .002$), and three-way interaction ($F(2, 74) = 4.00$, $p = .023$) remained significant. These results suggest differences between item ratings cannot account for this pattern of effects.

When Task Order was included as a between subjects variable with RTs, the main effect of Task, $F(1, 22) = 3.45$, $MSE = 21715.41$, $p = .077$, and the 4-way interaction, $F(2, 44) = 2.48$, $MSE = 1280.24$, $p = .096$, approached significance, and the remaining interactions were not significant for RTs or errors, Fs(2, 44) < 2.54, ps > .124. As no effects reached significance with Task Order, the pattern of congruency effects should not be attributed to order effects.

3.8 Discussion

Experiment 3 replicated the congruency effects shown in Experiment 2, where the semantic categorization task showed speeded responding with congruent pairings (i.e., red-hot and blue-cold) compared to incongruent pairings (i.e., red-cold and blue-hot). Relative to neutral baseline, significant facilitation was revealed with congruent pairings, and significant interference was revealed with incongruent pairings. With colour categorization, congruency effects were also shown for combinations of red and hot meaning words, but not blue-cold combinations, also replicating Experiment 2. Furthermore, evidence for facilitation was found with red-hot combinations relative to green-hot and blue-hot, but the effect of interference was not found (see Figure 3.6).

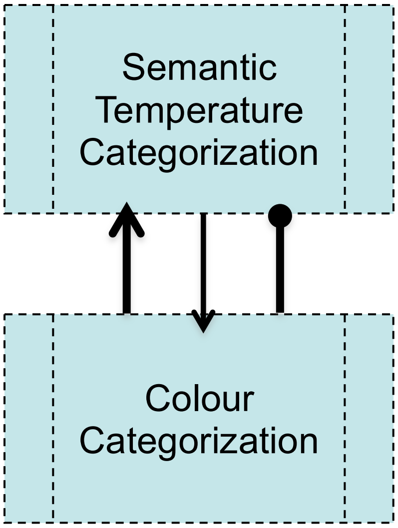

Figure 3.6. Relationship between colour and semantic temperature categorization. Evidence for facilitation (arrow ends) and interference (dotted ends) occurred from colour to semantic categorization for hot and cold words, while evidence for facilitation from semantics occurs only for red-hot colour categorizations.

Given the dissociation between cold-blue congruency effects between colour and semantic categorization tasks, two mechanisms likely support the association between colour and semantic temperature. The association between cold and blue found in semantic categorization supports an abstract association, as this effect was not found with colour categorization, and suggests the grounding of cold meaning in the colour blue. As the red-hot association occurred with colour categorization, this association appears to be driven by direct colour-association of hot words with the colour red. That is, a natural relationship likely exists between hot meaning words and red colour (e.g., ‘lava,’ ‘magma’), conceptually similar to the semantic Stroop effect (whereby ‘blood’ is a colour-associate for red; e.g., Lorentz et al. 2016). With the red-hot association also found with semantic categorization, these results also support the grounding of hot meaning in the colour red.

Overall, our results suggest a robust effect in the direction of colour processing affecting semantic processing, and a less consistent effect in the other direction. These results were consistent across Experiments 1, 2, and 3, in semantic temperature and emotion domains. This asymmetry supports the grounding of temperature-associated and emotion-associated words, whereby concepts described through metaphors may be represented through simulating perceptual processes like colour, a mechanism suggested by the experiential account of language comprehension (e.g., Zwaan, 2004) and grounded cognition (Barsalou, 2008).

However, abstract semantic domains do not only map onto colour processing. To describe the relationship of concept knowledge and perception beyond colour, we chose to investigate how metaphor may ground visual aspects of temporal processing. Linguists have discussed the importance of metaphors related to time, including ‘time as space’, ‘time as size’, and ‘time as motion’ (Lakoff & Johnson, 1999). Such primary metaphors have been show to span across cultures and languages (Ahrens & Huang, 2002). While a number of studies have explored physical correspondences between the domain of time, space, and motion, there is a dearth of literature examining metaphors of time relating semantic representations and visual aspects of duration. The aim of the following study, then, is to generalize and explore the dominance of either perceptual processing or semantic processing to the domain of time.

Chapter 4: Time Stroop Experiment

This chapter is based on the following manuscript under revision, and has been adapted to be consistent with the overall thesis.

Lorentz, E., Gould., L., Ekstrand, C., & Borowsky, R. (2016). A Matter of Time: Time associate words show congruency with time perception in a Stroop-like paradigm. Under review at *Acta Psychologica.*

Whether judging a safe interval to cross the street or recalling how long an event lasted as an eyewitness, the passage of time and its perception has regular impact on how we navigate our world and make decisions regarding duration. The perception of time, however, does not rely on a single sense organ, nor does it rely on an external stimulus like vision or audition. Rather, the passage of time relies on an integration of internal and external cues, including planning and executing motor sequences, viewing the progression of visual stimuli, and the unfolding of mental operations (Fraisse, 1984; see Grondin, 2010, for a review). According to a theory of magnitude (ATOM), time duration may share representation with size and numerosity as a broader construct of magnitude (Walsh, 2003). We aimed to test whether semantic representation also recruits magnitude in the semantic domain of time using a Stroop-like task.

The representation of size and numerosity have previously been integrated in ATOM, where congruency effects in judgments of spatial size and numerosity indicate their shared representation in the common format of magnitude (Walsh, 2003). Recently, ATOM has been extended to accommodate properties of luminance (Kadosh, Kadosh, & Henik, 2008) and duration (Hayashi, Kanai, Tanabe, & Sadato, 2015; Vicario, 2011), bringing support for shared representations of magnitude across space, time, and quantity (see Bueti & Walsh, 2009 for a review). For instance, Vicario (2011) has shown that digit primes can bias estimating the midpoint of a duration (i.e., time bisection) for suprasecond intervals, whereby low digits decrease the bisection point, and larger digits increase the bisection point. Xuan, Zhang, He, and Chen (2007) also showed that observers judged larger magnitudes to last longer with the increased number of dots, size of open squares, luminance of squares, and larger numeric digits using a Stroop-like paradigm.

Similarly, Dormal, Seron, and Pesenti (2006) found evidence for numerosity interference in a duration discrimination task, whereby participants indicated if a target duration was smaller or larger than a standard duration. The number of dots displayed for the target duration produced both interference and facilitation in duration judgements relative to a neutral number of dots, particularly when discriminating smaller differences from the standard durations compared to larger differences. In the reverse task, where participants judged differences in the number of dots from a standard, duration differences did not produce congruency effects with numerosity discriminations. With the possibility of asymmetries between these modalities, these effects support a common mechanism in representing magnitudes between modalities. Indeed, common neural mechanisms have also been implicated across time and numerosity, particularly in the parietal cortex (Pinel, Piazza, Le Bihan & Dehaene, 2004; Pouthas et al. 2005).

The study of time has also been investigated with semantic representation in relation to metaphor. For instance, spatial attention and time estimation show an association depending on the spatial metaphors a language employs. Casasanto and colleagues (2004) compared different language speakers who used either a dominant linear metaphor in their native language (e.g., long time; English and Indonesian) to speakers using a dominant quantity metaphor (e.g., much time; Greek and Spanish) in their native language on a time estimation task, using both growing lines (linear) and filling containers (quantity) as stimuli. Responses to these non-linguistic stimuli varied depending on the speaker's native language, whereby growing lines only modulated time estimations for speakers using linear metaphors and filling containers only modulated time estimations for speakers using quantity metaphors. These results suggest native language can influence the spatial representations we use with time in memory and cognition.

Semantic concepts of time have also been studied more directly using semantic categorizations. When categorizing word tense as either past or future, a bias in spatial attention was revealed, whereby participants were faster to categorize past tense words presented to the left than the right, and the future tense showed the reverse relationship (Torralbo, Santiago, &

Lupianez, 2006). A similar relationship was also shown linking the past to the back and future to the front of space, indicating these spatial-temporal biases are malleable to orienting in the participant's perspective. Such categorizing tasks provide a method to investigate common representation between modalities, and reveal the association between semantic processing and perceptual processes.

Time perception is a rapid process that shows common mechanisms with motor production and regulating action generation (Keele, Pokorny, Corcos, & Ivry, 1985; Rao, Mayer, & Harrington, 2001) and it also moderates the integration of sensory information, suggesting that it is automatic (Poppel, 1997). As time-perception operates on a low level, it is a potential candidate in grounding high level semantic processing of time associated words, in keeping with the experiential account of language comprehension. According to Ahrens & Huang, (2002), the words we use to describe time involve metaphors around amount (e.g., a lot of time), motion (e.g., time is passing by), and velocity (e.g., time is slowly dragging on; see also Lakoff & Johnson, 1999). By using words with these metaphorical association to time, we asked if duration and semantic time-associated words would interact in a similar way to temperature and colour.

We employed a Stroop-like task in which time associated words (e.g., quick, infinite), were presented for both short (200 ms) and longer (367 ms) durations. Short and long duration-associated words were paired with 200 ms and 367 ms durations to create congruent and incongruent stimuli for duration and semantic categorization tasks.

4.1 Hypotheses

Given support for common magnitude representations ATOM (e.g., Walsh, 2003), we hypothesized an association of time perception and semantic meaning through common magnitude representations. Given the literature supporting magnitude cues like numerosity, luminance, or spatial cues impacting time estimates (e.g., Xuan et al. 2007), we hypothesized that congruency effects would be dominant in the duration categorization task over the semantic categorization task. However, to the extent that abstract concepts like time words share processing with basic perception (e.g., Barsalou, 2008), congruency effects were hypothesized in the semantic categorization task. Notably, research in time perception has dominantly relied on error analyses (see Grondin, 2008). Our analyses will thus focus on both RTs and errors for better comparison to the present experiments and the past literature.

4.2 Method

4.2.1 Participants

Twenty-eight undergraduates (*M* = 19.75 years; *SD* = 2.63, range 17-28) were recruited from the University of Saskatchewan psychology participant pool and given course credit for their participation. All participants had normal or corrected-to normal vision, and spoke English as their first language. The experiment was approved by the University of Saskatchewan Research Ethics Board.

4.2.2 Stimuli

Twenty-four short meaning words and 24 long meaning words (see Appendix C) were selected through a systematic search of synonyms for ‘brief,’ ‘instant,’ ‘forever,’ and ‘eternal’ based on their metaphorical associations with time (i.e., spatial and magnitude associations). Lists were matched for length, $t(46) = .965$, $p = .339$ and word frequency, $t(46) = .982$, $p = .331$ using eLexicon (Balota, Yap, Cortese, Hutchison, Kessler, Loftus et al. 2007). Each word, shown in size 18 Arial font and white colour, was presented once for 200 ms and once for 367 ms for 96 total stimuli in a task block.

4.2.3 Apparatus. The apparatus was identical to Experiment 1.

4.2.4 Design and Procedure

The experiment design and procedures were the same as the previous experiments with the following exceptions. To minimize the potential for expectations to affect time perception, the fixation-cross remained on the screen for a randomly selected interval (350 ms, 400 ms, or 450 ms) before it was replaced by the stimulus on a black background. Participants also used button press (e.g., press left for 200 ms duration, or right for 367 ms duration) instead of vocal response, comparable to prior research (e.g., Dormal et al. 2006; Xuan et al. 2007), with left and right buttons to make duration categorizations. Because of the difficulty of the task with time categorization, participants were also presented with feedback after each trial, including 'correct,' in blue font and 'incorrect,' in yellow font, which remained until the participant pressed both buttons to initiate the next trial. Pressing both buttons reduced the effect of prior trial responses on the current trials. Assignment of buttons was counterbalanced across participants, and short meaning and brief duration responses did not switch buttons between blocks.

4.3 Results

Errors and correct median RTs were analyzed by subjects using a 2 x 2 x 2 repeated measures ANOVA, with Duration (200 ms, 367 ms), Meaning (short, long), and Task (duration categorization, semantic categorization) as repeated measures factors (see Figures 4.1 and 4.2). By-items, errors and correct median RTs were also analyzed with Meaning as a between-items factor (see Table 4.1 for ANOVA results). Overall, the main effect of Duration indicated faster responding with 200 ms duration trials than 367 ms duration trials, and the main effect of Meaning indicated faster responding with short meaning words than long meaning words. The

Duration by Meaning interaction, consistent across both RTs and errors, by subjects and by items, supports congruency effects. However, this pattern appears more consistent for duration categorization (see Figure 4.2) than semantic categorization (see Figure 4.1), supported by a three-way interaction on RTs by subjects. Additionally, the interaction between Duration and Task indicated that longer durations produced larger RTs, but particularly for Duration categorization.

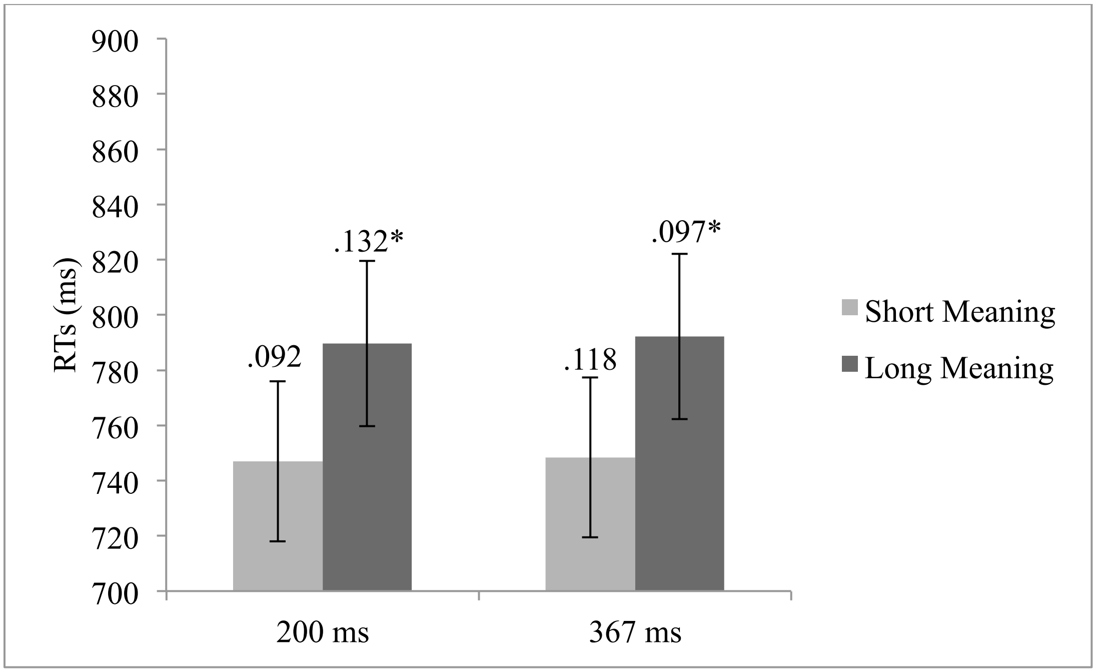

Figure 4.1. Semantic time categorization RTs as a function of Duration and Time Meaning. Errors are displayed above as a proportion out of 1 (95% CI = .021). RT have a CI of 29.0 ms.

Table 4.1

ANOVA Results for Experiment 4 by-subjects (s) and by-items (i) for Median RTs and Errors

Experiment 4	F_s	F_i	$(df_1, df_2)_s$	$(df_1, df_2)_i$	MSE_s	MSE_i	p_s	p_i
Latency								
Duration	14.01	69.16	(1, 27)	(1, 46)	4669.74	1386.14	*.001	*< .001
Meaning	4.39	4.70	(1, 27)	(1, 46)	5979.71	6926.44	*.046	*.035
Task	.23	.02	(1, 27)	(1, 46)	43294.61	7831.65	.634	.894
Duration x Meaning	4.73	6.93	(1, 27)	(1, 46)	2802.85	13886.14	*.038	*.012
Duration x Task	13.06	10.90	(1, 27)	(1, 46)	4432.54	1727.55	*.001	*.002
Meaning x Task	3.14	3.05	(1, 27)	(1, 46)	8268.89	7831.65	.088	.087
Duration x Meaning x Task	6.36	.91	(1, 27)	(1, 46)	2240.55	1727.55	*.018	.345
Errors								
Duration	19.23	26.52	(1, 27)	(1, 46)	.007	.005	*< .001	*< .001
Meaning	.03	.01	(1, 27)	(1, 46)	.004	.009	.863	.912
Task	11.96	13.78	(1, 27)	(1, 46)	.002	.011	*.002	*.001
Duration x Meaning	22.80	14.34	(1, 27)	(1, 46)	.003	.005	*< .001	*< .001
Duration x Task	27.38	23.58	(1, 27)	(1, 46)	.006	.002	*< .001	*< .001
Meaning x Task	.70	.28	(1, 27)	(1, 46)	.005	.011	.409	.599
Duration x Meaning x Task	.60	.34	(1, 27)	(1, 46)	.004	.006	.445	.563

Note: MSE = mean square error; df = degrees of freedom; * = significant at $p < .05$.

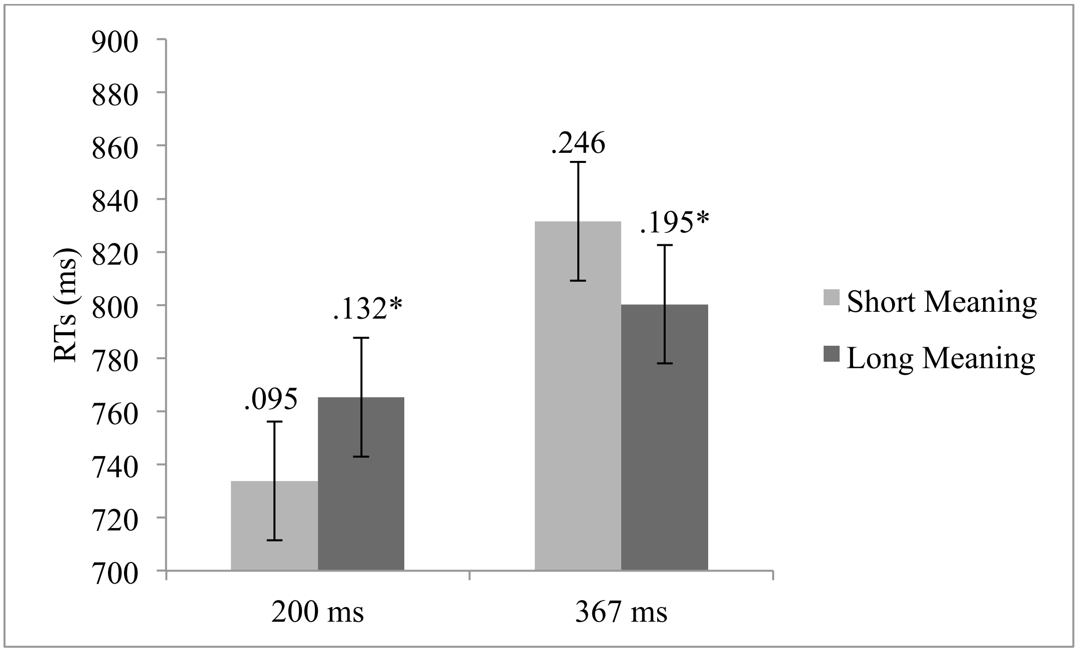

Figure 4.2. Duration categorization RTs as a function of Duration and Time Meaning. Errors are displayed above as a proportion out of 1 (95% CI = .034). RTs have a 95 % CI of 22.3 ms.

Bayesian analyses (e.g., Masson, 2011) evaluated the average congruency effect on errors for each task. The average duration categorization congruency effect of .044 was compared to 0 in a one-sample *t*-test, $t(27) = 3.13$, $p = .004$, with a Bayes factor for the null hypothesis of 8.47, a posterior probability of .894, indicating positive to strong evidence for the alternative hypothesis (a congruency effect), or a posterior probability of .106 for the null hypothesis. The semantic categorization congruency effect of .031 was compared to 0 in a one-sample *t*-test, $t(27) = 3.55$, $p < .001$, with a Bayes factor for the alternative hypothesis of 21.90, a posterior probability of .956, indicating strong evidence for the alternative hypothesis, or a congruency effect.

Bayesian analyses also evaluated the average congruency effect on RTs for each task. The average duration categorization congruency effect of 31.35ms was compared to 0 in a one-sample *t*-test, $t(27) = 4.15$, $p < .001$, with a Bayes factor for the null hypothesis of 92.71, a posterior probability of .989, indicating strong to very strong evidence for the alternative hypothesis (a congruency effect), or a posterior probability of .011 for the null hypothesis. The semantic categorization congruency effect of -.56ms was compared to 0 in a one-sample *t*-test, $t(27) = -.051$, $p = .96$, with a Bayes factor for the null hypothesis of 6.846, a posterior probability of .873, indicating positive evidence for the null hypothesis, or no congruency effect.

We compared short and long word lists on ratings of the strength of their temporal association (see Appendix C). Short words had mean rating of 3.96, and long words had a mean rating of 3.86, although these lists did not differ significantly $t(46) = .805$, $p = .425$.

Importantly, when Task Order was included as a between subjects variable, only a Task by Task Order interaction was marginally significant with RTs, $F(1, 26) = 3.60$, $MSE = 39492.16$, $p = .069$. As the remaining interactions with Task Order were not significant for either RTs or errors, $Fs(1, 26) < 1.65$, $ps > .210$, congruency effects should not be attributed to carry-over effects between tasks.

4.4 Discussion

With duration categorization, both errors and RTs strongly supported a congruency effect, such that word meaning systematically affected duration categorization. With short durations, short meaning words yielded faster RTs and fewer errors than long meaning words, and with long durations, long meaning words had shorter RTs and smaller error rates than short meaning words. Thus, this pattern of results reflects the typical Stroop effect, but in the domain of time. As RTs directly mirrored the pattern of errors with duration categorization, they reinforce the interpretation of consistent congruency effects, whereby semantic processing automatically and systematically affects temporal decisions.

The effect of time duration on semantic categorization also supports the interpretation of congruency effects at the level of errors. Within each duration, a congruency effect emerged between word types, and within each word type, congruency effects emerged between durations. The evidence with RTs was mixed, however. The main effect of word type emerged consistently at short and long durations, such that short meaning words were categorized faster than long meaning words. It may be the case that long meaning words were generally more abstract that short meaning words. While this possibility could be followed up with additional research (e.g., ratings on concreteness or tangibility), this potential difference should not bias congruency effects to one task over another. Rather, differences between word lists would be expected to generate effects at the level of a main effect, and not an interaction between word meaning and duration. On average, there was no congruency effect found with semantic categorization on RTs, but there was on duration categorization. This direction of effects supports the dominant processing of semantic meaning in the domain of time (see Figure 4.3).

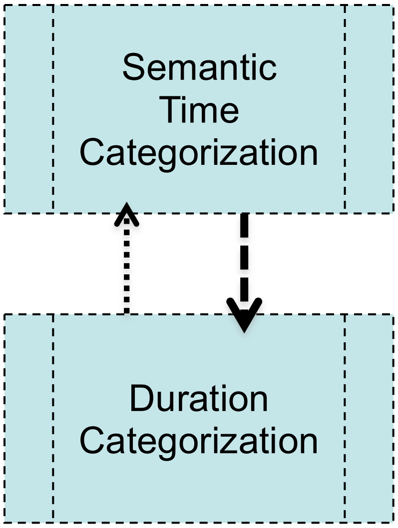

Figure 4.3. Relationship between duration and semantic time categorization. Evidence for both congruency effects occurs from semantic processing to duration categorization for both 200 ms and 367 ms intervals, while evidence for congruency effects occurs only with errors from temporal processing to semantic categorizations.

A strong influence of semantic processing on duration categorization extends the automaticity account of reading into the modality of time perception, even when word meaning was not task relevant. These congruency effects in time perception indicate that time perception tasks are susceptible to word reading influence, whereby a natural relationship occurs between the meaning of words and the physical durations that were categorized, similar to naming in the typical Stroop task (Algom et al. 2004). The word, albeit irrelevant, activates related duration representations, and indicates a connection between semantic processing and duration monitoring, likely overlapping as magnitude representations considering ATOM (Pinel et al. 2004; Walsh, 2003). As monitoring of the duration of events is constantly ongoing, time perception is considered both automatic and low-level (Näätänen, Syssoeva, & Takegata, 2004; Poppel, 1997). When made task relevant, our results corroborate others in that time perception becomes susceptible to irrelevant information like numerosity (e.g., Dormal et al. 2006; Vicario, 2011), spatial information (Torralbo et al. 2006) and, in the present case, semantic word processing. While access to semantic meaning is considered a slower, more high-level process than perceptual processes (Dehaene, Changeux, Naccache, Sackur, & Sergent, 2006), words have been shown to be processed through to semantic meaning, even in the absence of awareness (Lorentz, Gould, Mickleborough, Ekstrand, Boyer, Cheesman & Borowsky, 2015).

The observed congruency effects with semantic categorization suggest time perception processes interact bidirectionally with the semantic network, at least with errors. However, this interaction was not symmetric in the case of RTs. In the semantic categorization task, different durations did not produce consistent congruency effects. The robust effect on errors supports automatic, mandatory processing of temporal features with subsequent impact on semantic categorization, and further reflects the automatic monitoring of duration. This direction of effects supports the experiential account of language comprehension (Barsalou, 2008, Zwaan, 2004), whereby perceptual representations support representations of conceptual knowledge (see also Allport, 1985; Gibbs, 2005). It may be that metaphorical phrases like ‘gone in a flash’ to represent brief are not just useful in explaining the abstract notion of time, but are essential in supporting its semantic representation.

Chapter 5: General Discussion

The common use of metaphor in language indicates a relationship between abstract concepts and basic perceptual processes (Barsalou, 2008; Lakoff & Johnson, 1999; Zwaan, 2004). However, the robust Stroop effect also demonstrates automaticity in accessing semantic meaning in performing perceptual tasks. Four experiments reported here describe the underlying mechanisms supporting the shared representation of perceptual processes and conceptual meaning.

5.1 Summary of Major Findings

Experiment 1 tested the dominant direction of influence in the ‘mad is red’ and ‘sad is blue' metaphor of emotion. In two tasks, participants categorized colour (red or blue) and semantic meaning (mad or sad), where the irrelevant cue (either colour, or word meaning) provided congruent or incongruent trials. Consistent congruency effects on RTs emerged in semantic categorization, but not colour categorization. By errors, this pattern was supported with semantic categorization, but was not consistent in colour categorization. Overall, a dominant influence of colour processing on semantic categorization is suggested, as more consistent congruency effects occurred with semantic categorization (see Figure 2.3).

Experiments 2 and 3 tested the dominant direction of perceptual and conceptual influences with the 'red is hot' and 'blue is cold' metaphor of temperature. Experiment 2 provided a conceptual replication of Experiment 1, and Experiment 3 employed neutral trials with the colour green, testing the role of facilitation and interference in each task while replicating Experiment 2. In Experiment 2, congruency effects were consistent in semantic categorization for both hot and cold meaning words across RTs and errors. With colour categorization, congruency effects were only obtained with hot words (see Figure 3.3). In Experiment 3, the same congruency effects occurred, and the effects of facilitation and interference were consistent across hot and cold words with semantic categorization. With colour categorization, a facilitation effect was found for hot words in the colour red compared to green, but congruency effects were not consistent across hot and cold words (see Figure 3.6).

Experiment 4 tested the dominant direction of influence in temporal processing with words metaphorically associated with time, including 'quick' and 'slow.' In semantic categorization, congruency effects were only consistent by errors. In time categorization, however, congruency effects were consistent in short and long meaning words across RTs and errors, indicating the dominant processing of semantic representations influencing perceptual judgments in the domain of time (see Figure 4.3). Overall, each experiment provides new information about the direction of interactions between physical and conceptual processing (see Figure 5.1).

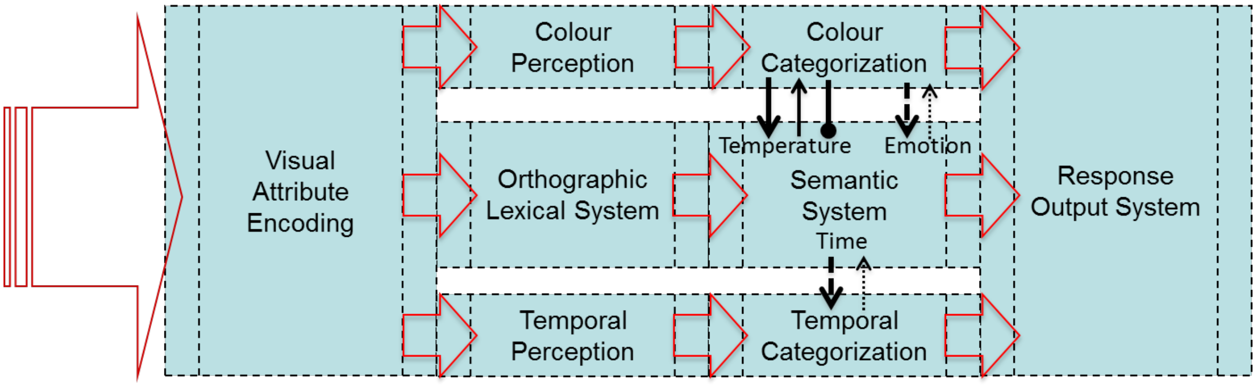

Figure 5.1. A model of visual attribute encoding, reading and semantic processing, perceptual categorization, and response output. This model reflects whole-word and visual attribute processing in parallel, with evidence from four experiments represented by black arrows connecting the semantic system and attribute categorization leading to response output. Dominant influences (dashed lines) and facilitation and interference (solid lines) are shown by line thickness, indicating asymmetries between visual attribute processing and semantic processing before final categorization responses.

In Figure 5.1, a combined model of reading and visual perception categorization, visual features including letter shape, colour, and temporal attributes are first encoded and feed forward into perception and recognition systems. In comparison to the dual-route model of reading (e.g., see Anton et al. 2014) this model focuses on the ventral route, whereby common words are recognized using the orthographic lexical system. For instance, the word ‘scald’ is recognized in the orthographic-lexical system and mapped onto conceptual knowledge in the semantic system (see also Coltheart, Rastle, Perry, Langdon, & Zeigler, 2001, reviewing the dual-route cascaded model). In parallel with word recognition, other visual attributes are encoded, perceptually processed, and recognized in long-term memory for categorization. In this model, words or visual attributes may be categorized in the absence of physical attributes or semantic information, respectively, as indicated by the three processing routes. However, the connections between perceptual categorization (i.e., colour, temporal attributes) and the semantic system in our model reflect the automatic, obligatory influence of related, but task irrelevant information when present. Dominant processes, and asymmetries in processing are indicated by the line thickness of connections, indicating that connection strength varies by visual attribute and semantic domain.

5.2 Evidence for Grounded Cognition

In the domain of colour perception, our results provide strong support for the theory of grounded cognition, namely, that semantic concepts are embodied in overlapping sensorimotor representations (Allport, 1985). Although colour was irrelevant to semantic categorization tasks, and manipulated independently, it automatically and systematically influenced the semantic decisions about temperature and emotion meaning in the form of congruency effects. In the domain of temperature, we also saw evidence for facilitation and interference compared to a neutral baseline. Our evidence suggests that the physicality of abstract concepts in colour is not just helpful in their conceptualization, but rather that these associations are obligatory. In this way, our results support the theory that concepts are built on physical metaphor (Lakoff & Johnson, 1999), and that language comprehension is to some extent embodied in modal processes (Zwaan, 2004).

Congruency effects with colour in the abstract domains of emotion and temperature also bring counterevidence to purely amodal theories of semantic representation. While Quillian's (1968) hierarchical model of semantic memory is intuitively based on efficiency, with propositional organization, the evidence supports modal overlapping in representations, with similarity between simulated features mapping the relationship between concepts (e.g., Rips et al. 1973). However, our evidence indicates these attributes are not represented purely symbolically. Rather, the metaphorical relationship (e.g., ‘feeling blue’) between concepts and perceptual properties like colour indicates grounding semantic meaning in sensorimotor representations (Barsalou, 2008).

This grounded theory of cognition fits well with theories of brain evolution. For instance, the representation of abstract concepts builds onto pre-existing structures that are more stimulus-oriented in their representations (MacLean, 1990). In the course of development, abstract representations are also thought to build upon sensorimotor representations (Piaget & Inhelder, 1969). Indeed, Johnson’s (1996) theory of conflation suggests that abstract and concrete concepts in metaphors like ‘red-hot’ are first conflated as children learn them, and only distinguished in a later period of differentiation. Overall, this asymmetry between abstract and concrete representations supports how conceptual knowledge may have a perceptual basis.

5.3 Evidence for Stroop Effects in Perceptual Tasks

The Stroop effect occurs when differences between a word’s semantic meaning (e.g., ‘banana’) and its physical attribute (in this case blue font colour) create delayed responding or increased errors with incongruent compared to congruent pairs (Algom, Chajut, & Lev, 2004; Lorentz et al. 2015; MacLeod, 1991). Thus, the Stroop effects reported here for perceptual categorization reflect a learned relationship between the semantic referent of the carrying word and the target visual attribute, whether colour or duration. Thus, congruency effects in the perceptual task likely reflect more literal, concrete associations (e.g., ‘magma’ in red) as opposed to metaphorical associations (e.g., ‘grief’ is blue). When performing perceptual tasks with colour and temporal categorization, we found congruency effects particularly for hot words across red, blue, and green colours, within each duration for time words comparing short and long meaning words, and with some words in the ‘mad’ meaning list that replicate prior emotion-related Stroop effects (i.e., ‘anger’; replicating effects shown by Sutton & Altarriba, 2008). These Stroop effects dissociated from congruency effects in the semantic categorization tasks across experiments, suggesting the mechanism driving Stroop effects is not the same mechanism driving semantic categorization congruency effects. While Stroop effects appear to reflect direct associations between semantic meaning and physical attributes, Grady (1997) describes the relationship of primary metaphors where concepts recruit sensorimotor attributes in their representation. To further investigate these differences, a measure of colour association, in addition to our ratings of semantic association, may help to disentangle the mechanism of these congruency effects.

As Stroop effects in our perceptual categorization tasks appear to be independent from congruency effects in semantic categorization, an alternative interpretation may posit an amodal mechanism driving Stroop effects, while a modal mechanism is driving congruency effects in the semantic categorization task. Louwerse and Jeuniaux (2010) bring evidence that both linguistic factors (amodal) are independent from modal processing when performing different categorization tasks. When participants performed semantic relatedness judgments for words, typical and atypical word order (e.g., ‘beer-foam’ versus foam-beer) predicted RTs and errors, in addition to modal factors like spatial congruency (e.g., ‘foam’ on top and ‘beer’ on bottom versus ‘foam’ on bottom and ‘beer’ on top). Interestingly, Louwerse and Jeuniaux (2010) found that linguistic factors were more predictive when participants responded to word pairs than to matched picture pairs, and spatial iconicity was more predictive with picture stimuli than linguistic factors. Our perceptual and semantic categorization tasks, revealing both visual attribute and semantic influences, fits well with their results in suggesting both task and stimulus attributes can be predictive of congruency effects driven by modal and semantic factors. Still, a future direction for the present experiments could identify both linguistic and modal measures to capture their independent contributions to congruency effects.

5.4 Boundary Conditions, Limitations, and Future Directions

While a clear association, and asymmetry in influence, is present in our data for colour-based metaphors, there are boundaries in their support for grounded cognition. First, the data may not establish a purely semantic locus where congruency effects occur. For instance, it is possible that colour cues activate responses like ‘hot’ or ‘mad,’ demonstrating a response conflict instead of activating competing long-term memory representations (Zhang, & Kornblum, 1998). One future direction to investigate this hypothesis might include a lexical-decision task. If participants are faster to verify a word is correctly spelt when presented in a congruent colour, this evidence would remove the possibility of overlapping responses and further isolate a semantic locus (e.g., Meier et al. 2004).

A second boundary condition for the hypothesis of grounded cognition involves the relative automaticity of perceptual processes. Both colour processing and temporal processing are considered relatively low-level, fast processes (e.g., Lorentz, et al. 2015; Pöppel, 1997), although the processing of time does not appear to produce interference effects to the same extent as colour or numerosity (Dormal et al. 2005). In this way, the processing of different perceptual modalities is not equally mandatory, at least not in a way systematically producing congruency effects. This may relate to the stage of processing that is affected by congruency effects. While colours may activate semantic or response representations by their association with semantic concepts, the same may not be true about temporal processing. In this way, the grounding of concepts does not appear to occur equally across all modalities of visual perception.

The present experiments have strong ties to the prior research in their respective domains of literature. For instance, experiments in the domains of emotion (e.g., Fetterman et al. 2012; Sutton & Altarriba, 2008) and temperature (e.g., Ho et al. 2014) have included larger word lists and overt naming, which the present experiments reflect directly. Our experiment regarding temporal processing also directly ties to prior research (e.g., Dormal et al. 2006; Xuan et al. 2007) by using button press and focusing on errors analysis. In reflecting the approaches in the prior literature, the present experiments have the advantage of comparing back to research in their respective domains.

Even so, the methods between the domains of time, temperature and emotion do not perfectly overlap, presenting limitations in generalizing our results across domains. For instance, the emphasis on error analyses in time research (Grondin, 2008) is not reflected in traditional Stroop paradigms, where RTs are the primary unit of analysis (e.g., Besner, Stolz, & Boutilier, 1997; Lorentz et al. 2016; MacLeod, 1991). Thus, congruency effects on errors, but not RTs, may be differently described in the emotion experiment as less robust, and more reliable in the time experiment. By analyzing both RTs and errors, as well as generalizing to both items and subjects, our design bridges these disparate literatures to come to common conclusions around the direction of influence more broadly. It would be prudent, though, to conduct follow-up experiments independently emphasizing RTs (at the cost of errors) and accuracy (at the cost of RTs) to see if both measures still reflect the same underlying mechanism (e.g., Meier et al. 2004). Generally, errors and RTs reflect the same processing, especially at a response selection stage, although qualitative differences may emerge with perceptual processing mechanisms (Pashler, 1989). Similarly, vocal and button press responses also show largely the same patterns of results, although there are reported instances of response modality differences (see MacLeod, 1991, for a review). Indeed, the generated conflict when processing multiple sources of information shows similar areas of activation in the anterior cingulate cortex between response modalities (Barch, Braver, Akbudak, Conturo, Ollinger, & Snyder, 2001; Lorentz, Mickleborough, Mendez, Gould, Ekstrand, Ellchuk, & Borowsky, submitted).

Another area for future research involves the use of neutral baselines. Stimuli using temperature words showed facilitation and interference with semantic categorization, and facilitation with colour categorization. While this additional level of detail is useful, the initial investigations employed two meanings and two colours or durations for two main reasons. First, neutral conditions may dilute the congruency effect, reducing the overall power (Lorentz et al. 2016). Secondly, the neutral condition creates inherent differences between tasks (i.e., 2 responses in the semantic categorization task, 3 responses in the colour categorization task in Experiment 3, which should only be used when a comparison to equal responses is possible (i.e., Experiment 2). An important future direction may explore how congruency effects would manifest in relation to facilitation and interference with the remaining modalities (e.g., with emotion words in Experiment 1, currently being pursued). Finally, we note the present experiments used only two levels of visual attributes and two levels of semantic meaning, creating an equal proportion of congruent and incongruent stimuli. Although more levels may be possible (e.g., 4 with Sutton and Altarriba, 2008), this allows more ways to create incongruent than congruent stimuli, whereby multiple baselines must be used to separately estimate facilitation, interference, and contingency effects (i.e., differences in repetition between conditions; Lorentz et al. 2016). Thus, the present experiments maximized the effects with the simplest design.

Another limitation inherent to the perceptual and semantic categorization tasks involves potential differences in task difficulty. Although Task effects were not consistently found, colour categorizations were typically faster than semantic categorizations in Experiments 1, 2, and 3. Such differences caution the interpretation of asymmetries in congruency effects. One explanation is that a deeper level of processing may occur with semantic categorizations than perceptual categorizations (e.g., Craik & Lockhart, 1972). If this is the case, modal visual attributes appear to influence deeper cognitive tasks, whereas word meaning and linguistic factors may influence shallower processes (see also Louwerse & Jeuniaux, 2010). One approach to overcoming these differences is the process dissociation procedure (Lindsay & Jacoby, 1994). Similar to the approach we took in reversing the tasks in each experiment, the process dissociation procedure involves independently manipulating either the perceptual or conceptual elements to estimate their contribution to congruency effects. For instance, Lindsay and Jacoby (1994) found that asymmetries in facilitation and interference were eliminated when degrading the colour quality of Stroop stimuli relative to non-letter control items. Moreover, colour contributions remained constant when they manipulated the proportion of congruent stimuli, despite the changing contribution of word reading processes. Thus, the process dissociation procedure would provide a complimentary next step in describing the mechanism of the association described in the present experiments, while using a second method to extend the present results.

Another notable limitation in the present experiments includes the use of new word lists to represent different semantic domains. While the ratings and accuracy measures converge to indicate each word fit its category as one meaning or the other in the vast majority of cases, some were less clear (e.g., ‘numb’ for cold, or ‘sluggish’ for short meaning). It may be valuable to only include highly rated (e.g., greater than 3.5 out of 5) in future studies. Even so, the use of ratings for association raises the question of what these ratings really capture. Words may possess attributes of saliency, intensity, or other related attributes of strength in the semantic network, suggesting ratings of association is just one of multiple important word qualities (Sutton & Altarriba, 2008), each related to semantic categorization. Another limitation of these lists involves their association with metaphor. Metaphors are often used in common phrases like 'over my head' (Lakoff & Johnson, 1999), whereas words on their own may be examined more literally, lending themselves to congruency effects in the perceptual categorization task more than the semantic categorization task. For instance, consistent perceptual congruency effects in the time categorization task may reflect the literal association between words and durations, in addition to their metaphorical relationship. Although the words chosen reflect the 'time is motion' metaphor (e.g., quick, rapid, speedy; continual, enduring, slow), other words may reflect more literal associations (e.g., brief, instant; forever). Errors in the time experiment suggest a bi-directional relationship of duration and semantic processing influences, indicating automatic processing of these cues regardless of the task. Thus, it would be beneficial to separate the influences of direct and metaphorical associations in subsequent experiments. An interesting future direction may employ phrases with metaphorical relationships to physical attributes, while continuing to have participants categorize their meaning. This would serve to isolate the metaphorical relationship from the direct associations found with some words.

Lastly, the use of mixed congruent and incongruent trials can produce differences on a trial-by trial basis that may be of interest in future experiments. For instance, participants in traditional Stroop tasks show slowed responding and increased accuracy following an error, known as the Rabbitt/Laming effect (Laming, 1979; Rabbitt, 1966), which similarly occurs following a conflict trial, known as the Gratton effect (Gratton, Coles, & Donchin, 1992). When presenting feedback, these effects may be enhanced (Kluger & DeNisi, 1996), suggesting our experiment on time may have emphasized accuracy more than the experiments without feedback. Thus, an important direction to pursue would include emphasizing either reaction time or accuracy to test whether each measure consistently reflects the same effect, and/or how speed accuracy tradeoffs may differ between modalities (e.g., Meier et al. 2004).

5.5 Conclusion

These four studies sought to describe the dominant direction of processing influence underlying shared perceptual and conceptual processing in the semantic network. Using both perceptual and semantic categorization tasks, evidence for processing influence was found in both directions, although marked asymmetries were found across domains. Perceptual categorizations show congruency effects with semantically related, but not task relevant words when a direct association exists between co-occurring domains. With semantic categorization, congruency effects occurred more consistently when more abstract associations (e.g., ‘anger’ is ‘mad’; ‘frigid’ is ‘cold’) took place. Overall, these results support the notion of grounded cognition, providing a framework for relating sensorimotor attributes and experience with linguistic and semantic processing.

References

Ahrens, K., & Huang, C. R. (2002). Time passing is motion. *Language & Linguistics*, *3*, 491.

Algom, D., Chajut, E., & Lev, S. (2004). A rational look at the emotional Stroop phenomenon: a generic slowdown, not a stroop effect. *Journal of experimental psychology: General*, *133*(3), 323.

Allport, D. (1985). Distributed memory, modular subsystems and dysphasia. In S. K. Newman & R. Epstein (Eds.), *Current Perspectives in Dysphasia,* 32-60. Edinburgh: Churchill Livingstone.

Anton, K. F., Gould, L., & Borowsky, R. (2014). Activation of lexical and semantic representations without intention along GPC-sublexical and orthographic-lexical reading pathways in a Stroop paradigm. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 40*(3), 623. http://dx.doi.org/10.1037/a0035154

Balota, D. A., Yap, M., Cortese, M., Hutchison, K., Kessler, B., Loftis, B., Neely, J. H., Nelson, D. L., Simpson, G .B., & Treiman, R. (2007). The English lexicon project. *Behavior Research Methods, 39*, 445–459. doi:10.3758/BF03193014

Barch, D. M., Braver, T. S., Akbudak, E., Conturo, T., Ollinger, J., & Snyder, A. (2001). Anterior cingulate cortex and response conflict: effects of response modality and processing domain. *Cerebral Cortex*, *11*(9), 837-848.

Barsalou, L. W. (2008). Grounded cognition. *Annual Review of Psychology, 59*, 617-645. doi: 10.1146/annurev.psych.59.103006.093639

Barsalou, L., & Hale, C. (1993). Components of conceptual representation. From feature lists to recursive frames. In I. Van Mechelen, J. Hampton, R. Michalski & P. Theuns (eds.), *Categories and Concepts: Theoretical Views and Inductive Data Analysis.* Academic

Press 97-144.

Berkowitz, L. (1993). Aggression: Its causes, consequences, and control. Mcgraw-Hill Book Company.

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

Boyatzis, C. J., & Varghese, R. (1994). Children's emotional associations with colours. *The Journal of genetic psychology*, *155*(1), 77-85.

Bueti, D., & Walsh, V. (2009). The parietal cortex and the representation of time, space, number and other magnitudes. *Philosophical Transactions of the Royal Society of London B: Biological Sciences, 364*(1525), 1831-1840.

Carr, T. H. (1992). Automaticity and cognitive anatomy: Is word recognition" automatic"?. *The American journal of psychology*, 201-237.

Casasanto, D., Boroditsky, L., Phillips, W., Greene, J., Goswami, S., Bocanegra-Thiel, S., … & Gil, D. (2004). How deep are effects of language on thought? Time estimation in speakers of English, Indonesian, Greek, and Spanish. *In Proceedings of the 26th annual meeting of the Cognitive Science Society* (186-191).

Changizi, M. A., Zhang, Q., & Shimojo, S. (2006). Bare skin, blood and the evolution of primate colour vision. *Biology letters, 2*(2), 217-221. doi: 10.1098/rsbl.2006.0440

Coltheart, M., Rastle, K., Perry, C., Langdon, R., & Ziegler, J. (2001). DRC: a dual route cascaded model of visual word recognition and reading aloud. *Psychological review*, *108*(1), 204.

Collins, A. M., & Loftus, E. F. (1975). A spreading-activation theory of semantic processing. *Psychological review, 82*(6), 407. http://dx.doi.org/10.1037/0033-295X.82.6.407

Craik, F. I., & Lockhart, R. S. (1972). Levels of processing: A framework for memory research. *Journal of verbal learning and verbal behavior*, *11*(6), 671-684.

Dehaene, S., Changeux, J. P., Naccache, L., Sackur, J., & Sergent, C. (2006). Conscious, preconscious, and subliminal processing: a testable taxonomy. *Trends in cognitive sciences, 10*(5), 204-211. doi:10.1016/j.tics.2006.03.007

Dennett, D. C. (1988). Quining qualia. *Consciousness in modern science*.

Dormal, V., Seron, X., & Pesenti, M. (2006). Numerosity-duration interference: A Stroop experiment. *Acta psychologica, 121*(2), 109-124.

Fetterman, A. K., Robinson, M. D., & Meier, B. P. (2012). Anger as “seeing red”: evidence for a perceptual association. *Cognition & emotion, 26*(8), 1445-1458. doi:10.1080/02699931.2012.673477

Fraisse, P. (1984). Perception and estimation of time. *Annual review of psychology, 35*(1), 1-37.

Gibbs Jr, R. W. (2005). Embodiment and cognitive science. Cambridge University Press.

Glenberg, A. M., & Gallese, V. (2012). Action-based language: A theory of language acquisition, comprehension, and production. *Cortex*, *48*(7), 905-922.

Grady, Joseph, 1997. Foundations of Meaning: Primary Metaphors and Primary Scenes. Ph.D. dissertation, University of California, Berkeley.

Gratton, G., Coles, M. G., & Donchin, E. (1992). Optimizing the use of information: strategic control of activation of responses. *Journal of Experimental Psychology: General*, *121*(4), 480.

Grondin, S. (2008). Methods for studying psychological time. Psychology of time, 51-74.

Grondin, S. (2010). Timing and time perception: a review of recent behavioral and neuroscience findings and theoretical directions. *Attention, Perception, & Psychophysics, 72*(3), 561-582.

Hansen, T., Olkkonen, M., Walter, S., & Gegenfurtner, K. R. (2006). Memory modulates colour appearance. *Nature neuroscience*, *9*(11), 1367-1368.

Hayashi, M. J., Kanai, R., Tanabe, H., & Sadato, N. (2015). Time reproduction and numerosity interaction in the parietal cortex: some missing links. *Arm and Hand Movement: Current Knowledge and Future Perspective, 4*(45), 150.

Hemphill, M. (1996). A note on adults' colour–emotion associations. *The Journal of genetic psychology, 157*(3), 275-280. doi :10.1080/00221325.1996.9914865

Ho, H. N., Van Doorn, G. H., Kawabe, T., Watanabe, J., & Spence, C. (2014a). Colour-temperature correspondences: When reactions to thermal stimuli are influenced by colour. *PloS one, 9*(3), e91854. doi:10.1371/journal.pone.0091854

Johnson, C. (1996). Learnability in the acquisition of multiple senses: SOURCE reconsidered. In *Annual Meeting of the Berkeley Linguistics Society* (Vol. 22, No. 1, pp. 469-480).

Kadosh, R. C., Kadosh, K. C., & Henik, A. (2008). When brightness counts: The neuronal correlate of numerical–luminance interference. *Cerebral Cortex, 18*(2), 337-343.

Kahneman, D., & Chajczyk, D. (1983). Tests of the automaticity of reading: Dilution of Stroop effects by colour-irrelevant stimuli. *Journal of Experimental Psychology: Human perception and performance*, *9*(4), 497.

Kan, I. P., Barsalou, L. W., Olseth Solomon, K., Minor, J. K., & Thompson-Schill, S. L. (2003). Role of mental imagery in a property verification task: fMRI evidence for perceptual representations of conceptual knowledge. *Cognitive Neuropsychology*, *20*(3-6), 525-540.

Keele, S. W., Pokorny, R. A., Corcos, D. M., & Ivry, R. (1985). Do perception and motor production share common timing mechanisms: A correlational analysis. *Acta psychologica, 60*(2), 173-191.

Klein, G. S. (1964). Semantic power measured through the interference of words with colour-naming. *The American journal of psychology*, 576-588. doi: 10.2307/1420768

Kluger, A. N., & DeNisi, A. (1996). The effects of feedback interventions on performance: a historical review, a meta-analysis, and a preliminary feedback intervention theory. *Psychological bulletin*, *119*(2), 254.

LaBerge, D., & Samuels, S. J. (1974). Toward a theory of automatic information processing in reading. *Cognitive psychology*, *6*(2), 293-323.

Lakoff, G., & Johnson, M. (1999). Philosophy in the flesh. New York, NY: Basic Books.

Lambie, J. A., & Marcel, A. J. (2002). Consciousness and the varieties of emotion experience: A theoretical framework. *Psychological Review*, *109,* 219-259. http://dx.doi.org/10.1037/0033-295X.109.2.219

Laming, D. (1979). Autocorrelation of choice-reaction times. *Acta Psychologica*, *43*(5), 381-412.

Lewinski, R. J. (1938). An investigation of individual responses to chromatic illumination. *The Journal of Psychology*, *6*(1), 155-160. doi:10.1080/00223980.1938.9917592

Lindsay, D. S., & Jacoby, L. L. (1994). Stroop process dissociations: The relationship between facilitation and interference. *Journal of Experimental Psychology: Human Perception and Performance*, *20*(2), 219.

Locke, J. (1690/1924). *An essay concerning human understanding*. New York: Prometheus Books.

Loftus, G. R., & Masson, M. E. (1994). Using confidence intervals in within-subject designs.

Psychonomic bulletin & review, 1(4), 476-490. doi:10.3758/BF03210951

Lorentz, E., Ekstrand, C., Gould, L., & Borowsky, R. (in press). Red-hot! How colour and semantic temperature processing interact in a Stroop-like paradigm. *Visual Cognition*.

Lorentz, E., Gould., L., Ekstrand, C., & Borowsky, R. (2016). A Matter of Time: Time associate words show congruency with time perception in a Stroop-like paradigm. Under review at *Acta Psychologica.*

Lorentz, E., Gould., L., Ekstrand., E., Mickleborough. M., & Borowsky. R. (2016). Feeling blue' and 'seeing red': Evidence grounding discrete emotions in colour perception from a Stroop paradigm. Under review at *Cognition and Emotion.*

Lorentz, E., Gould, L., Mickleborough, M., Ekstrand, C., Boyer, M., Cheesman, J., & Borowsky, R. (2015). All in one fell Stroop: Examining consciousness thresholds with a multiple response paradigm. *Psychology of Consciousness: Theory, Research, and Practice, 2*(2), 111-133. http://dx.doi.org/10.1037/cns0000049

Lorentz, E., McKibben, L., Ekstrand, C., Gould, L., Anton, K., & Borowsky, R. (2016). Disentangling genuine Stroop effects in reading from contingency effects: On the need for two neutral baselines. *Frontiers in Psychology: Cognitive Science, 7*(386). doi: 10.3389/fpsyg.2016.00386

Lorentz, E., Mickleborough, M., Mendez, I., Gould, L., Ekstrand, C., Ellchuk, T., & Borowsky, R. (2016). Using Presurgical fMRI with Emotional Stimuli to Aid Electrode Implantation in Deep Brain Stimulation for Treatment Resistant Depression. Under submission at *Neuromodulation.*

Louwerse, M. M., & Jeuniaux, P. (2010). The linguistic and embodied nature of conceptual processing. *Cognition, 114*(1), 96-104.

MacLean, P. D. (1990). *The triune brain in evolution: Role in paleocerebral functions*. Springer Science & Business Media.

MacLeod, C. M. (1991). Half a century of research on the Stroop effect: an integrative review. *Psychological bulletin, 109*(2), 163. http://dx.doi.org/10.1037/0033-2909.109.2.163

MacLeod, C. M. (1992). The Stroop task: The" gold standard" of attentional measures. *Journal of Experimental Psychology: General*, *121*(1), 12.

Mahon, B. Z., & Caramazza, A. (2008). A critical look at the embodied cognition hypothesis and a new proposal for grounding conceptual content. *Journal of physiology-Paris*, *102*(1), 59-70.

Martin, (2001). Functional neuroimaging of semantic memory. In R. Cabeza & A. Kingstone (Eds.), *Handbook of Functional Neuroimaging of Cognition,* 153-186. Cambridge, MA: MIT Press.

Masson, M. E. (2011). A tutorial on a practical Bayesian alternative to null-hypothesis significance testing. *Behavior research methods, 43*(3), 679-690. doi:10.3758/s13428-010-0049-5

Masson, M. E., Bub, D. N., & Breuer, A. T. (2011). Priming of reach and grasp actions by handled objects. *Journal of Experimental Psychology: Human Perception and Performance*, *37*(5), 1470.

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

McKenna, F. P., & Sharma, D. (1995). Intrusive cognitions: An investigation of the emotional Stroop task. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 21*(6), 1595. http://dx.doi.org/10.1037/0278-7393.21.6.1595

Meier, B. P., Moeller, S. K., Riemer-Peltz, M., & Robinson, M. D. (2012). Sweet taste preferences and experiences predict prosocial inferences, personalities, and behaviors. *Journal of Personality and Social Psychology*,*102*(1), 163.

Meier, B. P., Robinson, M. D., & Clore, G. L. (2004). Why good guys wear white automatic inferences about stimulus valence based on brightness. *Psychological science, 15*(2), 82-87. doi: 10.1111/j.0963-7214.2004.01502002.x

Morgan, G. A., Goodson, F. E., & Jones, T. (1975). Age differences in the associations between felt temperatures and colour choices. *The American journal of psychology,* 125-130. doi: 10.2307/1421671

Näätänen, R., Syssoeva, O., & Takegata, R. (2004). Automatic time perception in the human brain for intervals ranging from milliseconds to seconds. *Psychophysiology, 41*(4), 660-663.

Neely, J. H. (1977). Semantic priming and retrieval from lexical memory: Roles of inhibitionless spreading activation and limited-capacity attention. *Journal of experimental psychology: general*, *106*(3), 226.

Neely, J. H. (1991). Semantic priming effects in visual word recognition: A selective review of current findings and theories. *Basic processes in reading: Visual word recognition, 11*, 264-336.

Nelson, D. L., McEvoy, C. L., & Schreiber, T. A. (2004). The University of South Florida free association, rhyme, and word fragment norms. *Behavior Research Methods, Instruments, & Computers, 36*(3), 402-407. http://dx.doi.org /10.3758/BF03195588

Newhall, S. M. (1941). Warmth and coolness of colours. *The Psychological Record*.

Niemic, C. P., & Warren, K. (2002). Studies of emotion. A Theoretical and Empirical Review of Psychophysiological Studies of Emotion. *Journal of Undergraduate Research, Rochester, 1*(1), 15-19.

Osgood, C. E., Suci, G. J., & Tannenbaum, P. H. (1957). The measurement of meaning. Urbana, IL: University of Illinois Press.

Pashler, H. (1989). Dissociations and dependencies between speed and accuracy: Evidence for a two-component theory of divided attention in simple tasks. *Cognitive Psychology*, *21*(4), 469-514.

Piaget, J., & Inhelder, B. (1969). The psychology of the child. New York: Basic Books.

Pinel, P., Piazza, M., Le Bihan, D., & Dehaene, S. (2004). Distributed and overlapping cerebral representations of number, size, and luminance during comparative judgments. *Neuron, 41*(6), 983-993.

Plaut, D. C. (1999). Systematicity and specialization in semantics. In *Connectionist Models in Cognitive Neuroscience* (pp. 184-193). Springer London.

Pöppel, E. (1997). A hierarchical model of temporal perception. *Trends in cognitive sciences, 1*(2), 56-61.

Pouthas, V., George, N., Poline, J. B., Pfeuty, M., VandeMoorteele, P. F., Hugueville, L., ... & Renault, B. (2005). Neural network involved in time perception: an fMRI study comparing long and short interval estimation. *Human brain mapping, 25*(4), 433-441.

Proffitt, D. R. (2006). Embodied perception and the economy of action. *Perspectives on psychological science*, *1*(2), 110-122.

Pulvermüller, F. (1999). Words in the brain's language. *Behavioral and brain sciences*, *22*(02), 253-279.

Pylyshyn, Z. W. (1981). The imagery debate: Analogue media versus tacit knowledge. *Psychological review*, *88*(1), 16.

Pylyshyn, Z. W. (1984). *Computation and cognition*. Cambridge, MA: MIT press.

Quillian, M. R., (1968). Semantic memory. In M. Minsky (Ed.), Semantic Information Processing, MIT Press, Cambridge, MA, pp. 227–270.

Rabbitt, P. M. (1966). Errors and error correction in choice-response tasks.*Journal of experimental psychology*, *71*(2), 264.

Rao, S. M., Mayer, A. R., & Harrington, D. L. (2001). The evolution of brain activation during temporal processing. *Nature neuroscience, 4*(3), 317-323.

Richter, T., & Zwaan, R. A. (2009). Processing of colour words activates colour representations. *Cognition, 111*(3), 383-389. doi:10.1016/j.cognition.2009.02.011

Rips, L. J., Shoben, E. J., & Smith, E. E. (1973). Semantic distance and the verification of semantic relations. *Journal of verbal learning and verbal behavior*, *12*(1), 1-20.

Santiago, J., Ouellet, M., Román, A., & Valenzuela, J. (2012). Attentional factors in conceptual congruency. *Cognitive Science, 36*(6), 1051-1077. doi: 10.1111/j.1551-6709.2012.01240.x

Shearn, D., Bergman, E., Hill, K., Abel, A., & Hinds, L. (1990). Facial colouration and temperature responses in blushing. *Psychophysiology, 27*(6), 687-693. doi:10.1111/j.1469-8986.1990.tb03194.x

Stroop, J. R. (1935). Studies of interference in serial verbal reactions. *Journal of experimental psychology*, *18*(6), 643.

Sutton, T. M., & Altarriba, J. (2008). Emotion words in the mental lexicon: A new look at the emotional Stroop effect. *The Mental Lexicon, 3*(1), 29-46. http://dx.doi.org/10.1075/ml.3.1.04sut

Thompson-Schill, S. L., Kan, I. P., & Oliver, R. T. (2006). Functional neuroimaging of semantic memory. *Handbook of functional neuroimaging of cognition*, *2*, 149-190.

Tulving, E. (1972). Episodic and semantic memory 1. *Organization of Memory. London: Academic*, *381*(4).

Vicario, C. M. (2011). Perceiving numbers affects the subjective temporal midpoint. *Perception*, *40*(1), 23-29.

Westerbeek, J., & Mutsaers, K. (2008). Depression Narratives. *Literature and medicine, 27*(1), 25-55. http://dx.doi.org/10.1353/lm.0.0017

Williams, L. E., Huang, J. Y., & Bargh, J. A. (2009). The scaffolded mind: Higher mental processes are grounded in early experience of the physical world. *European Journal of Social Psychology, 39*(7), 1257-1267. doi:10.1002/ejsp.665

Xuan, B., Zhang, D., He, S., & Chen, X. (2007). Larger stimuli are judged to last longer. *Journal of vision*, *7*(10), 2-2.

Zhang, H., & Kornblum, S. (1998). The effects of stimulus–response mapping and irrelevant stimulus–response and stimulus–stimulus overlap in four-choice Stroop tasks with single-carrier stimuli. *Journal of Experimental Psychology: Human Perception and Performance*, *24*(1), 3.

Zwaan, R. A. (2004). The immersed experiencer: Toward an embodied theory of language comprehension. *Psychology of learning and motivation*, *44*, 35-62.

Zwaan, R. A., & Yaxley, R. H. (2003). Spatial iconicity affects semantic relatedness judgments. *Psychonomic bulletin & review*, *10*(4), 954-958.

Appendix A

Characteristics of Emotion Word Stimuli

Word	Emotion	Log Frequency HAL	Length	Emotional Association Rating (1-5 scale)	Word Categorising Congruency (ms)	Colour Categorising Congruency (ms)
alone	sad	10.84	5	3.27	39	39
crying	sad	8.836	6	3.18	16	-1
depressed	sad	8.341	9	4	-14.5	-10
despair	sad	8.263	7	3.7	40	30
distress	sad	7.617	8	3.18	147	-50
empty	sad	9.866	5	3.33	74	0
gloom	sad	8.053	5	2.94	24.5	5
grief	sad	8.27	5	3.61	52.5	-40
helpless	sad	7.885	8	3.39	21	1.5
hopeless	sad	7.588	8	3.42	22.5	2
hurting	sad	8.15	7	2.97	-49	-37
isolated	sad	8.862	8	3.42	51.5	-16
lonely	sad	8.491	6	3.18	50	-7.5
miserable	sad	8.08	9	3.82	-22.5	-33.5
missed	sad	10.205	6	2.76	123.5	34.5
mourn	sad	6.347	5	3.64	35	-3
regret	sad	8.515	6	3.24	0	-42
rejected	sad	9.086	8	3.24	-11	-44
sorrow	sad	7.321	6	3.36	55.5	5
sorry	sad	11.445	5	2.64	25	0.5
tragic	sad	7.932	6	3.7	51.5	-55
unhappy	sad	8.241	7	2.88	-27	-18.5
weep	sad	6.717	4	3.12	52	-62
withdrawn	sad	7.686	9	3.27	91	-75
	M	**8.44**	**6.58**	**3.30**	**35.3**	**-15.7**
angry	mad	9.506	5	3.76	53	130.5
despise	mad	7.256	7	4.03	79	42.5
detest	mad	6.159	6	3.85	186.5	-32
enraged	mad	6.319	7	4.3	25.5	68.5
frustrated	mad	8.472	10	3.03	-4.5	-6.5
fury	mad	8.259	4	4.3	5.5	-8
harsh	mad	8.499	5	3.42	-49.5	28.5
hate	mad	10.695	4	4.33	-7.5	26
hostile	mad	8.444	7	4.09	49	-5
irritated	mad	7.131	9	3.18	-21.5	52.5
livid	mad	5.204	5	4.64	126.5	45
loathing	mad	6.877	8	4.3	-1	89.5
malicious	mad	7.416	9	4.39	-16	-2.5
mean	mad	11.93	4	3.18	114	0
outraged	mad	7.12	8	4.42	8.5	24.5
rage	mad	8.747	4	4.45	18	8
resent	mad	7.546	6	3.27	236	51
scorn	mad	6.891	5	3.73	21	31
spite	mad	8.67	5	3.48	-56.5	-23
vicious	mad	8.176	7	4.58	98	-13
vile	mad	7.91	4	4.24	25	12
violent	mad	9.416	7	4.42	41.5	15.5
wicked	mad	8.345	6	4.06	-72.5	54.5
wrath	mad	8.638	5	4.33	-18	30
	M	**8.07**	**6.13**	**3.99**	**35.0**	**25.8**

Appendix B

Characteristics of Temperature Word Stimuli

Word	Temperature	Log Frequency HAL	Length	Temperature Strength Rating (1-5 scale)		Word Categorizing Congruency (ms)		Colour Categorizing Congruency (ms)	
				Exp 1	Exp 4	Exp1	Exp 4	Exp1	Exp 4
arctic	cold	7.452	6	4.71	4.46	24	-54	-1	-85
blizzard	cold	7.651	8	4.00	4.38	9.5	118	30.5	17.5
breeze	cold	7.659	6	2.42	3.38	19.5	-51	21.5	-16.5
brisk	cold	5.852	5	2.71	3.25	58	76.5	-17.5	-36
chill	cold	8.185	5	2.96	3.92	-8.5	79.5	38.5	21
cool	cold	10.811	4	3.00	3.71	140.5	-13	33	33
february	cold	11.018	8	3.00	3.25	70.5	37	-11.5	18.5
flurry	cold	6.581	6	3.38	3.63	42.5	219	51.5	13
freeze	cold	8.645	6	4.08	4.58	62	48.5	9	5
fridge	cold	7.377	6	3.17	3.71	72	24.5	-81	-31.5
frigid	cold	5.694	6	3.88	4.08	-34	53	0	-60
frosty	cold	5.855	6	3.42	4.17	44	-5	23	-49.5
frozen	cold	8.95	6	4.33	4.54	-21	23	12	-24.5
icy	cold	9.255	3	3.92	4.46	114.5	-68	-48	-8
january	cold	10.482	7	3.21	3.13	102	38.5	5	38
nippy	cold	4.431	5	3.00	3.04	-63.5	23	-36	8
numb	cold	7.148	4	3.00	2.79	28.5	-46	-38	-41
polar	cold	8.203	5	4.33	4.08	7	47	46.5	33
shiver	cold	6.544	6	3.13	4.17	-22	71	-32.5	-34.5
sleet	cold	4.942	5	3.21	3.08	-15	66	-41.5	-8
snowy	cold	6.631	5	3.42	3.96	85	24.5	-10.5	13
tundra	cold	8.549	6	3.67	3.54	-12.5	-39	23.5	6.5
windy	cold	7.07	5	2.50	2.96	-56	36.5	-65.5	-70.5
winter	cold	10.082	6	4.04	4.46	24	30	-54.5	-20.5
	M	**7.711**	**5.625**	**3.44**	**3.78**	**28.13**	**30.81**	**-5.98**	**-11.63**
blaze	hot	8.432	5	4.33	4.17	63.5	77	10	14.5
boiling	hot	8.032	7	4.63	4.58	35	11	-43	45.5
broil	hot	4.727	5	3.75	3.96	58	48.5	-49.5	51
fever	hot	8.336	5	3.46	4.13	67	22.5	66	88.5
flame	hot	10.453	5	4.58	4.58	-33	28	26	-12
grill	hot	7.31	5	3.58	4.08	113	65.5	27	50
inferno	hot	7.881	7	4.96	4.38	9	2	72	-1.5
july	hot	10.063	4	3.46	3.67	86	-19	17	7
lava	hot	7.922	4	4.83	4.83	112.5	33.5	31.5	9.5
magma	hot	6.17	5	4.75	4.75	8	109.5	83	19.5
molten	hot	6.458	6	4.83	4.13	13.5	.5	73	-9
piping	hot	6.988	6	4.29	3.33	36	55.5	-21.5	46
roast	hot	7.302	5	3.46	4.08	41	-32	50.5	53
scald	hot	4.745	5	4.46	3.96	54	-6	27.5	24
scorch	hot	5.242	6	4.63	4.17	49	15	41	35.5
sear	hot	5.58	4	3.96	3.54	40	39	-14	-72
simmer	hot	7.022	6	3.58	3.58	152.5	58.5	18.5	33
spicy	hot	6.742	5	3.96	3.96	84	30	43.5	28.5
steam	hot	8.562	5	3.92	4.08	-0.5	80	45.5	22.5
summer	hot	10.448	6	3.71	4.17	173	54	40	-15.5
sun	hot	11.213	3	4.46	4.67	58	55.5	-6.5	16.5
torch	hot	7.794	5	4.33	4.17	46	32	59	19.5
tropical	hot	8.977	8	4.17	4.08	26.5	96	-6	13.5
volcano	hot	8.024	7	4.88	4.42	-18.5	-44	59.5	57
	M	**7.684**	**5.375**	**4.21**	**4.14**	**53.06**	**33.85**	**27.08**	**21.81**

Note. HAL = Hyperspace Analogue to Language (see Balota et al. 2007); *M* = mean.

Appendix C
Characteristics of Temporal Word Stimuli

Word	Time Meaning	Log Frequency HAL	Length	Association Rating (1-5 scale)
abrupt	short	6.578	6	3.71
accelerated	short	8.122	11	3.96
brief	short	9.878	5	4.21
dashing	short	6.366	7	3.75
fast	short	11.252	4	4.54
flashing	short	7.903	8	3.75
fleeting	short	6.153	8	3.39
hasty	short	6.911	5	3.64
hurry	short	8.544	5	4.25
immediate	short	10.358	9	4.39
instant	short	9.165	7	4.43
quick	short	10.537	5	4.21
racing	short	9.621	6	4.07
rapid	short	9.119	5	4.46
running	short	11.886	7	3.79
rushed	short	7.757	6	4.39
short	short	11.467	5	4.14
snappy	short	6.787	6	4.07
speedy	short	7.302	6	4.46
sudden	short	9.075	6	4.18
swift	short	8.3	5	3.75
transient	short	7.202	9	2.54
vanishing	short	6.922	9	3.25
zooming	short	6.26	7	3.75
	M	**8.48**	**6.54**	**3.96**
constant	long	9.802	8	3.82
continual	long	7.431	9	4.04
delayed	long	8.513	7	4.14
enduring	long	7.217	8	3.68
eternal	long	9.292	7	4.36
extend	long	9.146	6	3.89
forever	long	9.938	7	4.57
freeze	long	8.645	6	3.54
gradual	long	7.442	7	3.86
infinite	long	9.37	8	4.39
lasting	long	8.131	7	4.14
lengthy	long	8.104	7	4.43
linger	long	6.368	6	3.57
pause	long	8.48	5	3.64
perpetual	long	7.558	9	2.89
persist	long	7.361	7	3.25
prolong	long	6.486	7	4.29
remain	long	10.11	6	3.61
slow	long	10.686	4	4.64
sluggish	long	6.752	8	3.75
stay	long	10.8	4	3.61
stuck	long	9.971	5	3.32
sustain	long	8.042	7	3.61
wait	long	11.031	4	3.54