THE DEVELOPMENT AND COMPREHENSION OF CONVENTIONAL METAPHORS

By

MICHAEL T. MCKAY

A DISSERTATION PRESENTED TO THE GRADUATE SCHOOL
OF THE UNIVERSITY OF FLORIDA IN PARTIAL FULFILLMENT
OF THE REQUIREMENTS FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY

UNIVERSITY OF FLORIDA

2004

ACKNOWLEDGMENTS

I would like to thank the members of my doctoral committee for their support and guidance during the qualifying process, and through the development and completion of this dissertation. These include Dr. Ira Fischler, committee chair; Dr. Lise Abrams and Dr. Peter Delaney of the Department of Psychology; and Dr. Tracy Linderholm of the Department of Educational Psychology. They have all contributed to making this accomplishment an enjoyable and enlightening process. I would also like to thank the members of my immediate family: my wife Cathy, for her years of love, encouragement, and support; my sons, Jerome and Sean, for being continuous sources of pleasure and inspiration to me, and my mother-in-law, Oleta, for her sustained assistance. I would also like to thank my Mother and Father for their years of patience, guidance, and complete support. I am indebted to all of them.

TABLE OF CONTENTS

page

ACKNOWLEDGMENTS ii

LIST OF TABLES v

LIST OF FIGURES vi

ABSTRACT vii

CHAPTER

1 INTRODUCTION 1

Metaphor in Language and Thought 1
Classifying Metaphor Models 2
Conventionalization 7
The Influence of Conventionalization 9
The Present Investigation 11

2 DEVELOPMENT OF STIMULUS MATERIALS 14

Survey 1 Method 16
Participants 16
Materials 16
Rating Dimensions 17
Design 20
Apparatus and Procedure 21
Results and Discussion 23
Reading Times 23
Ratings 24
Decision Times 28
Attributes 30
Survey 2 (Probe Norms) Method 31
Participants 32
Materials and Apparatus 32
Procedure 32
Results and Discussion 33

3 EXPERIMENT 1-CONVENTIONALITY, FORM AND ATTRIBUTE PRIMING ... 35

Method 41
Participants 41
Materials 41
Design 43
Procedure and Apparatus 43
Results and Discussion 45
Figurative Meaning Decision Times 46
Lexical Decision Times 51
Conventionality Ratings 55

4 EXPERIMENT 2-PRACTICE, CONTEXT, AND CONVENTIONALIZATION 56

Method 61
Participants 61
Materials 61
Design 61
Procedure and Apparatus 62
Results and Discussion 64
Figurative Meaning Decision Times 65
Lexical Decision Times 69
Conventionality Ratings 77

5 GENERAL DISCUSSION 82

APPENDIX

A 200 SENTENCES USED IN SURVEY 1 (SENTENCE NORMS) 91

B QUESTIONS ASKED PARTICIPANTS IN SURVEY 1 94

C CATEGORIZATION AND AVERAGE RATINGS OF 200 SENTENCES USED IN SURVEY 1 97

D 400 WORDS USED AS STIMULUS MATERIAL IN SURVEY 2 104

E QUESTIONS/INSTRUCTIONS ASKED OF PARTICIPANTS IN SURVEY 2 106

F 96 SENTENCES AND ATTRIBUTE PROBES USED IN EXPERIMENT 1 107

G CONVENTIONALITY RATINGS 110

LIST OF REFERENCES 113

BIOGRAPHICAL SKETCH 118

LIST OF TABLES

Table page

2-1. Average Reading Times (in seconds) for Sentence Types based on Participants' Literal and Figurative Ratings to All Sentences..... 24

2-2. Pearson Correlation Matrix of Pilot Conventionality and Sentence Ratings..... 25

2-3. Pearson Correlation Matrix of Sentence Ratings for Novel Figurative Sentences... 26

2-4. Pearson Correlation Matrix of Sentence Ratings for Conventional Figurative Sentences..... 27

2-5. Decision Times (in seconds) for Sentence Types based on Participants' Literal and Figurative Ratings to All Sentences..... 29

3-1. Percent of Sentences Judged Meaningful and Meaningless for Conventional and Novel, Metaphors and Similes 45

3-2. Analysis of Novel and Conventional Topic and Vehicle Terms via the MRC Psycholinguistic Database..... 50

4-1. Average Difference Scores for the Conventionality Ratings of the Conventional Metaphors used in Experiment 2. (Column Titles minus Row Titles)..... 78

4-2. Average Difference Scores for the Conventionality Ratings of the Novel Metaphors used in Experiment 2. (Column Titles minus Row Titles)..... 80

LIST OF FIGURES

Figure page

3-1. Meaningfulness Decision Times (in sec) for Meaningful Sentences based on Conventionality Type (Conventional, Novel) and Form (Simile, Metaphor)..... 47

3-2. Meaningfulness Decision Times in Terms of Speed-per-word (in msec) for Meaningful Sentences based on Conventionality Type (Conventional, Novel) and Sentence Form (Simile, Metaphor). 48

3-3. Lexical Decision Times (in ms) following Meaningful Sentences based on Sentence Conventionality (Novel, Conventional) and Probe Type (High Salient, Low Salient, Unrelated)..... 53

4-1. Median Meaningfulness Decision Times (in ms) for All Sentences in Experiment 2 (Repetition and Varied Topic Conditions) and Experiment 1..... 65

4-2. Comparison of Average Meaningfulness Decision Times (in ms) for Novel and Conventional Sentences Judged as Meaningful in Experiment 2 (Repetition Condition) and Experiment 1. 66

4-3. Comparison of Average Meaningfulness Decision Times (in ms) for Novel and Conventional Sentences Judged as Meaningful in Experiment 2 (Varied Topic Condition) and Experiment 1. 67

4-4. Median Lexical Decision Times (in ms) Following Meaningful Sentences by Probe Type for Condition 1 (Repetition). (A) Lexical Decision Times from Experiment 1, (B) Lexical Decision Times for the Not Studied Condition, (C) Lexical Decision Times for the Studied Condition. 71

4-5. Median Lexical Decision Times (in ms) Following Meaningful Sentences by Probe Type for Condition 2 (Varied Topic). (A) Lexical Decision Times from Experiment 1, (B) Lexical Decision Times for the Not Studied Condition, (C) Lexical Decision Times for the Studied Condition. 73

4-6. Median Lexical Decision Times (in ms) Following Meaningful Sentences by Probe Type for Condition 3 (Context). (A) Lexical Decision Times from Experiment 1, (B) Lexical Decision Times for the Not Studied Condition, (C) Lexical Decision Times for the Studied Condition. 75

Abstract of Dissertation Presented to the Graduate School
of the University of Florida in Partial Fulfillment of the
Requirements for the Degree of Doctor of Philosophy

THE DEVELOPMENT AND COMPREHENSION OF CONVENTIONAL METAPHORS

By

Michael T. McKay

May, 2004

Chair: Ira Fischler
Department: Psychology

The Career of Metaphor Hypothesis contends that different processes are used to comprehend novel and conventional metaphors. It further predicts that a shift in cognitive processing occurs with the conventionalization of a novel metaphor. Developing a method to reliably indicate this shift during the conventionalization of a novel metaphor would support this view and may provide the means to reconcile differences between the two best empirically supported views of metaphor comprehension: Comparison (Structure Mapping) and Categorization (Class Inclusion). Two Surveys and two Experiments were designed and conducted with this goal in mind.

Survey 1 was designed to identify the factors that contribute to the comprehension of a metaphor and explored the contribution of those factors in determining the conventionality of a metaphor. While no single factor was established as the most important, support was provided for those factors identified in the existing literature. More importantly a corpus of qualitatively and quantitatively balanced stimulus materials

were developed for use in Experiments 1 and 2, as well as future experiments. Survey 2 further refined probe words that would be used in lexical decision tasks in Experiments 1 and 2.

Experiment 1 explored the use of ratings, form preference (simile or metaphor), and lexical decision tasks to dissociate the processes involved in the comprehension of novel and conventional metaphors. Lexical decision response times to different types of probes proved to be an effective tool in dissociating the two. Experiment 2 used conventionality ratings and lexical decision response times to determine the effects of different kinds of learning experience on the conventionalization of novel metaphors. Three study conditions (Repetition, Varied Topic, and Context) were used with the Variable Topic condition having the greatest influence on the conventionalization of novel metaphors, which directly supported the hypotheses of the Career of Metaphor Hypothesis. It is speculated that the conventionalization of metaphors through varied topic presentation can be useful in creating the most effective metaphors. These findings have applications in a variety of domains including education, advertising, rhetoric, and marketing.

CHAPTER 1
INTRODUCTION

In this dissertation, a series of experiments are described that explore the factors that facilitate metaphor comprehension. Of particular interest are the notions that different processes are involved in the comprehension of novel and conventional metaphors, and that a shift in cognitive processing corresponds with the conventionalization of a novel metaphor.

Metaphor in Language and Thought

The word "metaphor" is derived from the Greek root words "meta" (meaning “to transfer”) and "pherein" (meaning “to bear”). A metaphor is a figure of speech in which a word or phrase literally denoting one kind of object or idea is used in place of another to suggest a likeness or analogy between them (e.g., "History is a mirror."). The subject of a metaphor is frequently referred to as the “target,” “tenor,” or “topic” of the metaphor (e.g., “History”), and that to which the subject is being compared is termed the “base” or “vehicle” (e.g., “mirror”) (Richards, 1936; Verbrugge & McCarrell, 1977). The similarities between the topic and the vehicle are termed the “ground,” and the dissimilarities are referred to as the “tension.” The simplest form of a metaphor is thus “A topic is a vehicle” (e.g., “Creativity is a toaster.”). Metaphors in this form are frequently referred to as “nominal metaphors,” in which the first noun given is the topic and the second noun given is the vehicle. Many of the existing models of metaphor comprehension concern themselves with the relationship between the topic and the

vehicle. In the second basic type of metaphor, the subject of the sentence is the topic, and the verb of the sentence is the vehicle (e.g., “The dog flew across the yard.”). Metaphors in this form are referred to as “predicative metaphors.” This scope of this proposal will be limited to nominal metaphors; however, it is commonly assumed that predicative metaphors operate by the same mechanisms as nominal metaphors (see Glucksberg, 2001, p. 49).

The strong claim by some (e.g., Gibbs, 1994a, 1994b; Lakoff & Johnson, 1980; Lakoff & Turner, 1989) is that metaphors may be the primary (if not the only) mechanism for reasoning with abstract concepts. A softer claim is that metaphors allow us to structure vague or ambiguous ideas in terms of more concrete realms of experience. In line with this softer claim, metaphors warrant study because they are highly pervasive in language, and are a major source of conceptual change and learning (Gentner & Bowdle, 2001). As such, the understanding of the underlying processes of metaphor comprehension is of interest to a variety of domains (e.g., science, education, journalism, mass communication, etc.).

Classifying Metaphor Models

A survey of the existing literature on metaphor processing and comprehension drawn from the domains of linguistics, philosophy, and psychology, revealed over twenty nominally distinct models. Many of these models appear to be minor variations of each other; still others are drastically different. The first step in synthesizing the large number of metaphor comprehension models in the literature is to organize them. One approach is to group them into two functionally related categories based on what appears to be their primary purpose and focus. The first

category can be referred to as “Sequencing” or “Timing” oriented models and the second category can be referred to as “Process” oriented models.

Sequencing or Timing models are concerned with the order of events in metaphor comprehension. These models take what might be termed a classic information-processing approach to metaphor comprehension, in that they are primarily concerned with the order in which literal and figurative senses of utterances are processed or activated, and less concerned with details of the how figurative language is understood. In this sense, they are the more global of the approaches. Much of this work has been framed by analogous controversies about the access of multiple meanings of polysemous words; indeed, for very familiar metaphors, the situation may be identical, as figurative meanings of words frequently used as metaphors become lexicalized.

There are several possible approaches for understanding a metaphor that may have a nonliteral (figurative) as well as a literal meaning. The first approach involves accessing the literal interpretation before the figurative, which is accessed only if a literal interpretation fails. This model, often called the "Standard Pragmatic Model" because of its claim that the literal meaning is derived via a linguistic process, and if that fails, pragmatic knowledge is used to search for a figurative sense that makes sense. A variation to this approach involves accessing both the literal and figurative interpretations, in that order, but with figurative analysis not conditional on the failure of a literal parse. In a second approach, derived largely from “modular” models of language processing, literal and figurative senses are analyzed in parallel from the outset, with one or the other becoming available first depending on a variety of

factors. A third approach, often referred to as the “Direct Access” model, involves using whatever context is available (which may include prior experience and associations with the specific metaphor) to determine which interpretation to first access. A fourth approach involves accessing a preliminary, neutral interpretation compatible with both the literal and figurative interpretations.

For the most part, empirical studies of the timing of literal and figurative senses has disconfirmed the Standard Pragmatic Model, and provided strong support for at least parallel access, and in some cases, direct (and contextually constrained) processes (for a detailed review, see McKay, 2003). For example, Glucksberg, Gildea, and Bookin (1982) found that a meaningful figurative sense slowed the decision that a sentence did not have a literally true sense. This metaphor interference effect could not occur if analysis of figurative senses began only when the literal analysis failed.

“Process” oriented models are concerned with the specific component processes that are involved with metaphor comprehension, rather than the sequence or order in which literal and figurative analyses occur. It is this approach that is the focus of the present dissertation.

Two general classes of Process models have emerged from the research, which have come to be termed the “Comparison” and the “Categorization” approaches, respectively. The terms refer to the core process by which metaphors are held to be understood. Ultimately, they seem to derive from the contrast between simile and metaphor forms for figurative languages, which have their literal

counterparts in comparison statements-“A bat is like a bird”-and categorization statements-“A robin is a bird.”

In “Comparison Models,” the topic and vehicle representations are understood at roughly the same level of abstraction. Metaphoric statements are thought to be interpreted as comparisons: “My lawyer is a shark” is understood as a simile-“My lawyer is like a shark.” Comprehension thus involves discovering the common features, or the common relational structure, between the topic and vehicle concepts. The comparison may involve a small set of attributes of the topic and vehicle, or entire schematic conceptual frameworks, which are aligned in some fashion as part of the comparison process (Gentner & Bowdle, 2001). This view is frequently referred to as “domain mapping,” wherein attributes or predicates from the domain of the vehicle term are mapped onto the domain of the topic term. Work in this area has been most directly motivated by work on similarity judgments (e.g., Tversky, 1977) and on analogical reasoning (e.g., Gentner & Markman, 1997).

Gentner and others have hypothesized that it is unfamiliar or novel metaphors that invoke a comparison process. In the comparison process, domain mapping occurs in a manner similar to that proposed in the “Structure Mapping Theory” (Falkenhainer, Forbus, & Gentner, 1989; Gentner, 1983) and the Career of Metaphor Hypothesis (Gentner & Bowdle, 2001). During the “alignment” stage attributes that do not “fit” are actively inhibited/suppressed. Those that remain sufficiently activated are mapped during the “projection” to an ad-hoc domain created in working memory. Comprehension occurs when this mapping is complete enough to allow the listener/reader to make sense of the topic-vehicle comparison. Depending on several

factors such as the time allowed to process a particular discourse, the amount of context provided, and the comprehender's background knowledge, a complete mapping of all the relationships between the domains may not occur. However, partial mappings of attributes and relationships between the domains can result in a working comprehension of the utterance.

In Categorization Models, the topic representation is subordinate to the vehicle representation in terms of abstractness. The vehicle term is thought to be polysemous, referring to both a domain-specific concept (e.g., "shark") and to a prototype exemplar, which is the epitome of a more abstract category (e.g., "vicious" and "calculating entities"). The process is one that is more akin to categorical analysis and property attribution and the processing of polysemous words, than it is to analogical reasoning. Metaphor comprehension is thus characterized as a search for properties (attributes or characteristics) of the vehicle concept that can plausibly be attributed to the topic. For example the metaphor "Our love has been a roller coaster ride" takes the properties of "a roller coaster ride" (e.g., "exciting," "scary," "fast," "full of ups and downs," etc.) and attributes them to "love."

Literal categorization statements have to be made between concepts existing at different levels of a taxonomic hierarchy, in the correct order, reflecting the subordinate-superordinate structure (e.g., "Tin is metal" not "Metal is Tin") (Glucksberg & Keysar, 1993). The non-reversibility of metaphors (cf., "My surgeon is a butcher" vs. "My butcher is a surgeon") has played an important role in theories of metaphor comprehension. Glucksberg (2001), for example, sees non-reversibility following very naturally from the categorization view. According to Glucksberg's

category approach, the interpretation of a metaphor is seen as an interaction between the category that is prototypically associated with the vehicle and the attributes or characteristics associated with the topic. Since topic and vehicle swap roles in a reversed metaphor, it naturally follows that the change in prototypical category results in a different interpretation.

There are however other explanations that can account for the asymmetry between the interpretations. Gentner et al. (2001) propose three alternative explanations: (a) *Initial temporal asymmetry*, which contends that processing begins by accessing or abstracting information from the vehicle and then projecting that information from the vehicle to the target; (b) *Initial processing asymmetry* (Glucksberg, McGlone, & Manfredi, 1997), which contends that processing begins simultaneously with both the topic and vehicle terms, but is differentiated initially in role-specific ways (i.e., processing of the vehicle term is different than processing of the topic term); and (c) *Initial symmetry followed by processing asymmetry*, which contends that that the initial stage involves a role neutral alignment that is followed by a directional process of inference projection. Based on numerous studies (Gentner & Wolff, 1997; Glucksberg, McGlone, & Manfredi, 1997; Wolff and Gentner, 1992; Wolff & Gentner, 2000) Gentner and colleagues have concluded that *initial symmetry followed by processing asymmetry* is the most likely alternative explanation.

Conventionalization

Gentner and colleagues have recently suggested that as a novel metaphor is seen or heard more frequently it becomes increasingly conventional, and there is a shift in the mode of processing from comparison to categorization (Gentner & Bowdle, 2001; Gentner & Wolff, 1997). This account (The Career of Metaphor

Hypothesis) is similar to a number of recent proposals that contend that the interpretation of novel metaphors involves “sense creation” and that the interpretation of conventional metaphors involves “sense retrieval” (e.g., Blank, 1988; Blasko & Connine, 1993; Giora, 1997; Martin, 1992; Turner & Katz, 1997). The notion that novel and conventional metaphors are processed differently, and that a shift in cognitive processing corresponds with the conventionalization of a novel metaphor is what motivated the experiments in this dissertation.

Before proceeding, it is important to distinguish the terms “familiarity” and “conventionality” (as defined by Wolff & Gentner, 2000). Familiar figures of speech are those where the topic and vehicle concepts have been encountered together many times in the past. An example of a familiar metaphor is “Life is a journey.” In contrast, conventional metaphors are those that contain stock vehicles-that is, vehicle terms that are closely associated with a stable metaphoric category. A stable metaphoric category is one that has well-established interpretations (e.g., the "pig" in "John is a pig" typically connotes “sloppiness” or being “dirty,” and any number of topics can be inserted without changing the interpretation). In a conventional metaphor the topic-vehicle pairing may be novel, even though the vehicle term is strongly associated with a stable metaphoric category. The distinction is important because a figure of speech can be familiar without being conventional (e.g., “Life is like a box of chocolates”) and conventional without being familiar (e.g., “Pawnshops are gold mines”) (Chiappe & Kennedy, 2001).

In discussing the distinction between conventionality and familiarity, Gentner and Bowdle (2001) claim that conventionalization arises from repeated figurative

uses of a given vehicle term, so that it acquires a specific metaphoric meaning. Familiarization involves repeated exposures to specific topic-vehicle pairings, as has occurred with the well-known metaphor “Time is money.” In essence, familiarization leads to the creation of stock expressions as opposed to stock vehicle terms.

Bortfeld and McGlone (2001) contend that a comprehensive theory of metaphor must necessarily be able to account for the fact that metaphors may be interpreted in fundamentally different ways depending on specific circumstances (such as context, time allowed for comprehension, etc.). They propose that the notion of different processing sets may be used to account for much of the variability in metaphor interpretation. Their claim is that qualitatively different interpretations may be the product of different metaphor processing sets. In this sense "processing sets" are similar to "mental sets" and result in the application of different processes to interpret conventional and novel metaphors. Following this logic, the attributional (category) and domain-mapping (comparison) models can be viewed not as competing comprehensive models or metaphor interpretations, but rather as descriptions of distinct processing sets that are activated in different interpretational contexts. Just as our interpretations of a given literal phrase structure or polysemous word can be dramatically influenced by context and processing sets, our interpretations of metaphorical language might be similarly influenced by context and goal changes.

The Influence of Conventionalization

Before proceeding, it is important to clearly define what is meant by “novel” and “conventional” metaphors. As previously mentioned, “conventional” metaphors are those that have been encountered enough times so that the vehicle term acquires a

specific and “stable” metaphoric meaning. “Novel” metaphors are those that have not been encountered enough times to form a specific or “stable” metaphoric meaning (e.g., “The wind is an armadillo”). We typically make few comparisons with “armadillos” and most of us do not have a stable metaphoric meaning for this vehicle.

The prediction made by the Career of Metaphor Hypothesis that novel metaphors are processed by comparison (sense creation) and that conventional metaphors are processed by categorization (sense retrieval) has important theoretical consequences. It means that the degree of conventionalization of a metaphor might be used to identify what processes are involved in its comprehension. In pursuit of this end Bowdle and Gentner (1995, 1999) (as reported in Gentner et. al., 2001) tested what Bowdle terms the “grammatical concordance assumption,” namely that the simile form (e.g., “A forest is like a harmonica”) invites a comparison and will be the preferred form for processing of novel figurative statements, while the metaphor form (e.g., “Divorce is a nightmare”) invites a search for a prestored category (which does not exist in novel metaphors) and will be the preferred form for processing of conventional figurative statements.

Bowdle and Gentner collected comprehension times for novel and conventional figurative phrases presented as metaphors or similes. As expected, conventional figuratives were interpreted faster than novel figuratives overall. These results are consistent with other results that show that conventional expressions are comprehended more rapidly than novel expressions (Blank, 1988; Blasko & Connine, 1993). More importantly, Gentner and Bowdle (2001) found the predicted interaction

between conventionality and grammatical form: Novel similes were faster than novel metaphors, but conventional metaphors were faster than conventional similes.

The Present Investigation

Two surveys were conducted to develop stimulus materials and two experiments were conducted to explore how novel and conventional metaphors are comprehended, and to directly study the conventionalization process itself. Survey 1 was designed to identify the factors that contribute to the comprehension of metaphors, and how those factors influenced ratings of the conventionality of a metaphor. A second goal of Survey 1 was the development of a corpus of stimulus sentences that would allow counterbalancing on the factors identified in previous metaphor comprehension research as being important for understanding (e.g., Comprehensibility, Literalness, Figurativeness, Familiarity, Interpretability, Aptness, and Similarity). The lack of control over these dimensions is a source of concern with the findings of earlier metaphor research. A third goal was the identification of attributes associated with the figurative interpretations of the stimulus sentences. These attributes were important for the development of probe words that could be used in lexical decision tasks in Experiments 1 and 2. Survey 2 was conducted to further refine the probe words developed in Survey 1.

An issue that is central to demonstrating that metaphors are conventionalized is how this conventionalization process can be experimentally determined. Experiment 1 examined several measures that could possibly distinguish novel and conventional metaphors. These measures included direct subjective ratings on the seven dimensions defined by the literature as being important to metaphor comprehension, and the relative speed of comprehension of simile and metaphor

forms. This was done in part to attempt a replication of the findings of Bowdle and Gentner (1995, 1999) (described above), with the more carefully controlled materials developed in Survey 1.

Experiment 1 also examined the feasibility of using differential lexical priming as an indicant of conventionality. Numerous metaphor comprehension studies have used priming manipulations (Blasko & Connine, 1993; Galinski & Glucksberg, 2000; Gernsbacher et al., 2001; Gibbs et al., 1997; Glucksberg et al., 2001, Kemper, 1989). It was predicted that the scope of priming for comprehended novel metaphors should include a broader range of attributes for the vehicle (and possibly the topic) as a result of the assumed alignment and mapping process involved. This prediction was based on Gentner's notion that there should be a broader and deeper semantic activation of concepts associated with novel metaphors compared to the less broad and more shallow activations associated with a conventional metaphor. The specific pattern of priming hypothesized and found for highly salient and less salient attributes of the vehicle terms is discussed in Chapter 3.

Since conventionality is a theoretically important factor in metaphor comprehension, it would be useful to find a reliable method to induce conventionalization experimentally. Experiment 2 used conventionality ratings and lexical priming to examine the conventionalization of novel metaphors. Participants were given practice in comprehending metaphors with specific vehicles in a study phase. Three conditions of practice were compared which were hypothesized to produce different degrees of conventionalization. In one case (repetition encoding), the same metaphor was presented three times for comprehension decisions. In a

second case (variable topic encoding), three presentations of the same vehicle each paired with different topics were presented for comprehension decisions. As with other forms of concept learning and perceptual discrimination (see, e.g., Dukes & Bevan, 1967) the varied encoding was predicted to result in greater "tuning" of the metaphoric conceptual sense of the vehicle, facilitating the process of conventionalization and resulting in higher vehicle ratings of conventionality.

In the third case (context) condition, a vehicle was presented in a single metaphoric frame, but a preceding context was presented to help "make sense" of the metaphor. Context has been used to facilitate metaphor comprehension (e.g., Gerrig & Healy, 1983, who found that the on-line comprehension of metaphors is facilitated by prior discourse context; and Pynte et al., 1996, who obtained EEG results that supported a context-dependent account of metaphor comprehension). The results of Experiment 2 are discussed in Chapter 4.

The implications of being able to identify two separable processes in the comprehension of novel and conventional metaphors, and potential applications of the conventionalization process are discussed in Chapter 5.

CHAPTER 2
DEVELOPMENT OF STIMULUS MATERIALS

One problem in the metaphor comprehension literature is the high degree of variability in the nature of the stimulus materials that have been used. Katz et al. (1988) attempted to rectify this by providing a corpus of 204 literary (selected from poetry) and 260 nonliterary (experimenter generated) metaphors on ten psychological dimensions that represented ratings of comprehensibility, some perceived metaphoric qualities, imagery values, familiarity, and tenor-vehicle relatedness. Although experimental counter-balancing can be conducted using these ten dimensions, the metaphors vary greatly in their syntactic and semantic complexity (e.g., "Humiliation is a curtain" and "Tranquility is a woodland river winding through hills in solitude"). This may be one of the reasons why the stimulus material in the corpus has not been used more frequently in studies of metaphor comprehension.

With this in mind, a new corpus of 200 stimulus sentences was compiled for use in Survey 1. Ratings were solicited from participants on several dimensions drawn from the numerous studies of figurative language that have used quantitative ratings or judgments about words, phrases, and sentences. These included phrase level ratings of (a) Aptness, (b) Comprehensibility, (c) (Familiarity, (d) Similarity, (e) Figurative Interpretability, (f) Figurativeness, and (e) Literalness, and word level ratings of Concreteness/Abstractness. The intention was to select materials that would create a gradient on each of the measured dimensions. The goal was to establish materials that would allow precise manipulations of these variables in future studies.

In addition to rating sentences on these dimensions, participants were also asked to provide rank-ordered attributes of the words. Specifically participants were asked to list attributes of the vehicle that informed their interpretation of the topic (e.g., when considering the sentence "The performer is an elephant" what qualities or attributes of "elephant" allow you to understand the figurative meaning of the sentence?). These attributes were constrained thus by the context of the stimulus sentence and were relational (relating the vehicle to the topic).

All sentences were in the nominal form (e.g., "A topic is a vehicle") and none contained adjectives. These sentences afford precise manipulation of the variables that have been implicated as being important in prior research on the comprehension of metaphors. Admittedly, one shortcoming of this stimulus material is that these sentences may not be representative of the processing that occurs with more complex metaphors. However, they do allow the methodical exploration of the processes involved in their comprehension, without confounding the results with additional semantic and syntactic variables. After reliable results have been obtained with these materials they can be modified in order to test expanded (longer metaphors with adjectives, e.g., "Creativity is a shiny toaster.") and embedded metaphors (metaphors contained in the context of a sentence, e.g., "When talking about creativity my art teacher always preached that creativity is a toaster and that ideas 'popped out' when they were ready.").

Independent ratings of novelty and conventionality for these metaphors were collected from the experimenter and four dissertation committee members to provide a preliminary measure of conventionality, assess its reliability, and explore its relation to

the variables previously mentioned. The rating dimensions are described in more detail in the next section.

Survey 1 Method

Participants

Two hundred participants (129 female) whose first language was English were recruited from undergraduate students taking General Psychology or Cognitive Psychology classes at the University of Florida. Participants were 18 years of age and older (*M* = 19.7 years) and on average had completed one year of college (*M* = 13.4 years). General Psychology students received two credits toward a class research requirement for their participation. Cognitive Psychology students received extra credit toward their class grade, as determined by their instructor.

Materials

The sentences used in this study were selected and modified from a database of over 2100 metaphors that the experimenter had compiled from the stimulus materials used in previously reported metaphor studies. An attempt was made to include metaphors that ranged widely in conventionality, from very novel to very conventional (see above). Each sentence was rewritten in its simplest form, (e.g., "A topic is a vehicle”-known as the nominal or minimal form). Four preliminary categories of 50 sentences each were created. One category-Literal Only (LO)-contained what the experimenter believed were sentences with only literal and no figurative meanings (e.g., “A beet is red.”). A second category-Figurative Only (FO)-consisted of sentences the experimenter believed were only figurative meanings (e.g., “A billboard is a wart.”). The third category-Literal and Figurative (LF)-contained what the experimenter believed to be both literal and figurative interpretations (e.g., “My brother is blue.”). The last

category-Nonsense (N)-was comprised of sentences for which the experimenter had no meaningful figurative of literal meaning. The sentences were selected so that as a corpus they would contain a range of values for each of the judged dimensions. The stimulus sentences and the questions used to solicit ratings are contained as Appendices A and B.

Rating Dimensions

Sentence ratings were obtained by asking participants to rate sentences on a scale of 1 to 5 in response to specific questions.

Aptness. Sometimes referred to as “metaphor goodness” or “sensibility”, aptness has been frequently employed as a measure of how good or apt metaphors appear (Tourangeau & Sternberg, 1981, 1982) or how pleasing the comparisons made by the metaphors were (Katz et al., 1988). Chiappe and Kennedy (2001) defined aptness as the extent to which a comparison captures important features of the topic. There are several variables that have been proposed to affect the perceived aptness of a metaphor. These include the salience and number of attributes or clusters of attributes that are being mapped, the amount of time the listener/reader has to perform the categorization or classification process, the degree of similarity between the topic and the vehicle, and the amount of context provided to shape the interpretation (particularly in the case of novel metaphors).

Aptness has been manipulated in many studies related to metaphor comprehension (Blasko & Connine, 1993; Chiappe & Kennedy, 2001; Gernsbacher et al, 1995; Gernsbacher, Keysar, Robertson, & Werner, 2001; Gerrig & Healy, 1983; Katz et al., 1988; Tourangeau & Rips, 1991; Tourangeau & Sternberg, 1981, 1982). In a typical study, Gildea and Glucksberg (1983) found that, with no supportive context, interference from figurative senses on judgments of literal meaning only occurred when the metaphors

had been rated as highly apt. Aptness may be the best “bottom-line” indicator of whether or not a metaphor has stimulated a rich comparison or mapping process for novel metaphors.

Familiarity. Familiarity is based on the reader's sense of familiarity with the ideas expressed by the metaphor. Several studies have involved manipulations of familiarity (Blank, 1988; Blasko & Briihl, 1997; Blasko & Connine, 1993; Chiappe & Kennedy, 2001; Cronk et al., 1993; Katz et al., 1988; Pynte et al., 1996). The familiarity of metaphors and related attributes (see below) have been a particularly important construct in recent models of metaphor processing, since it is argued that there may be qualitatively different modes of comprehension invoked for familiar and unfamiliar metaphors.

Chiappe and Kennedy (2001), for example, argue that when people are unfamiliar with a topic-vehicle pair, they have to derive the relevant properties that make up the ground of the statement. As they encounter a topic-vehicle pairing several times, the relevant category (the set of relevant properties from the ground of the statement) may be represented directly in long-term memory. This would enable the specific topic-vehicle pairing to be understood via a categorization process. Because of the potential correlation between familiarity and conventionality (see above), it may be important to identify the degree of familiarity of given stimulus sentences so that it can be controlled.

Similarity. May be obtained directly by having participants rate the similarity between the meanings of the topic and vehicle terms within a metaphorical expression. Not surprisingly, the notion of similarity between topic and vehicle has been fundamental to some of the most influential models of metaphor comprehension (e.g., Tversky, 1977;

Ortony, 1979; see below). Numerous studies have investigated the effects of similarity on metaphor comprehension (Connine, Blasko, & Wang, 1994; Glucksberg & Keysar, 1990; Goldvarg & Glucksberg, 1998, Medin, Goldstone, & Gentner, 1990, Ortony, 1979; Rips, 1975). Chiappe and Kennedy (2001) argue that as familiarity increases, there may be an increase in judged similarity between the topic and vehicle, the result of an increase in the number of relevant common features. The notion is that as we become more familiar with a comparison, the salient differences between the two concepts may become less prominent and the relevant common features may become more prominent. Chiappe and Kennedy (2001) found experimental evidence for this notion, but did not assess the conventionality of the vehicle terms. Because similarity can exert an influence on conventionalization similar to that exerted by familiarity it was important to collect similarity ratings in the corpus so that similarity could be accounted for in the conventionalization process.

Metaphoricity. Metaphoricity ratings measure the perceived metaphoric content of sentences. Typically they involve the judgment of literalness and figurativeness on a single dimension. Ortony (1979) found that people could make reliable judgments about degrees of metaphoricity, suggesting that there is a continuum from the literal to the nonliteral. Since sentences can have both literal and figurative senses (e.g., “My father is a preacher”), it would seem more precise to consider literalness and figurativeness as two distinct dimensions, which is the approach taken in this dissertation. Gentner and Bowdle (2001) contend that metaphoricity arises from the tension generated by juxtaposing concepts from unrelated but potentially relatable domains. The greater the distance between the topic and the vehicle, the more metaphoric the statement will seem

(so long as the concepts are relatable). Several metaphor comprehension studies have involved metaphoricity (Cacciari & Tabossi, 1988; Coulson & Van Petten, 2002; Gentner & Bowdle, 2001; Giora & Fein, 1999; Glucksberg et al., 2001; Katz et al., 1988). While all of the variables discussed have been implicated as being important in various studies of metaphor comprehension, the author feels that the most important variable in the "wisdom" of a metaphor is aptness. Therefore the correlations between variables that are predicted to be of the greatest interest are those between similarity, familiarity and aptness. Familiarity is predicted to correlate highly with aptness; the more familiar a metaphor becomes the more appropriate or "good" it will seem. Similarity is predicted to have a non-linear relationship with aptness. If the similarity between the topic and vehicle terms is too low, then the "tension" cited by Ortony (1979) as necessary in a metaphor will not exist. If the similarity between the topic and vehicle terms is too great then they are likely members of the same category and the sentence will be interpreted as a literal sentence instead of a metaphor.

Design

The design can be thought of as a 4 (sentence type: Literal Only, Figurative Only, Literal and Figurative, and Nonsense) x 2 (conventionality type: Novel and Conventional) within subjects design. In the case of Sentence Type, sentence assignment was based on the Literal and Figurative ratings averaged across the 20 participants who saw each sentence. Literal Only (LO) sentences were those that averaged greater than 3 on the "literal" scale and less than 3.01 on the "figurative scale." Figurative Only (FO) sentences were those that averaged less than 3.01 on the literal scale and greater than 3 on the figurative scale. Sentences thought to have Literal and Figurative meanings (LF) were those that averaged greater than 3.0 on both the literal and figurative scales.

Finally, Nonsense sentences (N) were those that received average ratings of 3.0 or less on both scales. Conventionality Type was determined from the independent pilot ratings.

Apparatus and Procedure

The experiment was written in Visual Basic and was presented using a computer monitor, keyboard, and mouse. At the beginning of the experiment participants were asked to provide the following demographic data: gender, age, highest level of education completed, and current educational goal. Names were not associated with subject numbers and the participant's privacy was protected.

To familiarize them with the task, participants underwent a two-sentence tutorial (one sentence with a literal only meaning "An apple is a fruit" and a sentence with both literal and figurative meanings "The performer is an elephant"). The tutorial provided feedback for each of the possible ratings as well as examples of meaning interpretations and attributes. The tutorial took approximately 15 minutes to complete.

Participants were then be presented with 20 individual sentences that they were asked to rate using 5-point Likert scales and responding to questions/instructions on: (a) comprehensibility (“A sentence can have many meanings. Rate how easy or difficult it is for you to understand the first meaning you thought of when you read the sentence above.”); (b) literalness (“Does this sentence have a literal meaning? Rate the sentence on how literal you interpret its meaning to be.”); (c) figurativeness (Does this sentence have a non-literal or figurative meaning? Rate the sentence on how figurative you interpret its meaning to be.”); and (d) familiarity (“How commonly or frequently have you read or heard the comparison being made between the first noun and the last word in the sentence?”).

When participants indicated that a sentence had some degree of figurative meaning (2 or greater on the Likert rating) they were asked to make additional ratings on that sentence using 5-point Likert scales on: (a) figurative interpretability (”Rate how easily you 'get the point' of the figurative meaning in the sentence above.”); (b) aptness (“How apt is this sentence as a metaphor? In other words rate how good a metaphor this sentence makes?”); and (c) similarity (“How similar in meaning are the first noun and the last word in the sentence?”).

Additionally, for each sentence they identified as figurative they were asked to comply with the following instructions:

> In your own words type up to three distinct meaning(s) of this sentence. Indicate in the box following the sentence if your interpretation is a literal or a figurative one. Remember to type the first meaning that comes to mind in the 'white' box, the second meaning that comes to mind in the 'yellow' box, and the third meaning in the 'blue' box. If the sentence is meaningless, type 'meaningless'.

Participants were then asked to identify which of these meanings were literal interpretations and which were figurative interpretations.

For the sentences they typed and identified as having figurative interpretations, participants were asked to comply with the following instructions:

> For your FIGURATIVE interpretations of the sentence, please type AS MANY attributes or qualities (NOT ASSOCIATIONS) of the last word in the sentence that you can. Try to list those attributes of the last word in the sentence that you applied to the first noun in the sentence to make it meaningful. Following each attribute is a number box - use this to assign a ranking of: '1' to the most important, '2' to the next most important, etc.

Rating judgments and the times to make those decisions were recorded, as were their generated meaning and attribute inputs. At the end of the session, the rationale behind the experimental design was explained during debriefing.

In a separate effort, conventionality/novelty ratings of the vehicle terms were solicited from the author and his dissertation committee members (henceforth referred to as “the conventionality pilot ratings”). Participants evaluated vehicle term for their degree of conventionality on a 5-point Likert scale. A rating of “1” was awarded for very novel vehicles and a rating of “5” was awarded for very conventional vehicles. Conventionality was determined by having the participants think about the uses of the vehicle term in various metaphors they knew and then making a determination as to whether the term had acquired domain-general meanings that make it a “stock” or “stable” vehicle term (as suggested by Gentner and Bowdle, 2001).

Results and Discussion

Averaged results are presented for the dependent measures, conventionality, and the number of attributes generated for each sentence categorized as figurative in Appendix C.

In all analyses, decision times occurring outside 3 standard deviations of the distribution of responses for the same type were rejected as outliers. Analyses of variance (ANOVA’s) were performed on all data using an alpha level of .05, making a sphericity-correction with the Huynh-Feldt (H-F) statistic. However, uncorrected degrees of freedom (df) will be presented throughout this dissertation for clarity.

Reading Times

Table 2-1 presents the means and standard deviations of the reading times for the four sentence types (Figurative Only, Literal and Figurative, Literal Only, and Nonsense). On average participants read Literal Only sentences faster, and Nonsense sentences slower than all others. Sentences that had both Literal and Figurative meanings were read slightly faster than those that had Figurative Only meanings.

Table 2-1. Average Reading Times (in seconds) for Sentence Types based on Participants' Literal and Figurative Ratings to All Sentences.

	Mean	*SD*	*n*
Literal Only	2.32	1.27	1278
Literal & Figurative	2.75	1.47	615
Figurative Only	2.87	1.57	1239
Nonsense	3.04	1.47	646

Reading times overall were quite varied, with 10.8% of the scores (462) exceeding the criteria of 3 *SD* above the mean. The uneven distribution of the remaining responses reflects the fact that many of the sentences selected by the experimenter as being Nonsense sentences were interpreted as figuratively sensible by participants. Additionally, sentences selected by the experimenter as having both Literal and Figurative meanings were frequently interpreted as having one or the other meaning.

A univariate ANOVA treating sentence as the random effect was conducted on reading times averaged by sentence across participants. There was a significant overall main effect of sentence type on reading time, $F(3,200) = 15.4$, $p < .001$. Planned pairwise comparisons using Dunnett's T3 for unequal variances indicated that Literal Only sentences were read significantly faster than all other sentence types and that the reading times for the other three sentence types were not significantly different from each other. The lack of difference in reading times between Figurative Only and Literal and Figurative sentences is one justification for combining these sentence types for study in Experiments 1 and 2.

Ratings

The primary purpose of Survey 1 was to determine the characteristics of a set of metaphors for use in subsequent studies. The relationship between the seven rated

dimensions and how these ratings in turn were related to conventionality was also of interest. To initially assess the relationship between these variables a Pearson Correlation Matrix was constructed (Table 2-2).

Table 2-2. Pearson Correlation Matrix of Pilot Conventionality and Sentence Ratings.

	COMP	LIT	FIG	FIGINT	FAM	SIM	APT
CONV	.42*	.09	.41*	.62*	.31*	.22*	.63*
COMP		.78*	-.28*	.75*	.93*	.55*	.59*
LIT			-.69*	.36*	.81*	.52*	.13+
FIG				.23*	-.44*	-.27*	.40*
FIGINT					.63*	.45*	.88*
FAM						.59*	.49*
SIM							.41*

Notes: * Correlation is significant at the 0.01 level.
+ Correlation is significant at the 0.05 level. CONV = conventionality. COMP = comprehensibility. LIT = literalness. FIG = figurativeness. FIGINT = figurative interpretability. FAM = familiarity. SIM = similarity. APT = aptness.

With the exception of Literalness and Pilot Conventionality, all of the rating dimensions were significantly correlated with each other. The strongest correlation was between Comprehensibility and Familiarity (r = .93), the weakest between Literalness and Pilot Conventionality (r = .09). Clearly, the presence and strength of these correlations makes it difficult to argue for the independence of some of the variables, at least, in determining comprehensibility. Some correlations appear to be driven by a common underlying decision, as with the strong positive correlations between aptness, figurative interpretability, and comprehensibility ratings (see Table 2-2).

In contrast to our claims about the independence of figurative and literal meaning, Literalness and Figurativeness were negatively correlated with each other (r = - .69). However, this may reflect the inclusion of many sentences intended to have only literal or figurative sense, rather than evidence for a single underlying literal-figurative dimension.

It may also help explain why the correlation between literal meaning and comprehensibility is strongly positive (+ .78), while that between figurative meaning and comprehensibility is weakly negative (- .28). The correlation between conventionality and familiarity was only moderate (+ .31), supporting the claims that these dimensions are relatively independent.

It is claimed that conventional and novel metaphors are understood through different processes. One way to address this is to see if the pattern of correlations obtained among the other variables is different for novel and conventional metaphors.

Therefore, two additional correlation matrices were created. The first looked at correlations between the seven dimensions of metaphor comprehension for all figurative sentences (Figurative Only and Literal and Figurative) containing vehicles that were identified as Novel by the Pilot Conventionality ratings (Table 2-3). The second looked at correlations between the seven dimensions of metaphor comprehension for all figurative sentences (Figurative Only and Literal and Figurative) containing vehicles that were identified as Conventional by the Pilot Conventionality ratings (Table 2-4).

Table 2-3. Pearson Correlation Matrix of Sentence Ratings for Novel Figurative Sentences.

	LIT	FIG	FIGINT	FAM	SIM	APT
COMP	.44*	.28*	.54*	.59*	.22*	.45*
LIT		-.10*	.21*	.40*	.27*	.23*
FIG			.11*	.15*	-.09^{+}	.08^{+}
FIGINT				.43*	.35*	.71*
FAM					.40*	.44*
SIM						.41*

Notes: n = 856
* Correlation is significant at the 0.01 level.
$^{+}$ Correlation is significant at the 0.05 level.

Table 2-4. Pearson Correlation Matrix of Sentence Ratings for Conventional Figurative Sentences.

	LIT	FIG	FIGINT	FAM	SIM	APT
COMP	.40*	-.06+	.49*	.60*	.18*	.34*
LIT		-.51*	.03	.38*	.21*	-.08*
FIG			.33*	-.18*	-.02	.33*
FIGINT				.35*	.18*	.66*
FAM					.24*	.28*
SIM						.14*

Note: n = 1782
* Correlation is significant at the 0.01 level.
+ Correlation is significant at the 0.05 level.

Comparing the two tables, most of the patterns of correlations are very similar, but there are a few notable exceptions. For Novel metaphors, Aptness is more strongly correlated with Familiarity and Similarity (r = .44 and .41 respectively) than are Conventional Metaphors (r = .28 and .14 respectively). Likewise, Similarity is more strongly correlated with Figurative Interpretability and Familiarity for Novel metaphors (r = .35 and .40 respectively) than for Conventional Metaphors (r = .18 and .24 respectively). This can lead us to infer that Similarity between the topic and vehicle terms is more important in Novel metaphors. It is also interesting to note the correlation between Familiarity and Figurativeness for Novel metaphors is weak, but opposite that for Conventional Metaphors (r = .15 compared to r = -.18).

In order to more precisely identify the dimensions that control variance in Aptness two regression series were built using Aptness as the response variable and the other five dimensions as predictor variables. Because we were interested in differences between the role these dimensions play in the processing of Novel and Conventional metaphors, a separate regression series was created for each set of sentences (as classified by the Pilot Conventionality ratings. Systematically each variable was dropped from the model and then added back in. This yielded the r-square change and provided a measure of the

variance in Aptness controlled by each of the predictor variables. The full model accounted for nearly 54% of the variance in Aptness (R^2 adjusted = .544) for Novel metaphors and 47% (R^2 adjusted = .468) for Conventional metaphors. Figurative Interpretability specifically accounted for 23% (R^2 change = .230) of the variance in Aptness for Novel metaphors and 19.5% for Conventional metaphors. Similarity accounted for almost 2% of the variance in aptness for Novel metaphors and almost 0% for Conventional metaphors. The other variables, Comprehensibility, Literalness, Figurativeness, and Familiarity accounted for less than 1% of the variance in aptness for both Novel and Conventional metaphors.

Both the correlation and regression analyses lead to the conclusion that a metaphor's Aptness is most affected by how well a reader "gets the point" of a metaphor (Figurative Interpretability). Similarity appears to exert a modest influence and accounts for a small portion of the variability in Aptness ratings. Comprehensibility, Literalness, Figurativeness, Familiarity, and Similarity have a minimal ability to explain the variance in Aptness ratings.

Decision Times

Table 2-5 reflects the decision times obtained for each of the rating responses. Decision time captures the amount of time involved in rereading the question and assigning a rating to the sentence.

Table 2-5. Decision Times (in seconds) for Sentence Types based on Participants' Literal and Figurative Ratings to All Sentences.

	Literal Only	Literal & Figurative	Figurative Only	Nonsense
Comprehensibility	4.15 (2.22)	4.80 (2.39)	5.09 (2.36)	4.80 (2.32)
	1266	*603*	*1170*	*644*
Literalness	3.78 (2.00)	4.56 (2.30)	5.12 (2.17)	4.16 (2.07)
	1281	*597*	*1188*	*640*
Figurativeness	3.80 (2.10)	3.87 (2.23)	3.76 (1.96)	3.69 (2.32)
	1216	*590*	*1247*	*636*
Figurative	4.59 (1.83)	4.05 (1.76)	4.07 (1.66)	4.01 (1.67)
Interpretability	*324*	*611*	*1245*	*428*
Aptness	5.03 (2.12)	4.35 (2.11)	4.31 (1.98)	4.02 (1.86)
	336	*611*	*1252*	*425*
Familiarity	5.89 (2.25)	5.62 (2.20)	5.43 (2.20)	4.96 (2.16)
	336	*591*	*1234*	*427*
Similarity	21.16 (32.74)	85.18 (33.82)	76.33 (36.10)	30.71 (35.96)
	1305	*559*	*1143*	*677*

Notes: *SD*'s are presented in parenthesis,
n's are presented in italics

Comprehensibility. Average decision times about the Comprehensibility of Literal Only sentences were fastest, compared to the other three sentence types. The slower decisions for the LF sentences, in particular, relative to the LO sentences is evidence for at least parallel processing of figurative and literal meanings.

Literalness. The impact of figurative meaning on literal reading is seen even more clearly in the slower times for this decision for Figurative Only compared to Nonsense (5.12 vs. 4.16 *s*), where the decision is negative, and the slower times for LF compared to LO (4.56 vs. 3.78), where the decision is positive. The slowing of decision about the absence of literal meaning by figurative meaning replicates earlier demonstrations of the "metaphor interference effect" (see Keysar, 1989, and Chapter 1), and is evidence against the Standard Pragmatic model described earlier.

Figurativeness. In contrast to decisions about literal meaning, decisions about figurative meaning are very similar for the different sentence types. This could be taken as evidence that in general, there are qualitative differences in how we understand literal and figurative meanings of statements.

Figurative interpretability. Decision times were somewhat slower to LO sentences (4.59 *s*) than to Nonsense sentences (4.02 *s*), indicating that literal meaning interfered with deciding that a sentence had no figurative sense, consistent with the parallel-access view. In contrast, decision times for FO and LF were about the same.

Aptness. Average decision times about the aptness of a sentence, as with figurative interpretability, were slowest for Literal Only sentences and fastest for Nonsense, with decisions for Figurative Only and sentences with both Literal and Figurative meanings intermediate and nearly the same.

Familiarity. Average decision times concerning the familiarity of a sentence was fastest for Nonsense sentences and slowest for Literal Only sentences, with times for FO and LF sentences again comparable and intermediate.

Similarity. Average Similarity decisions made to sentences with some type of Figurative meaning took two to three times longer than decisions to either Literal Only or Nonsense sentences. One explanation might be that the comparison triggered a more extensive mapping of sentences with Figurative meanings along the lines of Gentner's Structure Mapping Theory (Gentner and Bowdle, 2001).

Attributes

The attributes listed by participants in this study were intended for use as probes in lexical decision tasks in Experiment 1 and 2. The acquisition of salience ratings allowed them to be classified based on those ratings and frequency of occurrence as High

Salient Probes (those rated as most important in aiding readers to understand the figurative meaning of the sentence and occurring at greater frequency) and Low Salient Probes (those that were rated by readers as less important to their understanding the figurative meaning of the sentence and occurring with very low frequency). When reviewing the attributes, it became obvious that not all participants provided attributes that were exclusively related to the figurative meaning of the sentence. For example if the sentence was "The performer is an elephant," participants might provide you with the word "grey," which is related to a literal quality of an elephant. In order to identify and reject attributes that were related to the literal meanings of the topic and vehicle terms, or were lexical associates of those terms, a second Survey was conducted.

Survey 2 (Probe Norms) Method

The purpose of this study was to collect ratings, decision times, and attributes for words presented in isolation. Judgments of concreteness/abstractness and familiarity were measured to ensure that participants thought about the words before generating attributes. In addition to ratings of the variables just mentioned, participants were asked to provide rank-ordered attributes (qualities/properties) of the words.

The attributes generated in Surveys 1 and 2 were used to compare and contrast the knowledge associated with the literal readings of each word from the target sentence with the metaphoric and literal meanings elicited by those same words in the context of that sentence. Attribute rankings and frequency of occurrence were subsequently used to determine salience, which was a critical variable for the lexical decision probes used in Experiments 1 and 2.

Participants

256 participants (184 female) whose first language was English were recruited from students taking General Psychology or Cognitive Psychology classes at UF. Participants were 17 years of age and older (M = 19.5 years) and on average had completed one year of college (M = 13.4 years). General Psychology students received two credits toward a class research requirement for participation. Cognitive Psychology students received extra credit toward their class grade, as determined by their instructor.

Materials and Apparatus

The stimulus words were the 400 topic and vehicle terms that were used in Survey 1. In order to create word lists that could be administered in one hour, 16 lists with 25 words per list were created. The lists were formed by arranging all the target words alphabetically and then selecting the first word on the list, skipping 15, selecting the 17th word on the list, skipping 15, etc. After each list was formed the remaining target words were used following the same procedure. Each list was presented to 16 participants.

Participants in this study made two rating decisions using a five point Likert scale: (1) How concrete/abstract the word was; and (2) How familiar the word was to them. A copy of the questions used to solicit ratings is included as Appendix D and the full list of stimulus words is included as Appendix E.

Procedure

The experiment was written in Visual Basic and presented using a computer monitor, keyboard, and mouse. After reading and signing informed consent forms about the nature of the experiment, participants were asked to input the following demographic data: gender, age, highest level of education completed, and current educational goal.

Participants were then presented with the 25 individual words that they were asked to rate using a 5-point Likert scale on its: (a) abstractness/concreteness, and (b) how commonly or frequently they had read or heard the word. For each word they were asked to type 3-5 qualities or attributes of the displayed word (the difference between attributes and associations was stressed during the instructions). Then they were asked to rank-order the attributes they generated in terms of their importance to the concept(s) the word represented. Reading times for the words and decision times for the ratings were recorded, as were the participant's generated attributes.

Results and Discussion

Survey 1 resulted in the establishment of a corpus of stimulus sentences that can be used in future metaphor comprehension research. The ratings collected on the stimulus sentences will allow for better manipulations of the dimensions that the literature has indicated is important for metaphor comprehension.

For the purposes of this dissertation, the most important information garnered from Survey 2 is the rank ordered attributes of each word provided by the participants. It is assumed that there is a high correlation between the rank and salience of an attribute. The attributes obtained in Survey 2 were combined with the relational attributes obtained in Survey 1 in a database with over 18,000 entries. By selecting attributes that were given to sentences in Survey 1, but that did not occur as literal or lexical associates of the topic and vehicle terms in Survey 2, probe words related primarily to the figurative meaning of the sentence were established. These probe words were then assigned as High Salient or Low Salient depending on the participants frequency of responding with that probe word. These probe words were then used in Experiment 1 and 2 as part of an

attempt to compare and contrast the processes involved in the comprehension of Novel and Conventional metaphors.

The attribute data collected in Surveys 1 and 2 will be used in future research, for example to construct a knowledge base using the Structure Mapping Engine (SME). The Structure Mapping Engine is a computational model of similarity, comparison, and analogy making that is based on the claims of Structure Mapping Theory (Gentner, 1983; Gentner & Markman, 1997; Gentner et al., 2001) that metaphors can be modeled as extended structural mappings between domains for novel metaphors but not for conventional metaphors.

CHAPTER 3
EXPERIMENT 1-CONVENTIONALITY, FORM AND ATTRIBUTE PRIMING

Experiment 1 was designed to examine implications of the Career of Metaphor Hypothesis proposed by Gentner et al., (2001) through the pursuit of three goals. The first goal of Experiment 1 was to explore how metaphors that vary in their degree of conventionalization might be understood, and how this might be linked to differences with the component processes involved in their comprehension. A second goal was to explore the influence of presentation form (metaphor or simile) on the time it took to comprehend figurative statements. The third goal of Experiment 1 was to examine the extent to that attempts to comprehend conventional and novel metaphors primed different kinds of probe words that were strongly or weakly related to the figurative meanings of the metaphor.

The Career of Metaphor Hypothesis (Gentner & Bowdle, 2001) claims that different processes are used to comprehend novel and conventional metaphors. One prediction made by the Career of Metaphor Hypothesis is that novel metaphors are processed by comparison (sense creation) and that conventional metaphors are processed by comparison and/or categorization (sense retrieval). Gentner et al., (2001) claim this is due to the polysemous nature of a conventionalized vehicle term (since it has both a domain-specific meaning and a related domain-general meaning).

In pursuit of the first goal of this experiment it was important to acquire the conventionality ratings directly for each participant. Stimulus sentence conventionality was initially determined in this study using pilot conventionality ratings. A more direct

measure was obtained by asking participants to rate the conventionality of each sentence. Given that people have different knowledge bases and levels of exposure to any given metaphor vehicle, it was deemed important to determine how each individual interpreted a given sentence. This allowed the examination of the impact of group and individual conventionality ratings within the experiment, as well as by comparison with the conventionality pilot ratings. The comparison of different sources for these ratings is important for the development of valid and reliable measures of conventional/novel stimulus materials that can be used to explore the processing of metaphors.

As discussed in Chapter 1, the Career of Metaphor Hypothesis predicts a processing advantage for novel and conventional metaphors depending on whether they are in a simile or metaphor form. In order to test the proposed form-process link, Bowdle and Gentner (1995, 1999) (as reported in Gentner et. al., 2001) conducted an experiment comparing comprehension times for novel and conventional figurative statements in the form of metaphors or similes. Their goal was to test what Bowdle termed the “grammatical concordance assumption.” Namely that the simile form (e.g., "A forest is like a harmonica") by its nature invites a comparison and will be the preferred form for the processing of novel figurative statements, while the metaphor form (e.g., "Divorce is a nightmare") invites a search for a prestored category (which does not exist in novel metaphors) and will be the preferred form for processing of conventional figurative statements.

By measuring comprehension times for figurative sentences that the experimenters had judged to be either novel or conventional, by their definition, Gentner et al., (2001) found the predicted interaction between conventionality and grammatical

form: Comprehension times for novel similes were faster than novel metaphors, but conventional metaphors were faster than conventional similes. These differential findings for the processing of novel and conventional figuratives based on presentation form lend support to the Career of Metaphor Hypothesis, and contrast with the predictions of categorization (class-inclusion) models (e.g., Glucksberg, 2001; Glucksberg & Keysar, 1990; Glucksberg, McGlone, & Manfredi, 1997). Categorization models would claim that other factors such as (1) semantic world knowledge about the topic that constrains what types of characterizations are relevant and meaningful, and (2) knowledge about the vehicle and what it can epitomize, drive comprehension times. The models fail to account for the findings of Gentner et al., (2001).

In a second study, Gentner et al., (2001) presented participants with a range of figurative statements and asked them to rate their preference for simile versus metaphor forms. As predicted, their participants strongly favored the simile form for novel figuratives. However, they showed no preference between simile and metaphor forms for conventional metaphors.

Because of the apparent importance of form in determining the impact of conventionality, both were manipulated in Experiment 1. Using the novel and conventional stimulus material developed in Surveys 1 and 2, participants' comprehension times for simile and metaphor forms were used to determine if the interactions of simile vs. metaphor form and novel vs. conventional vehicles reported by Bowdle and Gentner could be replicated with the present materials, in which conventionality was determined more systematically and verified empirically.

The third goal of this experiment was to test an implicit prediction of the Career of Metaphor Hypothesis that the pattern of priming for words/concepts related to metaphoric meaning will be different for conventional metaphors than for novel metaphors. Although numerous metaphor comprehension studies have used priming manipulations (Blasko & Connine, 1993; Galinski & Glucksberg, 2000; Gernsbacher et al., 2001; Gibbs et al., 1997; Glucksberg et al., 2001, Kemper, 1989) this is to my knowledge the first use of priming specifically related to the conventionality of a figurative sentence.

Metaphors have been used as primes for decisions about individual words in several other studies, which have shown that metaphors do indeed activate concepts that are related to various senses of vehicle terms (Gernsbacher et al, 1995, 2001; Glucksberg, Gildea, & Bookin, 1982; Kemper, 1989; Keysar, 1989; Wolff & Gentner, 2000) Using a cross-modal priming technique that manipulated familiarity and aptness, Blasko and Connine (1993) individually presented metaphors such as “The stars are snowflakes” (e.g., “stars” and “snowflakes”) auditorily. They then immediately showed target words related to the ground of the metaphor (e.g., “unique”) or to unrelated concepts (e.g., “fifteen”). Target words were determined in an independent study where the participants were asked to generate the “single word that is the most central feature of the meaning of the metaphor” to each vehicle term presented in isolation. The most frequent responses to each vehicle term were selected as metaphor-related terms. Metaphors facilitated lexical decisions to metaphor-related words, although the effect was strongest for high-familiar, apt metaphors. Degree of conventionality was not controlled. Importantly, Blasko and Connine (1993) showed that the priming was the result of processing at the

sentence level, rather than the effect of the individual words in the metaphor, since in their Experiment 5, no priming was found when the topic and vehicle terms were presented sequentially with no sentence frame (e.g., star / snowflake / unique).

In a recent study of priming by individual words (Glucksberg et al., 2001), a metaphor-related vehicle term (e.g., “shark”) was followed by either a metaphor-related target word (e.g., “vicious”) or a literally related (and metaphorically-unrelated) target word (e.g., “swim”). Metaphor-unrelated terms were also presented by using vehicle terms from unrelated metaphors (e.g., “sunshine”) followed by the terms “vicious” or “swim.” Glucksberg, et al. found significant, and equivalent, priming for words related literally and figuratively to the prime. As Gernsbacher, et al. (2001) note, however, the metaphors from which the terms were drawn were conventional and familiar. It seems likely that familiarity and conventionality would impact priming differentially; a conventionalized vehicle, being polysemous, should prime its figuratively related targets; however, a vehicle of a familiar metaphor should not, out of its metaphor frame, prime its figurative sense (e.g., chocolate–uncertain).

The Career of Metaphor Hypothesis (Gentner et al., 2001) suggests a differential pattern of priming for words following novel versus conventional figuratives. Novel figuratives should evoke a broader and deeper activation of associated concepts and words than conventional figuratives. Conventional figuratives can be understood by accessing prestored words/concepts associated with the vehicle’s figurative meaning, while novel figuratives require activation of potential words/concepts, mapping of those concepts and the acceptance/rejection of that mapping before a meaning can be derived.

Using probe words that vary in their strength of association to the figurative meanings of the stimulus sentences, lexical priming might serve as an excellent method to identify the effects of conventionalization and the breadth of lexical activation that results from the different processes involved in conventional and novel metaphor comprehension. Experiment 1 used three types of probes to assess the depth of activation for conventional and novel metaphors: High Salient figurative, Low Salient figurative, and Unrelated. The degree of priming was assessed by comparison of the lexical decision response times to the related and unrelated probes. Three hypotheses were derived concerning the pattern of priming found as a function of salience and conventionalization, when the probes immediately following comprehension of the figurative meaning of a sentence:

H1: Significant priming for words strongly related to the figurative meaning of the vehicle (High Salient) would be obtained for both novel and conventional figuratives, reflecting the activation of concepts related to the figurative sense.

H2: The degree of priming for High Salient targets was predicted to be greater for the conventional figuratives than for the novel figuratives, since the conventional figurative meaning should be well learned and more automatically and strongly activated by the figurative use of the vehicle.

H3: The degree of priming for Low Salient targets was predicted to be small or absent for the conventional figuratives, as they would be more narrowly “tuned” to the conventional, dominant figurative sense. In contrast, priming for Low Salient targets for the novel figuratives was predicted to be larger, since this would follow a more thorough

exploration of the “conceptual space” of potential figurative senses of the vehicle and how that might be mapped to an attribute of the topic term.

One aspect of Gentner’s model, however, implies a very different prediction for the Low Salient targets for novel figuratives. If, as part of the projection and mapping process, inappropriate senses are actively inhibited, then the Low Salient targets may actually show negative priming, relative to the unrelated (and uninhibited) target words.

Method

Participants

120 participants (102 female) whose first language was English were recruited from undergraduate students taking General Psychology or Cognitive Psychology classes at the University of Florida. Participants were 18 years of age and older (*M* = 19.9 years) and on average had completed nearly two years of college (*M* = 13.8 years). General Psychology students received two credits toward a class research requirement for participation. Cognitive Psychology students received extra credit towards their class grade, as determined by their instructor.

Materials

Because we were interested in manipulating conventionality, the 120 sentences from Study 1 that received figurative ratings higher than 3.0 were rated in terms of their novelty/conventionality by the pilot conventionality ratings of five independent raters. The resulting 97 stimulus sentences were rank ordered by their pilot conventionality ratings. The 36 sentences that received an average pilot conventionality rating of 1 - 2.6 (on a 5-point scale) were considered novel metaphors. Thirty-six sentences were selected from the 56 sentences that received an average rating of 3.4 – 4.6 and were considered conventional metaphors (20 sentences were rejected as target sentences because they did

not have meaningful or awkward simile equivalents (e.g., “My sister is like cool”). 24 filler items were drawn from the remaining sentences that received figurative ratings greater than 3.0 and sentences that received figurative ratings less than 3.0.

Four types of probe words [High Salient Figurative (HSF), Low Salient Figurative (LSF), Unrelated Control (UC), or Nonword (N)] were developed from the attributes collected in Surveys 1 and 2. Probe words were selected using the importance (salience) ratings and their frequency of occurrence as attributes that were given by participants in Survey 1. These words were then matched with attributes that were listed in response to the presentation of topic and vehicle terms in Survey 2. Any attributes that occurred in both Surveys 1 and 2 to a given sentence and its topic and vehicle terms were rejected as probe word candidates. This was done because the purpose of Survey 2 was to identify attributes that might have lexical associations and/or be related literally to the topic and vehicle terms. In this way the probe words selected were more likely to be related to the figurative meaning of the sentence.

For example for the stimulus sentence: “A dilemma is a cactus.”-“painful” was selected as the high salient figurative probe word because it occurred four times and received an average importance rating of 3.6 (on a 5 point scale) in Survey 1, and was not listed in isolation as an attribute of either “dilemma” or “cactus” in Survey 2. The low salient probe word was “bothersome”-because it occurred once and was not rated in Survey 1, and was not listed in Survey 2. The unrelated control word was “looming” which did not occur as an attribute that was listed to either the sentence or topic and vehicle terms in Surveys 1 and 2.

Filler sentences were always presented with nonword probes (e.g., “A cigar is a skunk” was paired with the nonword “yelms”). Appendix E contains the metaphor form of the 96 Sentences used in Experiment 1 along with the four types of probe words used.

Design

A 2 (sentence form: metaphor, simile) x 2 (conventionality type: novel, conventional) x 3 (probe: high salient, low salient, unrelated control) within-subjects design was used.

Procedure and Apparatus

Two stimuli set variations were used. In one variation half the conventional and novel stimulus sentences were presented in simile form and half in metaphor form. In the second variation they were presented in the opposite form from that presented in variation one. Aptness, familiarity, and similarity gradients were counterbalanced as much as possible between variations.

The experiment was written in Visual Basic 6.0 and was presented using computer monitors, keyboards, and mice. At the beginning of the experiment participants were asked to provide the following demographic data: gender, age, and highest level of education completed. Names were not associated with subject numbers and the participant's privacy was protected.

A brief tutorial was used to instruct the participants on the sequence of events they were asked to perform. Tutorial data were not collected. The experiment consisted of two tasks-a lexical decision priming task and a conventionality-rating task.

Sentences were randomly presented on the screen and participants were asked to read each sentence and respond to the following question: "Does this sentence have a figurative meaning?" The participants were instructed to make one keypress if the target

sentence had a figurative meaning for them and another keypress if it did not. The keys were marked with green and red stickers. The time from presentation of the participant's response was saved as comprehension time.

One second after either key-press, one of four probe types appeared participants were asked to decide as quickly and accurately as possible whether or not the string constituted an English word. (The RSI of one second was chosen in light of pilot work suggesting that with shorter intervals, the decision and response to the sentence created potential interference with the subsequent lexical decision and response.) The time from presentation of the probe string to participant response was recorded.

In 75% of the trials (72/96) the target string was a word and on 25% the trials (24/96) it was a nonword. The 72 conventional and novel target sentences were always followed by one of the three types of probe word and the 24 filler sentences were always followed by a nonword. Three sets of meaningful probes were created to allow pairing of each stimulus sentence with each probe type, (each participant only saw one such pair). The next sentence in the stimulus set appeared 1500 milliseconds after the participant's response to the last lexical decision task. This cycle continued randomly with each sentence being presented until all 96 sentences and their associated probes had been viewed and responded to.

Upon completion of the sentence comprehension and identification phase, participants engaged in a distractor task for approximately 15 minutes. The distractor task was a reading comprehension test selected from the Nelson-Denny Reading Test, Form H (Brown, Fishco, & Hanna, 1993). Participants were then asked to rate the conventionality of the vehicle terms that were presented in the 72 target stimulus

sentences (using a Likert scale of 1 – 5, with “1” indicating highly novel sentences and “5” indicating highly conventional sentences). Participants received a brief tutorial on conventionality and made several practice ratings to ensure they understood the task.

Results and Discussion

Results are presented for each of the three measures of Experiment 1. In the first phase Figurative Meaning decision times and lexical decision response times to the various probes are considered. Because this study is concerned with the comprehension of figurative sentences, sentences that were indicated by participants as meaningless were not analyzed further. Table 3-1 displays the percent of sentences judged meaningful and meaningless for novel and conventional figuratives in simile and metaphor forms. Note that roughly 80% of conventional figuratives were judged meaningful, while less than 50% of the novel figuratives were judged meaningful.

Table 3-1. Percent of Sentences Judged Meaningful and Meaningless for Conventional and Novel, Metaphors and Similes

	Conventional Metaphor	Conventional Simile	Novel Metaphor	Novel Simile
Meaningful	82 %	81 %	40 %	46 %
Meaningless	18 %	19 %	60 %	54 %

For sentences that had been judged meaningful the comprehension times and decisions to the various types of probes were analyzed. As a test of the effects of form and its interaction with conventionality, differences between the comprehension times for sentences presented in conventional metaphor form are compared with those that were presented in simile form. For all analyses, sentence conventionality was determined by obtaining the average conventionality ratings across participants for each sentence. 37 of the stimulus sentences (with average ratings conventionality ratings greater than 3.0)

were identified as Conventional and 34 sentences (with average conventionality ratings less than 3.0) were identified as Novel.

In all analyses decision times occurring outside 3 standard deviations of the distribution of responses for the same type were rejected as outliers. It was determined that the best way of representing the data was by determining the median decision times for each participant and then averaging these times across subjects. This approach is in keeping with the recommendations made by Ratcliff (1993) and assures that each subject contributed one value to the analysis (thereby avoiding unequal numbers of responses in each of the response cells). Analyses of variance (ANOVA's) were performed on all data using an alpha level of .05, making a sphericity-correction with the Huynh-Feldt (H-F) statistic.

Figurative Meaning Decision Times

Figure 3-1 shows the mean decision time to judge sentences as figuratively meaningful, as a function of conventionality and form. Each participant saw half of the sentences in simile form and the other half in metaphor form.

A 2 (Conventionality Type) x 2 (Sentence Form) ANOVA conducted on these data indicate that decisions of meaningfulness are made faster to conventional metaphors than to novel metaphors, $F(1,64) = 65.63$, $p < .001$. These findings are consistent with those of previous researchers (e.g., Blank, 1988; Blasko & Connine, 1993; Gentner & Bowdle, 2001) who found that overall comprehension times were faster for conventional figuratives than for novel figuratives. Although no specific hypothesis was generated regarding speed of comprehension for novel versus conventional figuratives, it is one confirmation that the conventionality ratings derived from the participants is valid, and is

a reasonable approach to defining conventionality/familiarity of figurative uses of vehicles.

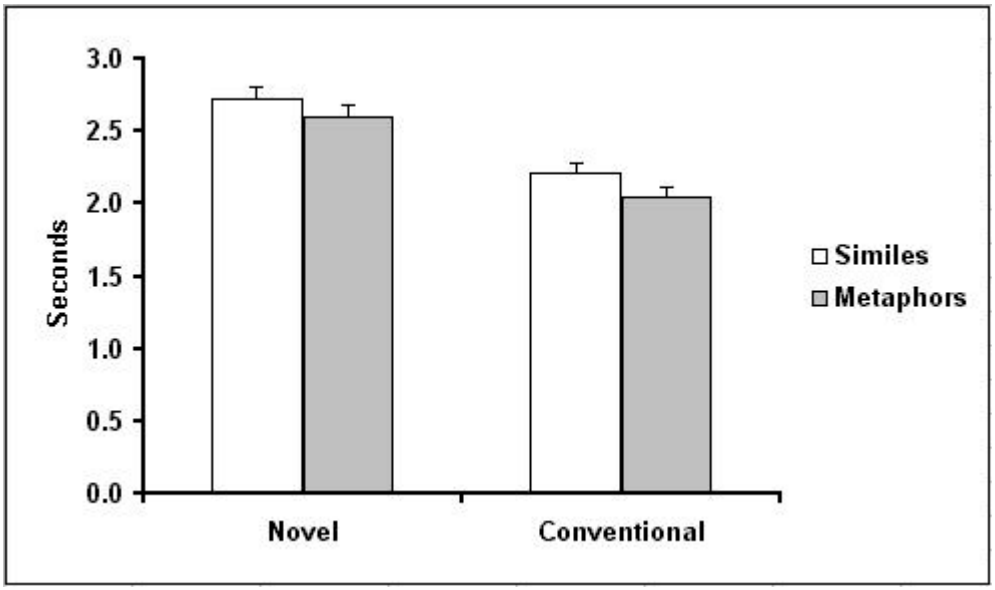

Figure 3-1. Meaningfulness Decision Times (in sec) for Meaningful Sentences based on Conventionality Type (Conventional, Novel) and Form (Simile, Metaphor) (Standard Error Measures are Indicated by Error Bars)

Participants also made meaningfulness judgments faster to sentences when they were in the form of a metaphor (2600 ms for novel metaphors and 2048 ms for conventional metaphors) compared to when they were in the form of a simile (2720 ms for novel similes and 2217 ms for conventional similes). The 2 x 2 ANOVA indicates that there is a main effect of Sentence Form, $F(1,64) = 8.56$, $p = .005$ indicating a small but reliable difference between the time it takes to make a judgment of meaningfulness for similes and metaphors. There was no indication of an interaction between Conventionality type and sentence form, $F(1,64) = 0.769$, $p < .385$.

This lack of interaction fails to replicate the “form preference” findings of Bowdle and Gentner (2001) that novel metaphors were read faster in simile form than in metaphor form with the opposite being true for conventional metaphors, which were read

faster in the categorization form. However, these analyses do not take into account for the word length of each sentence (3 – 6 words). On average, simile forms are longer than metaphor forms, and reading times should control for this superficial difference. To account for this, new analyses using meaningfulness decision times adjusted for sentence length were conducted. Figure 3-2 shows the meaningfulness decision times per word for Conventional and Novel figuratives in simile and metaphor forms.

After correcting for sentence length, the effects of Sentence Form are reversed. Meaningfulness decision times to both novel and conventional metaphors are made faster when the sentences are presented in simile form than when they are in the form of a metaphor. A 2 (Sentence Form: Metaphor, Simile - between subjects) x 2 (Conventionality Type: Novel, Conventional – within subjects) ANOVA showed main effects for Conventionality Type, $F(1,64) = 38.62$, $p < .001$, again showing faster times for conventional sentences of about 100 ms per word.

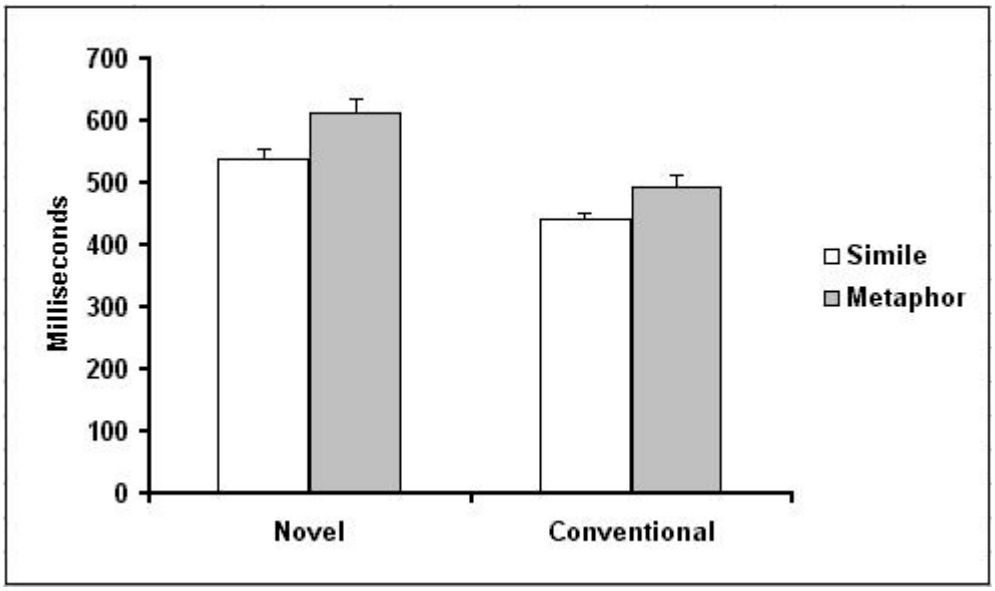

Figure 3-2. Meaningfulness Decision Times in terms of Speed-per-word (in msec) for Meaningful Sentences based on Conventionality Type (Conventional, Novel) and Sentence Form (Simile, Metaphor) (Standard Error Measures are Indicated by Error Bars)

There was also a main effect for Sentence Form, $F(1,64) = 11.48$, $p < .002$. Similes were processed approximately 65 ms a word faster than were metaphors.

Although there was some evidence for the form preference findings of Bowdle and Gentner (2001) in that there was a slightly larger advantage for novel similes over metaphors (c. 80 ms) than for conventional similes over metaphors (c. 50 ms), the interaction between Conventionality Type and Sentence Form did not approach significance, $F(1,64) = 0.133$, $p < .718$.

Even with the sentence length adjustment for word processing times, we were unable to fully replicate Bowdle and Gentner's (2001) "form preference" findings. In order to rule out the possibility that other factors such as familiarity, concreteness or word complexity were differentially influencing the effects of sentence type for conventional and novel figurative sentences, the topic and vehicle terms of all sentences were submitted to several analyses via the MRC Psycholinguistic Database (accessible at http://www.psy.uwa.edu.au/mrcdatabase/uwa_mrc.htm). The results of those analyses, coupled with the ratings of concreteness and familiarity obtained on those terms from Survey 2 are contained in Table 3-2.

The number of syllables and ratings of familiarity, concreteness, and imagability were comparable in all four cells of analysis. The various analysis of written and verbal word frequency also tended to be very similar with the exception of a few outliers that notably skewed the averages (e.g., the frequency of "time" is reported as 1599 in the Kucera-Francis written frequency table, 8599 in the Thorndike-Lorge written frequency table, and 415 in the Brown verbal frequency table – this serves to drive the averages up significantly in all three scales). Consequently we may conclude that the stimulus

material is fairly matched between cells, making it unlikely that these factors are influencing differences between novel and conventional metaphor processing. It is interesting to note that when the scales are matched, the ratings for familiarity and concreteness obtained in Survey 2 are very similar to those provided by the MRC database-thus providing a source of external validity to the ratings provided in Survey 2.

Table 3-2. Analyses of Novel and Conventional Topic and Vehicle Terms via the MRC Psycholinguistic Database.

	Conventional Topic	Novel Topic	Conventional Vehicle	Novel Vehicle
Number of syllables	2.22 (*1.07*)	2.00 (*0.99*)	1.75 (*0.97*)	2.25 (*0.87*)
Kucera-Francis written frequency (Range 1 – 69961)	147 (*322*)	39 (*38*)	32 (*63*)	44 (*145*)
Thorndike-Lorge written frequency (Range 1 – 30,000,000)	929 (*1876*)	325 (*381*)	242 (*481*)	208 (*852*)
Brown verbal frequency (Range 1 – 6833)	40 (*93*)	5 (*6*)	6 (*8*)	17 (*42*)
Familiarity rating (Range 100 – 700)	543 (*56*)	545 (*51*)	511 (*56*)	460 (*90*)
Familiarity ratings from Survey 2 (Range 1 – 5)	4.43 (*0.41*)	4.03 (*0.73*)	4.05 (*0.55*)	3.55 (*0.65*)
Concreteness rating (Range 100 – 700)	474 (*128*)	539 (*85*)	574 (*50*)	564 (*56*)
Concreteness ratings from Survey 2 (Range 1 – 5)	3.42 (*1.26*)	3.66 (*1.17*)	4.12 (*0.80*)	4.34 (*0.62*)
Imagability rating (Range 100 – 700)	516 (*98*)	564 (*60*)	577 (*49*)	554 (*60*)

Notes: *SD*'s are presented in italics in parenthesis.
Each scales range is denoted parenthetically below the scale name

One explanation for the lack of form preference for conventional metaphors might be that since the alignment and mapping between domains has been previously achieved, a simile form (which includes a subset of viable property attributions attained in the fuller mapping of the metaphor form) seems as likely an interpretation as the metaphor form.

In either case, the mental work involved in understanding the metaphor has already been done.

Lexical Decision Times

Figure 3-3 reflects the lexical decision response times to the two sentence conventionality types (novel or conventional) and the three probe types (high salient, low salient, and unrelated control). For the purpose of this analysis data from both similes and metaphors were combined together in order to provide more power. This decision was made owing to the small number of responses in each cell (3 per cell on average for both similes and metaphors), as well as to the absence of any predictions about lexical decision priming as a function of whether the figurative statement had been presented in simile or metaphor form – the Career of Metaphor Hypothesis claims that comprehension may be slowed by the "incongruent" form, but the nature and outcome of the comprehension process is similar regardless of form.

The data were reduced by obtaining the median response time for each participant for sentences they identified as meaningful and following correct lexical decisions to probe words for each of the six cells in the 2 x 3 (Sentence Conventionality x Probe Type) matrix. All median response times within each cell were then averaged together.

As can be seen in Figure 3-3, overall, correct lexical decisions are made faster to probes following conventional sentences than to novel sentences. A 2 x 3 (Sentence Conventionality x Probe Type) ANOVA conducted on these data confirmed that there was a main effect of Sentence Conventionality, $F(1,103) = 10.186$, $p < .005$. It appears, however, that this effect is driven largely by the substantial difference in decision times for the low salient probes following conventional versus novel sentences.

The main effect of Probe Type was also significant, $F(2,103) = 6.327$, $p = .002$, with High Salient probes responded to most quickly (c. 760 ms), Low salient probes intermediate (c. 790 ms) and Unrelated probes the slowest (c. 820 ms). The presence of significant priming for the High Salient probes confirms the first hypothesis that priming for attributes related to the vehicle's figurative sense would be obtained following comprehension of the sentence's figurative meaning. The similarity of the degree of priming for High Salient probes following novel and conventional sentences disconfirms the second hypothesis, which predicted greater priming following conventional sentences for these probes.

Most importantly, there was a significant interaction between Sentence Conventionality and Probe Type, $F(2,103) = 4.515$, $p = .014$. The cause of this interaction can be seen in Figure 3-3 in the overall pattern of responses for lexical decisions: Novel and conventional sentences show substantial, and very similar, priming for the High Salient probes (c. 65 ms relative to unrelated probes); but while priming for Low Salient probes is as great as that for High Salient probes following conventional sentences; there was no priming at all (in fact, a small and non-significant negative priming effect) for Low Salient probes following novel sentences.

Paired comparisons indicated priming effects by revealing significant differences in the lexical decision times between: High Salient Novel and Low Salient Novel, $t(112) = -3.03$, $p < .004$ (two-tailed); Low Salient Conventional and Unrelated Conventional Controls, $t(123) = -4.59$, $p < .001$ (two-tailed); and High Salient Conventional and Unrelated Conventional Controls, $t(123) = -4.16$, $p < .001$ (two tailed). The priming

effects for High Salient probes following Novel and conventional sentences were comparable.

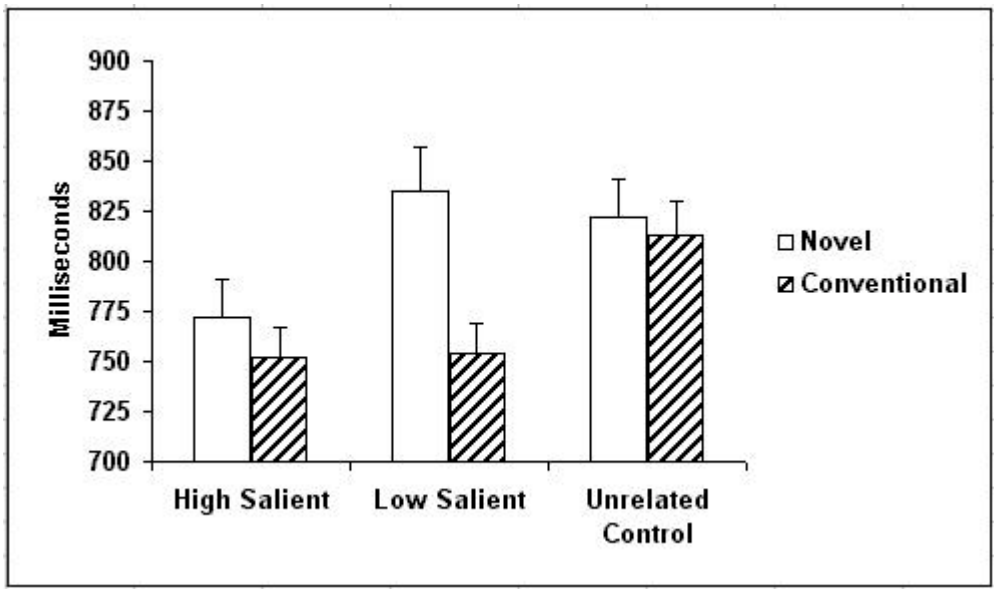

Figure 3-3. Lexical Decision Times (in ms) following Meaningful Sentences based on Sentence Conventionality (Novel, Conventional) and Probe Type (High Salient, Low Salient, Unrelated) (Standard Error Measures are Indicated by Error Bars).

This pattern of lexical decision times using the various types of probes supports the broad prediction of the Career of Metaphor Hypothesis that the pattern of priming for words/concepts related to metaphoric meaning will be different for conventional metaphors than for novel metaphors. The Career of Metaphor Hypothesis proposes that novel metaphors are processed by comparison (sense creation) and that conventional metaphors are processed by comparison and/or categorization (sense retrieval). The comparison process should take longer than the sense retrieval involved in categorization. Presumably in order to facilitate this process there must be a broader and deeper activation of related words and concepts. Support for this idea was seen earlier in the slower decision times for the sentences themselves. Using the notion of strength of activation, one can conjecture that the reason that there is very little difference between

the responses made to High Salient probes for novel and conventional figurative sentences has to do with the fact that as figurative attributes of the sentence they are highly related to the figurative meaning of the sentence and are primed by the figurative sentence which preceded them.

The pattern of responses to Low Salient probes is not explained as easily. However, some help comes from Gernsbacher's (1990) Structure Building Model of text comprehension which posits the inhibition of concept nodes in memory that are not relevant to the text representation as that representation is constructed, and enhancement of concept nodes that are relevant to the constructed text representation. In this model properties of metaphor vehicles irrelevant for metaphor comprehension are initially activated and then suppressed. According to Gentner et al., (2001) this process would only be required for figuratives that require a comparison process and would not be necessary for figuratives that use a categorization process.

The differential pattern of responses to low salient probes following figurative sentences shown in Figure 3-3 can be explained by combining Gernsbacher's and Gentner's ideas. Novel figuratives would facilitate a broader and deeper activation of associated concepts and words in order to arrive at an understanding of the sentence. Those concepts/words that were deemed irrelevant to the meaning of the sentence would be suppressed. When a figuratively related probe word (that presumably is or is related to the concepts and words just suppressed) is shown, the participant must overcome the suppression in order to respond to the lexical decision task. Overcoming this suppression could account for the differences in response time to the low salient probes following novel and conventional figurative sentences.

The results of the lexical decision task in Experiment 1 support the notion that a priming task may be able to differentiate between conventionalized and novel metaphors. This notion was tested further in Experiment 2, where novel metaphors were used as stimulus sentences and the lexical decision time of Low Salient probes was correlated with ratings of conventionality in several conditions that attempted to increase a given sentence's conventionality.

Conventionality Ratings

By comparing the averaged conventionality ratings for each sentence provided by the participants with the averaged Pilot Conventionality Ratings one gets a feel for how stable conventionality ratings are. There was a 91.6% match between the ratings of novel and conventional figurative sentences and a Pearson Correlation Coefficient of $r = 0.93$ between the averaged scores. The average overall standard deviation for the conventionality ratings was 1.08 indicating that the decision to combine ratings of 1 and 2 to create novel sentences and ratings of 3 and 4 to create conventional sentences resulted in reasonably stable ratings.

The stability of averaged ratings is important for the selection of sentences to be used in studies that explore the effects of conventionality on metaphor comprehension. Nonetheless, there are individual differences between subjects. As we saw in Survey 1, based on their experience with a figurative or metaphorical phrase, participants will assess the conventionality of each stimulus sentence differently. It is important to take these individual differences into account when high and low salient probes are used and lexical decision tasks are employed.

CHAPTER 4
EXPERIMENT 2-PRACTICE, CONTEXT, AND CONVENTIONALIZATION

The Career of Metaphor Hypothesis (Gentner & Bowdle, 2001) contends that different processes are involved in the comprehension of novel and conventional metaphors. It therefore implies that a shift in cognitive processing occurs during the conventionalization of a novel metaphor. Developing methods to reliably indicate this shift during the conventionalization of a novel metaphor would strongly support this view and may provide the means to reconcile differences between the Comparison (Structure Mapping) and the Categorization (Class Inclusion) views of metaphor comprehension.

Experiment 1 used comprehension time; direct ratings of novelty/ conventionality; and patterns of lexical decision times to determine the degree of conventionalization that existed for a given sentence and the consequences of that conventionalization. Several aspects of the results supported the claims of the Career of Metaphor Hypothesis: Conventional figuratives, which could be reliably distinguished from novel ones by ratings, demonstrated faster comprehension speeds, and strikingly different patterns of priming of subsequent words. Stronger support would be provided if the shift in the conventionalization of a novel vehicle could be observed “in vitro.” The goal of Experiment 2 was to explore the "conventionalization" of metaphors in an experimental setting by providing different kinds of experience with novel figurative sentences prior to assessing their degree of conventionalization.

Specifically, Experiment 2 looked at the effects of three contrasting types of experience (here termed: Repetition, Varied Topic Encoding, and Context) on the

conventionalization of novel metaphors. These are defined and described below. Because the patterns of responses to lexical decisions made following simile and metaphor forms of novel and conventional figuratives were similar to each other in Experiment 1, it was decided to focus on the metaphor form in this study.

Repetition. Repetition was selected as a condition based on Chiappe and Kennedy's (2001) claim that it is important to establish the difference between Familiarity and Conventionality. Gentner & Bowdle (2001) note that conventionalization involves repeated figurative uses of a given vehicle term, so that it acquires domain-general meaning. Familiarization, however, involves repeated exposures to specific topic-vehicle pairings. Presenting the same novel sentences three times each during the learning phase should allow participants to become very familiar with them. However, repetition is predicted to have very little effect on changes in the perceived conventionalization of the sentence. This predication is based on the notion of Structure Mapping Theory (Falkenhainer, Forbus, & Gentner, 1989; Gentner, 1983) that a lack of "mapping" sufficient for "sense creation" to occur takes place in repeated exposures to the same sentence. In short, the participants are not expected to do the mental work necessary to do the extensive and thorough mapping that is required for the conventionalization of a figurative sentence.

Varied Encoding. Very little work has been conducted in the metaphor comprehension literature employing varied topic encoding. In one study conducted by Bowdle and Gentner (1999) participants were presented with several novel figurative statements using the same vehicle by presenting their participants with two novel similes and asking them to create a third that was similar in meaning to the first two (e.g., "Doubt

is like a tumor" / "A grudge is like a tumor" / "____ is like a tumor"). Each participant saw eight of these sets, as well as other sets that contained literal similarity statements, metaphors, and categorizations. Then participants were asked to indicate their preference for one of two test sentences that contained the same vehicle in both simile and metaphor form (e.g., "An obsession is a tumor" / "An obsession is like a tumor"). Participants who saw two novel similes for a target vehicle were more likely to prefer the metaphor form, compared with participants who had seen the same base terms in literal similarity statements.

Bowdle and Gentner (1999) argued that carrying out a series of parallel figurative alignments facilitated the creation of an abstraction that favored the metaphor form. Bowdle and Gentner therefore suggested that the experience with varied figurative frames facilitated the process of conventionalization. However, they only used the preference test, trained only with similes, tested with a very short lag between variations, and did not manipulate the degree of variation during training.

This condition of Experiment 2 attempted to directly test the prediction that varied encoding could be used to achieve the conventionalization of metaphors in vitro. This was accomplished by using a wide range of topics in learning sentences that would encourage different mappings with each presentation but stabilize or conventionalize the vehicle term.

It was hypothesized that this manipulation will increase the conventionality ratings of novel metaphors compared the conventionality ratings for their cohort sentences in Experiment 1. (Sentences from Experiment 1 provided baseline measures of conventionality since they were only seen once). Additionally, the conventionality

ratings of conventional metaphors were hypothesized to remain similar to their cohort sentences in Experiment 1. This is because the categorization process involved in understanding conventional metaphors should not benefit from the opportunity to conduct multiple mappings (since the process does not require mapping). One last prediction is that the overall conventionality ratings given to metaphors in this condition should be higher than the conventionality ratings given to cohort sentences in the repetition condition.

Context. Perhaps not surprisingly, manipulations of context have been widely used in metaphor comprehension studies (e.g., Blasko & Briihl, 1997; Bortfeld & McGlone, 2001; Gerrig & Healy, 1983; Gildea & Glucksberg, 1983; Keysar, 1989; Peleg, Giora, & Fein, 2001; Pynte et al., 1996; Vu et al., 1998). The use of context, in both the local sense of the words and sentences preceding an utterance, and more globally, in terms of the comprehender's knowledge about the world and about the topic of discourse, has been an important part of the study of language comprehension in general. These "top-down" conceptually driven processes might be considered even more important in determining how figurative expressions are understood.

When metaphorical meanings are not readily available, then a minimal context can make them so (Gildea & Glucksberg, 1983). Evidence of this comes from the "metaphor interference effect" (slower rejection of literally false statements that have figurative senses) first shown by Gildea and Glucksberg (1983) only being found when the metaphoric meaning was primed by a preceding context. Given the critical role of context in comprehension generally, it seemed reasonable to expect that the process of metaphor conventionalization would be impacted by context, but there has been little

theoretical or empirical work on the effects of context upon the conventionalization of a metaphor.

If the Career of Metaphor Hypothesis is correct, then the responses to studied novel metaphors are expected to behave, to differing degrees depending on the learning manipulation, like unstudied conventional metaphors. Based on the findings of Experiment 1 this should be true for the measures of: (a) comprehension speed, (b) the pattern of lexical decision priming, and (c) assigned conventionality ratings. The variable topic learning condition was predicted to result in the greatest conventionalization of novel sentences. Therefore, the sentences in this condition should show the greatest similarity to the previously conventionalized metaphors. The repetition condition, in contrast, should provide less experience with the metaphoric domain of the vehicle and produce the least amount of change between the conventionalization of studied and non-studied novel metaphors. Repetition is predicted to increase familiarity. If novelty remains virtually unchanged in light of this increase in familiarity, then we should have compelling evidence for the dissociation of familiarity and conventionality.

The effects of the context condition depend on whether it facilitates the mapping and comparison process. If it makes conceptual mapping unnecessary, then those vehicles used with novel topics (the studied novel metaphors) might behave like unstudied novel metaphors. On the other hand, if context facilitates but does not replace the mapping process that occurs at the time of initial metaphor comprehension, in a sense “teaching” the mapping to the reader, then the studied novel metaphors should behave more like sentences in the varied topic condition.

Method

Participants

270 participants (186 female) whose first language was English were recruited from undergraduate students taking General Psychology or Cognitive Psychology classes at the University of Florida. Participants were 18 years of age and older (*M* = 19.4 years) and on average had completed at least one year of college (*M* = 13.3 years). General Psychology students received two credits toward a class research requirement for participation. Cognitive Psychology students received extra credit towards their class grade, as determined by their instructor.

Materials

Thirty novel metaphors and 30 conventional metaphors (identified by the Pilot Conventionality ratings from Survey 1 with concordant conventionality ratings from Experiment 1) were selected as the target stimulus sentences in this experiment. Twenty Filler items were drawn from the 72 filler items used in Experiment 1. Target sentences retained the same three probes (HSF, LSF, and UC) as they did in Experiment 1. Filler sentences were also associated with the same nonword probes as in Experiment 1.

Design

A 3 (Learning: repetition, variable topic, context) x 2 (Metaphor Type: novel, conventional) x 2 (Study: studied, not studied) x 3 (Probe type: high salient figurative, low salient figurative, unrelated control) mixed-subjects design was used. Learning was varied between subjects, while Metaphor Type, Probe Type and Study were within-subjects factors.

Procedure and Apparatus

Two stimuli set variations were used. In one variation half the conventional and novel stimulus sentences (15 each) were seen in the learning phase and half were not. In the second variation seen and unseen sentences were swapped.

The experiment was written in Visual Basic 6.0 and was presented using computer monitors, keyboards, and mice. At the beginning of the experiment participants were asked to provide the following demographic data: gender, age, and highest level of education completed. Names were not associated with subject numbers and the participant's privacy was protected.

A brief tutorial was used to instruct the participants on the sequence of events they were asked to perform. Tutorial data were not collected. The experiment consisted of two phases - a learning phase and a test phase. After the test phase participants were asked to rate the conventionality of the vehicle terms that were presented in the 60 target stimulus sentences on a scale of 1 – 5.

Learning Phase: There were three different types of learning phase, one for each experimental condition. In each condition half of the learning sentences were randomly presented on the screen (studied) half of the learning sentences were withheld (not-studied). Participants were asked to read each presented sentence and respond to the following question: "Does this sentence have a figurative meaning?" The participants were instructed to indicate if did or not with a key-press. The time from presentation of the sentence to the participant's response was recorded. Specific additional details of each condition are discussed below.

In the Repetition condition the learning phase consisted of the random presentation of 90 sentences. These 90 sentences were actually three repetitions of the 15 novel and 15 conventional target sentences.

In the Variable Topic condition, the learning phase also consisted of the random presentation of 90 sentences. These 90 sentences were made up of three variations of the 30 target sentences in which the same vehicles were paired with different topics. For example for the target sentence "Creativity is a toaster." in the next phase participants would see these three sentences randomly presented:

1st presentation: "Maturity is a toaster."

2nd presentation: "Impatience is a toaster."

3rd presentation: "Inspiration is a toaster."

Alternate topics were selected that shared the different figurative senses that were stressed in the target sentence. However, care was taken to avoid selecting topics that were synonymous or completely interchangeable with the topics of the target sentence. This was done to facilitate the hypothesized contextual mapping necessary for the conventionalization of the topic terms.

In the Context condition, 30 target sentences were presented following the presentation of a context sentence that was intended to make the target sentence more comprehensible and apt. When they finished reading the context, participants pressed the space bar and the associated target sentence appeared. For example:

"Incubation is the process by which thoughts or ideas are developed unconsciously and then suddenly surface in consciousness when they are done."

Target Sentence: "Creativity is a toaster."

Reading times were collected for both the context sentence and the target sentence.

Test phase: Immediately after completing the learning phase, participants began the test phase of the experiment, which was similar to the procedures of Experiment 1. There were 80 trials in the test phase consisting of: 15 studied and 15 non-studied novel and conventional metaphors, and 20 filler sentences. On each trial, as the metaphor was shown participants were asked to decide if the sentence had a figurative meaning. One second after keypress response, one of four probe types [High Salient Figurative (HSF), Low Salient Figurative (LSF), Unrelated Control (UC) or Nonword (N)] appeared. The participants were asked to decide whether or not the string of letters constituted an English word and indicate their decision with a keypress. The time from presentation of the probe to participant response was recorded. As in the previous surveys and experiments, responses outside of three standard deviations of the mean response time for each probe type were rejected as outliers.

Upon completion of the test phase, participants underwent the same distracter task as that presented in Experiment 1 (the Nelson-Denny Reading Test, Form H) for approximately fifteen minutes. Participants were then asked to rate the conventionality of the vehicle terms that were presented in the 60 target stimulus sentences using a scale of 1 – 5.

Results and Discussion

Results are presented for the three measures collected during Experiment 2. First, figurative meaning decision times to the sentences presented during the learning phase in the repetition and variable topic conditions are compared with the figurative meaning decision times to same metaphors in Experiment 1. Second, lexical decision response

times to the various probes presented in the test phase are compared across conditions. Finally, conventionality ratings from the final phase of the experiment are compared.

Based on recommendations made by Ratcliff (1993) all results are reported as median times averaged within each subject so that each participant only contributed one value to the analysis. This method was selected because it was thought to better represent the data and avoids confounds created by the disproportionate identification of sentences as novel or conventional by participants.

Figurative Meaning Decision Times

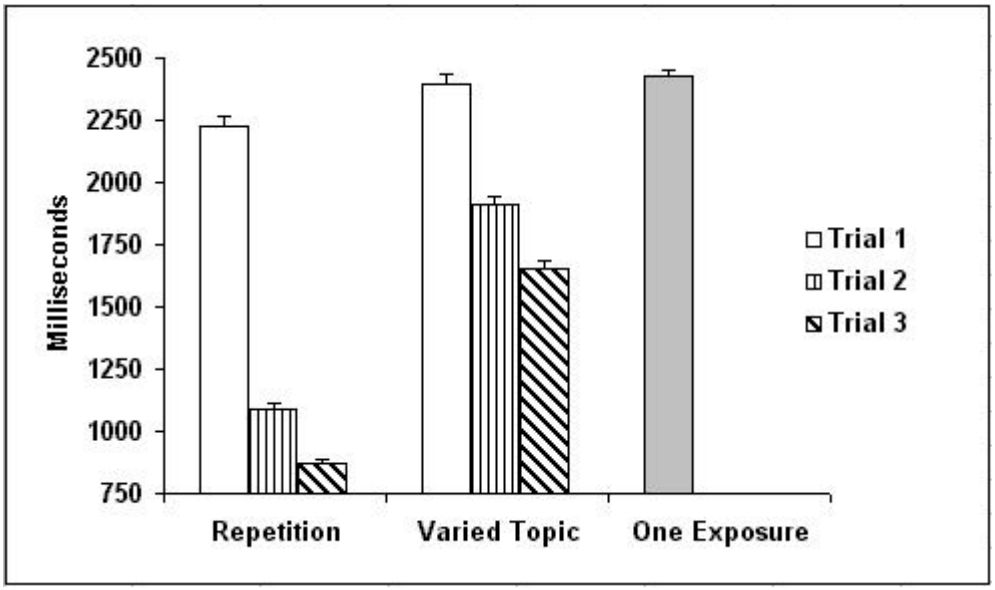

Figure 4-1. Median Meaningfulness Decision Times (in ms) for All Sentences in Experiment 2 (Repetition and Varied Topic Conditions) and Experiment 1.

Figure 4-1 compares figurative meaningfulness decision times for all sentences seen in Experiment 2 in the Repetition and Varied Topic conditions with the matched sentences from one exposure in Experiment 1. The first readings of each sentence took roughly the same amount of time in Experiment 1 and both conditions of Experiment 2. Second and third readings took much longer in the Varied Topic condition than in the Repetition condition. This was to be expected since familiarity with the same sentence

would lead to faster response times. The second and third readings in the Varied Topic condition were new sentences. Although this was not predicted, the decreasing reading times are probably attributable to the fact that only the topic changed from the previously read version, and that at least some of the conceptual mapping needed to understand the new metaphor would have already been done. It can also be inferred from the graph that participants expended more mental effort in the reading of the Varied Topic sentences than in the repetition condition. This is an important requisite if we are to believe Gentner's claim that additional mappings take place with varied sentences.

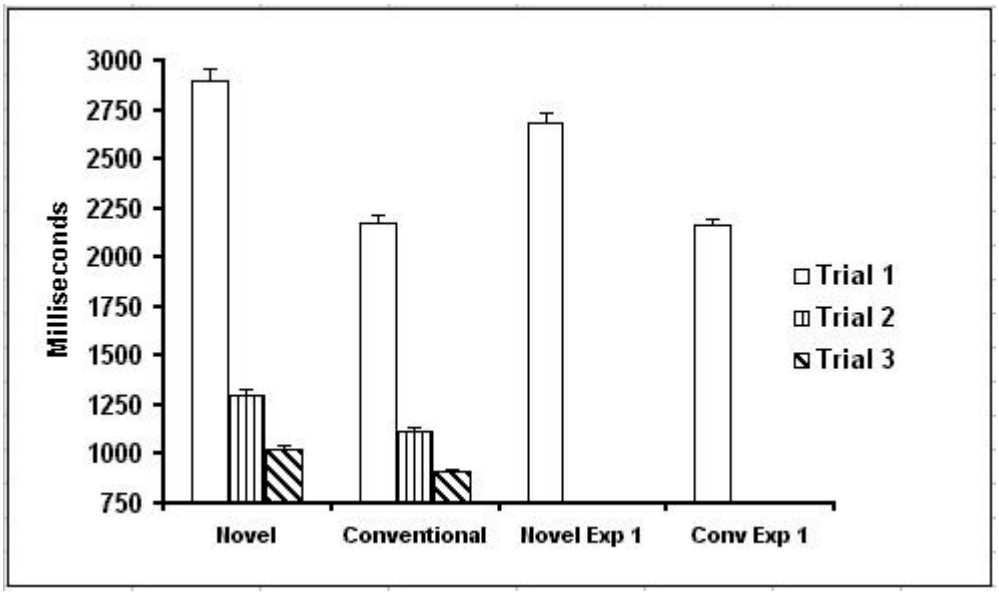

Figure 4-2. Comparison of Average Meaningfulness Decision Times (in ms) for Novel and Conventional Sentences Judged as Meaningful in Experiment 2 (Repetition Condition) and Experiment 1.

Because we are interested in the differential processing of Novel and Conventional metaphors, decision times were examined separately for novel and conventional metaphors. Figures 4-2 and 4-3 show the effects conventionality has on meaningfulness decision times in the Repetition and Varied Topic conditions respectively.

Figure 4-2 shows the same trend in meaningfulness decisions in the Repetition condition that we saw in the reading times for the novel and conventional sentences in Experiment 1. Conventional sentences are deemed meaningful faster than novel sentences. Very similar patterns exist for meaningfulness decisions about both novel and conventional metaphors in that both decrease drastically between the first and second readings of the sentence.

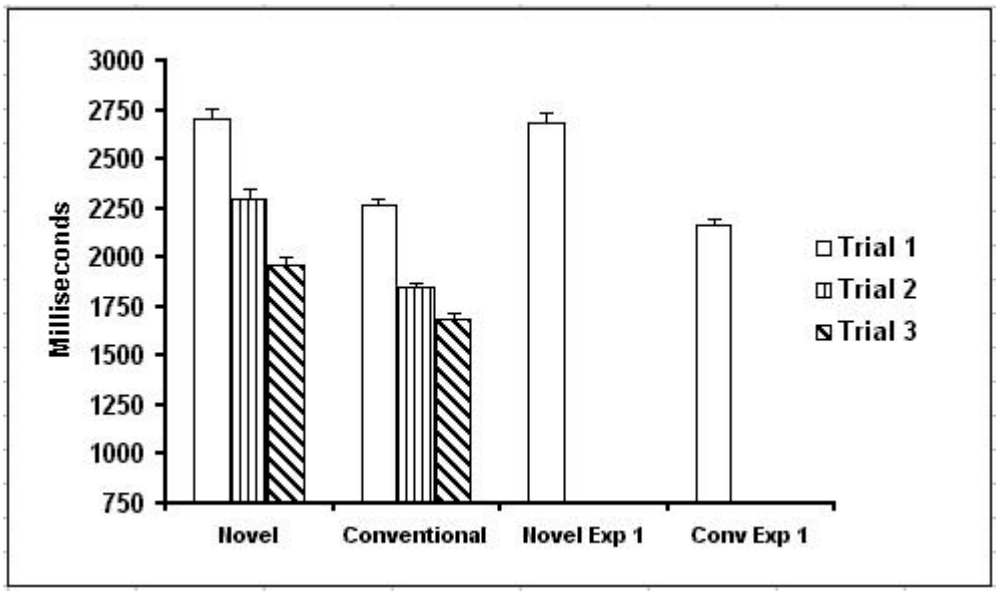

Figure 4-3. Comparison of Average Meaningfulness Decision Times (in ms) for Novel and Conventional Sentences Judged as Meaningful in Experiment 2 (Varied Topic Condition) and Experiment 1.

Figure 4-3 shows that subsequent readings result in faster meaningfulness decision times for both novel and conventional sentences in the Varied Topic condition as well. Additionally, it replicates the conventional versus novel processing advantage seen in Experiment 1. There is a trend toward a smaller effect of Trial for the Conventional sentences than for the novel sentences, but the interaction was not significant.

It was expected that in the Repetition condition, second and third exposures to the same metaphor would have very little impact on the perceived meaningfulness of that

metaphor. To explore this prediction, the percent of change that occurred with repeated exposures was calculated for each participant, and the average change across participants derived. For the most part, the overall proportion of sentences judged to be meaningful did not change across trial, somewhat surprisingly, either for novel or conventional sentences. Some sentences became more meaningful (the expected trend), yet others became less meaningful (the unexpected trend).

In the Repetition condition, including the control sentences, 11% changed across the three trials, with 6% percent becoming more meaningful with repeated exposure and 5% became less meaningful. These proportions were nearly identical when broken down by conventionality.

Shifts in meaningfulness were somewhat greater in the Varied Topic condition, with nearly 30% showing some kind of shift. Overall, 12% of all sentences (including control sentences) became more meaningful, while 17% became more meaningless. 14% of the conventional metaphors become more meaningful while 12% became more meaningless. Novel metaphors showed 9% became more meaningful and 19% became more meaningless.

One possible explanation is that as participants read a novel sentence like "Maturity is a toaster" (which 30% of the participants indicated became less meaningful after seeing the variants: "Impatience is a toaster" and "Inspiration is a toaster") they created a mapping that made this sentence meaningful. In order to do this for a vehicle like “toaster” which is more novel (with an overall conventionality rating of 1.68) they had to expend more mental effort. Evidence for this is provided by the finding that for those that found the follow on sentences to "Maturity is a toaster" more meaningless, the

median meaningfulness decision time was 2,750 ms compared to 2,060 ms (the median meaningfulness decision time for all novel sentences in the varied topic condition). It is possible that when the participants were presented with the second and third variations of the sentence, their prior mappings prevented them from creating new mappings that allowed them to find the sentences meaningful.

Lexical Decision Times

Figures 4-4, 4-5, and 4-6 show the lexical decision times to the three types of probes following novel or conventional sentences, for the Repetition, Varied Topic, and Context conditions respectively. To facilitate interpretation, each graph is presented in three parts. Part A (to the right) is a small inset of Figure 3-3 from Experiment 1 (the Lexical Decision Times (in ms) for figuratives following meaningful sentences based on Sentence Conventionality and Probe Type). This inset is provided to remind the reader of the lexical priming effect found in Experiment 1 and to allow comparison with the results of Experiment 2. Part B (upper left) shows the results for unstudied sentences and serves as a baseline measure within each condition. In each Figure, the pattern of decision times appears strikingly similar for unstudied sentences of Experiment 2 and the comparable sentences of Experiment 1. Part C (lower left) shows the results for the studied sentences. A comparison of the patterns is associated with each figure.

In the Repetition condition, a comparison of the pattern of lexical decision responses associated with Not-Studied sentences (Figure 4-4B) and those associated with Studied sentences (Figure 4-4C) shows a similar gross pattern for both. A pattern that is very similar to that seen in Experiment 2 (Figure 4-4A). Namely that lexical decisions made to Low Salient probes following meaningful conventional sentences are more similar to those made to high salient probes than to unrelated control sentences. While

lexical decisions made to Low Salient probes following meaningful novel sentences are more similar to those made to unrelated control sentences than to high salient probes.

In the Repetition condition, study generally decreased lexical decision times by roughly 25 ms. The two exceptions to that were; the lexical decision response times to Low Salient probes following conventional metaphors (which were the same in both the not-studied and studied conditions), and the increase in lexical decision response times to High Salient probes following novel sentences. This last trend is counter to the prediction made that repetition should not influence lexical decision response times to High Salient probes and will be discussed later.

Although the Repetition condition generally decreased lexical decision response times, caution must be used in drawing any conclusions. The lexical decision times to the Unrelated Control probes also decreased by roughly 25 ms, which is within the range of variability for this data as indicated by standard error measurements. For this reason the pattern of responses is deemed more important than specific amounts of decrease in lexical decision times. The variability in the data is attributable to the small number of responses per subject, per cell (3 on average). The overall pattern of responses indicates that the effects of Repetition did not change the strength of activation for the probe words, because the overall pattern of priming effects is the same.

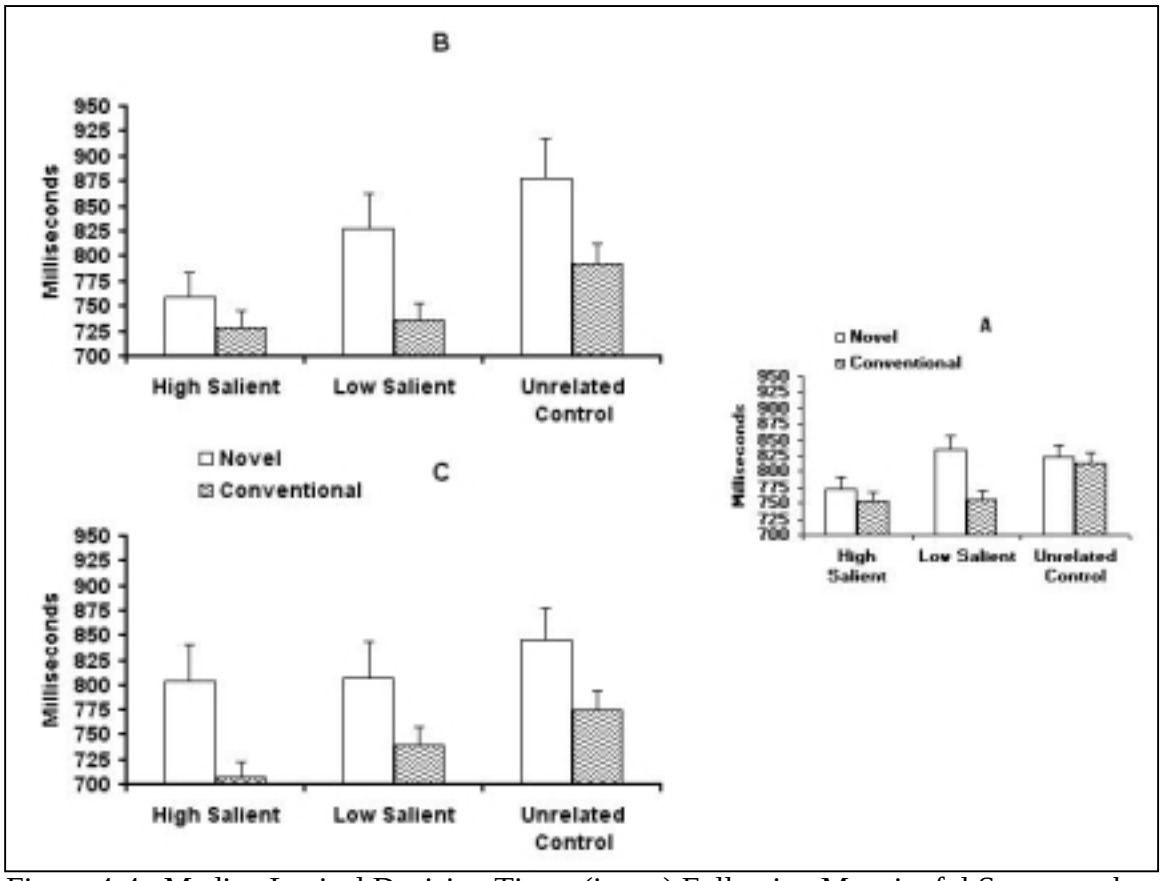

Figure 4-4. Median Lexical Decision Times (in ms) Following Meaningful Sentences by Probe Type for Condition 1 (Repetition). (A) Lexical Decision Times from Experiment 1, (B) Lexical Decision Times for the Not Studied Condition, (C) Lexical Decision Times for the Studied Condition. (Standard Errors are Indicated by Error Bars).

This interpretation is supported by the 3-way, 2 (Study Condition) x 2 (Sentence Conventionality) x 3 (Probe Type) ANOVA conducted on the lexical decision response times in the Repetition condition. Because there were 236 missing cells in this analysis, means for each of the 12 cell types were computed and substituted for the missing cells. The analysis run with the conditional means was nearly identical to an analysis using the grand mean for all conditions. For that reason only the results substituting the conditional means for missing cells are reported here. Main effects for Sentence Conventionality were found, $F(1,90) = 48.92$, $p < .001$ and for Probe Type, $F(2,90) = 20.02$, $p < .001$. Main effects were not found for Study Condition, $F(1,90) = 0.61$, $p < .5$,

nor any interaction effects (Study Condition x Conventionality, $F(1,90) = 0.21$, $p < 1.00$; Studied Condition x Probe Type, $F(2,90) = 1.13$, $p < 0.5$; Conventionality x Probe Type, $F(2,90) = 0.22$, $p <.81$; and Study Condition x Conventionality x Probe Type, $F(2,90) = 2.26$, $p < 0.2$.

Evidence of priming effects in the not studied, repetition condition are provided by paired comparisons between the lexical decision times of: Novel High Salient and Novel Unrelated Control, $t(89) = -4.53$, $p < .001$ (two tailed); Conventional Low Salient and Conventional Unrelated Control, $t(89) = -2.64$, $p < .011$ (two tailed); and Conventional High Salient and Conventional Unrelated Control, $t(89) = -3.27$, $p < .003$ (two tailed). Priming effects in the studied, repetition condition only occur for Conventional High Salient and Conventional Unrelated Control, $t(89) = -3.138$, $p < .003$ (two tailed).

In the Varied Topic condition, a comparison of the pattern of lexical decision responses associated with Not-Studied sentences (Figure 4-5B) and those associated with Studied sentences (Figure 4-5C) reflects the pattern hypothesized at the outset of this experiment. It was hypothesized that a decrease in the lexical decision response times would occur to Low Salient probes that followed novel metaphors. The study condition did in fact result in this predicted decrease. However, it also resulted in an increase in the lexical decision response times to High Salient probes and Unrelated Control probes following novel sentences, while the lexical decision response times to all three probes following conventional sentences remained virtually unchanged.

One possible cause for the unpredicted increase in lexical decision response times to both the High Salient and Unrelated Control probes in the Varied Topic condition is

response competition. Response competition might arise as the result of the multiple mappings involved in the comprehension of the varied novel sentences. This could be construed to be reflective of the conventionalization of the novel metaphors. The same pattern would not be expected for the conventionalized sentences, because they are already conventionalized, hence the multiple mappings do not occur and the response competition does not exist.

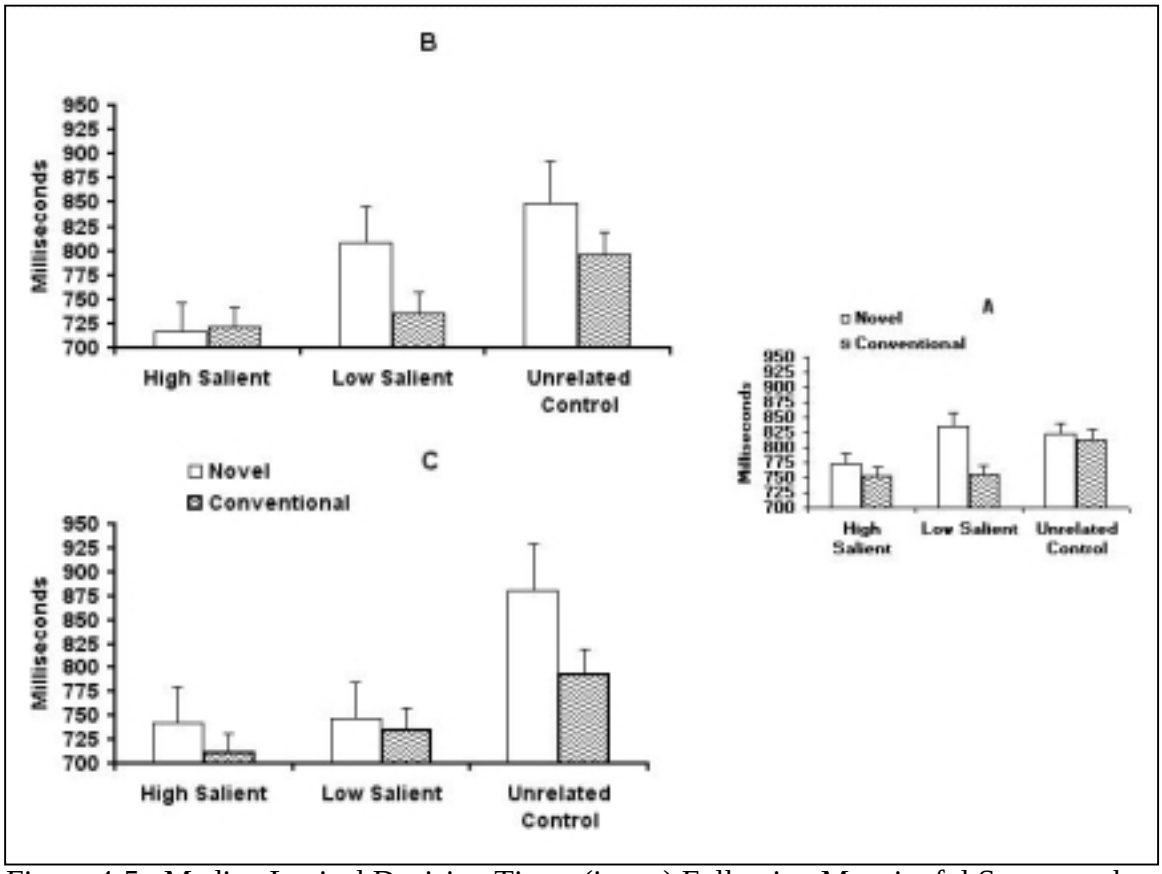

Figure 4-5. Median Lexical Decision Times (in ms) Following Meaningful Sentences by Probe Type for Condition 2 (Varied Topic). (A) Lexical Decision Times from Experiment 1, (B) Lexical Decision Times for the Not Studied Condition, (C) Lexical Decision Times for the Studied Condition. (Standard Errors are Indicated by Error Bars).

The overall pattern of responses in the Varied Topic condition indicates that its effects did cause a change the strength of activation for the Low Salient probe words. One cause of this could be the predicted conventionalization of the novel metaphors.

This interpretation is supported by the 3-way, 2 (Study Condition) x 2 (Sentence Conventionality) x 3 (Probe Type) ANOVA conducted on the lexical decision response times in the Varied Topic condition. Because there were 248 missing cells in this analysis, means for each of the 12 cell types were computed and substituted for the missing cells. The analysis run with the conditional means was nearly identical to an analysis using the grand mean for all conditions. For that reason only the results using the substituted conditional means are reported.

Main effects for Sentence Conventionality were found, $F(1,90) = 8.50$, $p < .005$ and for Probe Type, $F(2,90) = 35.37$, $p < .001$. Main effects were not found for Study Condition, $F(1,90) = 0.07$, $p < 1.0$, nor any interaction effects (Study Condition x Conventionality, $F(1,90) = 0.027$, $p < 1.00$; Studied Condition x Probe Type, $F(2,90) = 1.70$, $p < 0.2$; Conventionality x Probe Type, $F(2,90) = 2.12$; Study Condition x Conventionality x Probe Type), $F(2,90) = 2.68$, $p < 0.1$.

Evidence of priming effects in the not studied, variable topic condition are provided by paired comparisons between the lexical decision times of: Novel High Salient and Novel Unrelated Control, $t(89) = -4.32$, $p < .001$ (two tailed); Conventional Low Salient and Conventional Unrelated Control, $t(89) = -2.67$, $p < .01$ (two tailed); and Conventional High Salient and Conventional Unrelated Control, $t(89) = -3.25$, $p < .003$ (two tailed). These results are almost identical to those found in the not studied, repetition condition.

Priming effects in the studied, varied topic condition are shown in the following four paired comparisons: Novel Low Salient and Novel Unrelated Control, $t(89) = -3.99$, $p < .001$ (two tailed); Conventional Low Salient and Conventional Unrelated Control,

$t(89) = -2.55$, $p < .013$ (two tailed); Novel High Salient and Novel Unrelated Control, $t(89) = -4.60$, $p < .001$ (two tailed); and Conventional High Salient and Conventional Unrelated Control, $t(89) = -3.57$, $p < .002$ (two tailed). These results support the predictions made of the variable topic condition earlier.

In the Context condition, a comparison of the pattern of lexical decision responses associated with Not-Studied sentences (Figure 4-6B) and those associated with Studied sentences (Figure 4-6C) reflects a similar overall pattern to that found in the Varied Topic Condition with a few notable exceptions. The lexical decision response times to all probes following Not-Studied novel and conventional sentences are much longer (50 – 100 ms) in the Context condition than in the Varied Topic condition. Proportionally, the change in lexical decision response times is much less in the Context condition (relative to its own Not-Studied condition). However, in terms of absolute values the amount of change to each of the probes is very similar across sentence conventionality and probe type.

Due to the variability in the data (owing to small cell sizes per participant) it is the position of the author that the relative change in lexical decision response times is a better indicator of conventionalization than directly comparing the magnitude of change that occurred for each probe type. Having said that, the overall pattern of responses in the Context condition indicates that it was the least effective at conventionalizing novel metaphors. A comparison of the lexical decision response times to Low Salient probes following Studied novel sentences in the Context and Repetition conditions shows that the least amount of relative decrease occurred in the Context condition. This is contrary to the hypothesis that the lexical decision response times to Low Salient probes following

novel sentences for the Context condition would be greater than those for the Repetition condition but less than those for the Varied Topic Condition.

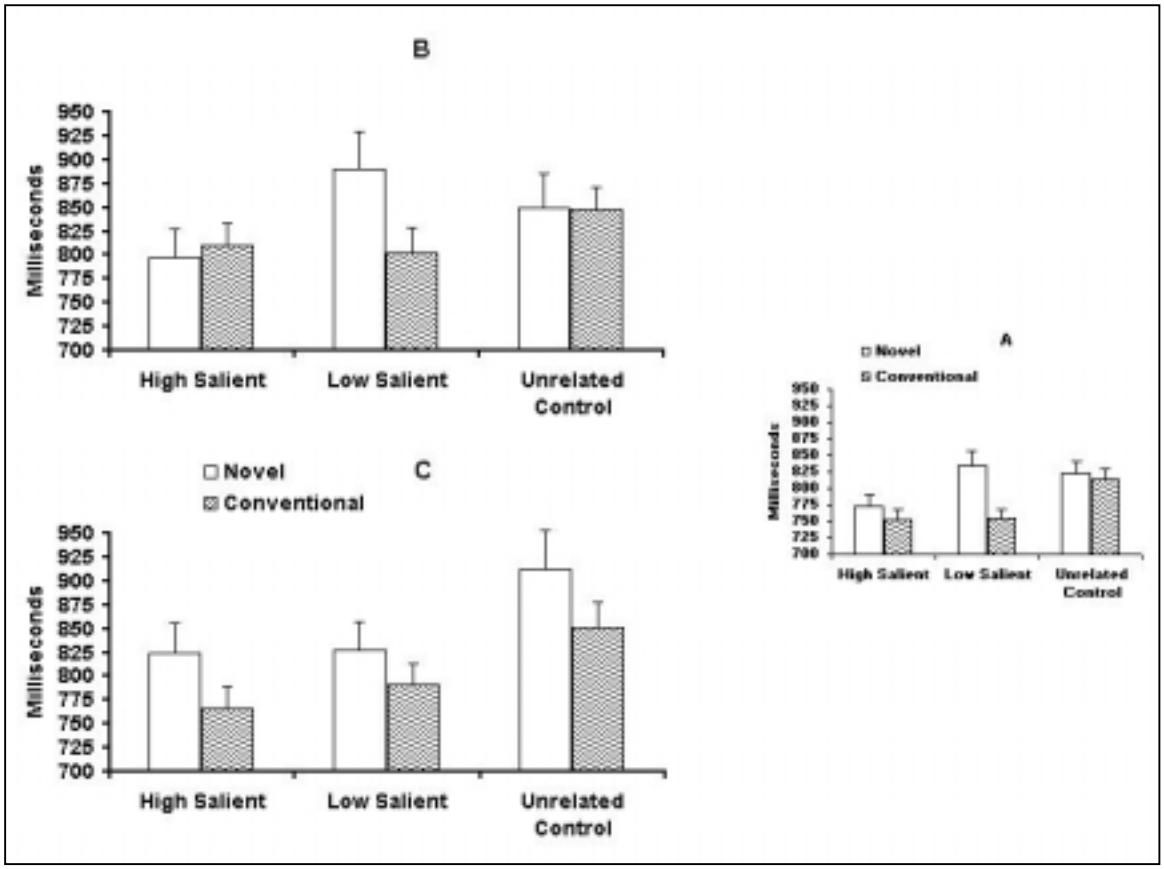

Figure 4-6. Median Lexical Decision Times (in ms) following Meaningful Sentences by Probe Type for Condition 3 (Context). (A) Lexical Decision Times from Experiment 1, (B) Lexical Decision Times for the Not Studied Condition, (C) Lexical Decision Times for the Studied Condition. Standard errors are indicated by error bars.

This interpretation is supported by the 3-way, 2 (Study Condition) x 2 (Sentence Conventionality) x 3 (Probe Type) ANOVA conducted on the lexical decision response times in the Context condition. Because there were 209 missing cells in this analysis, means for each of the 12 cell types were computed and substituted for the missing cells. The analysis run with the conditional means was nearly identical to an analysis using the grand mean for all conditions. For that reason only the results using the substituted conditional means are reported.

Main effects for Sentence Conventionality were found, $F(1,90) = 8.60$, $p < .005$ and for Probe Type, $F(2,90) = 9.88$, $p < .001$. The interaction between Study Condition and Probe Type approached significance, $F(2,90) = 2.90$, $p = .06$. As did the 3-way interaction of Study Condition x Conventionality x Probe Type, $F(2,90) = 2.89$, $p = 0.06$. Neither of these interactions was significant using the more conservative test of substituting missing cells with the grand mean. Main effects were not found for Study Condition, $F(1,90) = 0.097$, $p < 1.0$, nor any other interaction effects (Study Condition x Conventionality, $F(1,90) = 1.24$, $p < .5$; Conventionality x Probe Type, $F(2,90) = 1.20$, $p < .5$.

No evidence of priming effects are found in the not studied, context condition with paired comparisons between lexical decision times. Priming effects are found in the studied, context condition for the following four paired comparisons: Novel Low Salient and Novel Unrelated Control, $t(89) = -2.78$, $p < .008$ (two tailed); Conventional Low Salient and Conventional Unrelated Control, $t(89) = -2.22$, $p < .03$ (two tailed); Novel High Salient and Novel Unrelated Control, $t(89) = -2.87$, $p < .006$ (two tailed); and Conventional High Salient and Conventional Unrelated Control, $t(89) = -2.88$, $p < .006$ (two tailed).

Conventionality Ratings

Appendix G lists the conventionality ratings for each stimulus sentence averaged for: (1) Experiment 1 and all the Not-Studied sentences in each of the three conditions of Experiment 2; (2) Experiment 1; and (3) Each of the Experiment 2 conditions (Repetition, Varied Topic, and Context) independently for the studied and not-studied sentences. While variations between each column is small, the grand mean for the

sentences of Experiment 1 and the Not-Studied sentences should provide the most reliable measure of conventionality ratings, owing to its sample size (*n* = 198).

One way to assess the influence of the learning manipulations (Repetition, Varied Topic, and Context) used in Experiment 2 is to look at the differences in conventionality ratings. By comparing the conventionality ratings between the various conditions we can establish difference scores. Table 4-1 presents a matrix of the difference scores for Conventional metaphors, while Table 4-2 presents a matrix for difference scores for Novel metaphors. In both tables the scores are calculated by subtracting average conventionality ratings for not-studied sentences (the row headings) from the conventionality ratings for the studied sentences (the column headings). Positive scores indicate that there was an increase in sentence conventionality ratings as a result of study. Negative scores would indicate a decrease in conventionality ratings as a result of study.

Table 4-1. Average Difference Scores for the Conventionality Ratings of the Conventional Metaphors used in Experiment 2. (Column Titles minus Row Titles)

	Repetition Studied	Varied Topic Studied	Context Studied
Experiment 1	0.16	0.29	0.06
Experiment 2 Repetition Not Studied	0.04		
Experiment 2 Varied Topic Not Studied		0.18	
Experiment 2 Context Not Studied			0.03
Overall Average (Experiment 1 & 2) for Not Studied	0.12	0.25	0.02

Table 4-1 indicates that each of the three learning conditions had a different effect on conventionality ratings of Conventional metaphors. Repetition barely served to increase conventionality ratings when compared to its own Not-Studied control and only increased conventionality ratings by approximately one tenth of a point when using the conventionality ratings of Experiment 1 or the overall average for Not-Studied sentences in Experiments 1 and 2 as a baseline measure. The Varied Topic condition served to increase conventionality ratings by one fifth of a point when compared to its own Non-Studied control and approximately one quarter of a point when using the conventionality ratings of Experiment 1 or the overall average for Not-Studied sentences in Experiments 1 and 2 as a baseline measure. The Context condition barely changed conventionality ratings regardless of what baseline measure was used.

This observation is supported by the results of a 3 (Condition) x 2 (Sentence Conventionality) ANOVA that was conducted on the conventionality difference scores using the Not-Studied conventionality scores for each condition as its own baseline. The results indicated there is a main effect for Condition, $F(2,60) = 7.91$, $p < .001$. Planned pairwise comparisons using a Bonferroni adjustment for multiple comparisons found significant differences in conventionality difference scores between the Varied Topic and Repetition conditions ($p < .003$). The difference between the conventionality difference scores of the Varied Topic and Context conditions was marginally significant ($p < .011$). Establishing the Varied Topic Condition as the one condition that influenced the conventionalization of novel metaphors.

Table 4-2 indicates that each of the three learning conditions had a different effect on the conventionalization of Novel metaphors, but that two different patterns emerge

depending on whether the same-condition Not-Studied baseline or the Experiment 1/Experiment 1 and 2 averaged baseline measures are used for comparison. Using the Not-Studied controls we see a pattern in the direction of the pre-experimental hypothesis. Namely that the Repetition condition would effect the smallest changes in conventionality ratings, then Context, and that the largest changes would result from the Varied Topic manipulation.

Table 4-2. Average Difference Scores for the Conventionality Ratings of the Novel Metaphors used in Experiment 2. (Column Titles minus Row Titles)

	Repetition Studied	Varied Topic Studied	Context Studied
Experiment 1	0.15	0.16	0.06
Experiment 2 Repetition Not Studied	0.02		
Experiment 2 Varied Topic Not Studied		0.15	
Experiment 2 Context Not Studied			0.07
Overall Average (Experiment 1 & 2) for Not Studied	0.10	0.11	0.01

The amount of change that occurred in the Repetition and Varied Topic conditions is less for Novel metaphors than it was for Conventional metaphors. One possible explanation for the difference in the magnitude of change is that it is easier to further conventionalize, conventional metaphors than it is to conventionalize novel metaphors.

Using the Experiment 1/Experiment 1 and 2 averaged baseline measures we see similar effects on conventionalization (roughly a tenth of a point) by the Repetition and Varied Topic conditions but a minimal increase in conventionality ratings due to Context. The difference in effects depending on what baseline measure is chosen may reflect the sensitivity of conventionalization measures for novel metaphors. A less coarse scale and replication would help to clear this issue.

CHAPTER 5
GENERAL DISCUSSION

The study of metaphor comprehension has important implications for our understanding of the underlying processes of language comprehension and problem solving. This dissertation has made several contributions towards a fuller understanding of metaphor comprehension. The first contribution was the development of a corpus of stimulus material that can be used in future metaphor related studies. This material was developed in response to the criticisms of the wide diversity in existing materials. Both sets of materials (the sentences from Survey 1 and the topic and vehicle terms from Survey 2) and their associated ratings will be available for download via my website at Randolph-Macon College. By collecting ratings on the seven dimensions indicated by the existing literature as being important for metaphor comprehension, more precise manipulations and assessment of those dimensions can occur in future research.

The second contribution was the application of techniques to experimentally differentiate the theoretically distinguishable processes involved in the comprehension of novel and conventional metaphors. This dissertation looked at several techniques that seemed to be candidates for this purpose. Those techniques included: reading times (collected in Surveys 1 and 2), meaningfulness decision times (Experiments 1 and 2), form preference (Experiment 1), lexical decision times (Experiment 1), and direct assessment of vehicle conventionality by the solicitation of ratings (Experiments 1 and 2).

As seen in Survey 1, reading times were significantly different for what was identified post hoc as novel and conventional metaphors. While this may indicate that novel and conventional metaphors are processed differently, it does not necessarily indicate that two dissociable processes are involved. It may reflect the same underlying process, but require more time for the novel metaphor because of lack of familiarity with the comparison being made, necessitating a more exhaustive search of the lexicon for attributes that make the sentence work. Another explanation, which is predicted by the Career of Metaphor Hypothesis, is that the increased time involved in reading novel sentences is due to an alignment and "mapping" process that is not necessary for conventional metaphors. Reading times are not sufficient to indicate which explanation is more accurate.

It was also found in Survey 1 that the pattern of correlations among the various measures of metaphor comprehension was different in several ways between the metaphors that had been rated as novel versus conventional by independent judges. These differential patterns also suggest that the processes of understanding novel and conventional metaphors are qualitatively different, as claimed by the Career of Metaphor Hypothesis.

In both Experiment 1 and Experiment 2, meaningfulness decision times were longer for novel metaphors than for conventional metaphors. Again, while this may reflect a qualitative difference between the two types of metaphor, it does not strongly dissociate them. Lack of familiarity with the sentence may account for the increase in resources to “understand” the sentence. Again the "resource" and "mapping" explanations offered above could both account for this difference in decision times.

Meaningfulness decision times are not sufficient evidence to identify which explanation is more likely.

By adjusting sentence meaningfulness decision times by sentence length in Experiment 1 we were able to partially replicate the form preference reported by Bowdle and Gentner (2001). Their findings were that novel metaphors were read faster when given in the form of a simile (e.g., "Creativity is like a toaster.") than when given in metaphor form (e.g., "Creativity is a toaster"). Gentner and Bowdle (2001) hypothesized and found that the simile form invites comparison and the metaphor form invites categorization resulting in faster reading times for novel metaphors in a simile form compared to those in a metaphor form.

This hypothesis was extended to meaningfulness decision times, which would include reading time and the alignment and mapping process theorized in the Career of Metaphor Hypothesis. While this hypothesis was supported, we failed to fully replicate the associated prediction that meaningfulness decision times would be faster for conventional figuratives in metaphor form than for those figuratives in simile form. Our data, adjusted for sentence length, indicated that conventional figuratives as well as novel figuratives were read faster in simile form.

Stronger evidence for the Career of Metaphor Hypothesis comes from results of the lexical decision tasks and the directly obtained conventionality ratings. Together, these measures point to different processes being involved in the comprehension of novel and conventional metaphors.

First, in both Experiments 1 and 2 (Non-Studied conditions), lexical decisions to Low Salient probes displayed typically less priming, and sometimes inhibition

(Experiment1, and the Context condition of Experiment 2) when following novel metaphors. This contrasted to decision times for Low Salient probes following conventional metaphors, which showed larger and consistent positive priming relative to the unrelated probes.

According to Gentner and Bowdle (2001), the comparison or mapping process associated with the processing of novel metaphors should result in a deeper and broader activation of the semantic and lexical network than the categorical processes associated with conventional metaphors. Yet if this were true then it would seem that Low Salient probes should evoke greater priming following novel metaphors since they are more likely to be partially activated in the processing of a novel metaphor than in the processing of a conventional metaphor. But Gentner and Bowdle (2001) claim that during the "alignment" stage of the comparison process, attributes that do not "fit" are suppressed. Those that remain sufficiently activated are mapped during the "projection' stage to an ad-hoc domain created in working memory. Comprehension occurs when this mapping is complete enough to allow the listener/reader to make sense of the topic-vehicle comparison. From this it follows that Low Salient probe words, which have less dominant or weaker associations with the figurative meanings of the metaphor, are likely to be subject to some suppression.

One possible counter explanation for these results is the claim that the Low Salient probes for the novel sentences were as weakly related to the figurative meaning of the novel sentences as they were to the unrelated control probes (thus the reason for the similar lexical decision response times between them). Of course this potential criticism fails to account for why the lexical decision response times should be equally as fast for

Low Salient probe words as they are for High Salient probe words following conventional figuratives. The benefit of analyzing the data by subjects is that the same sentences were judged by some participants as conventional and by other participants as novel. Those participants saw the same Low Salient probe word, and depending on how they viewed the metaphor they had different lexical decision response patterns. The data showed this effect to be quite robust in all four tests (viz., Experiment 1 and the Not-Studied trials in each of the three learning conditions – Repetition, Varied Topic, and Context).

Having established that lexical decision tasks using High Salient, Low Salient, and Unrelated Control probe words are reliable in differentiating novel and conventional metaphors, we further explored the claim of the Career of Metaphor Hypothesis that conventionalization could be induced through varied mappings of vehicle terms. In Experiment 2, different learning experiences were created in order to explore the effects of repetition, varied topic encoding, and context on conventionality. Overall, despite an impressive reduction in comprehension times with repetition, this familiarization had little effect on conventionalizing novel metaphors, as judged by conventionalization ratings and by the pattern of lexical decision priming. This supported the hypothesis that increased familiarity does not lead to increased conventionalization.

The Varied Topic condition was included to directly test the Career of Metaphor Hypothesis by facilitating various mappings of vehicle concepts. According to the hypothesis, varied experience in trying to comprehend metaphors where the same vehicle appears with different topics should have resulted in the greatest increase in the conventionalization of novel metaphors in which that vehicle again appears. This was

supported both by the increased priming seen for Low Salient probes following novel sentences, which more closely resembled the pattern seen for probes following conventional sentences, and by the slight increase in the conventionality ratings of novel sentences seen in this condition.

Context was included to determine if it served to facilitate, or abbreviate, the mapping process that is hypothesized to occur as a vehicle becomes conventionalized. Theoretically this could be accomplished by enriching the semantic activation that occurs during the mapping process. It was hypothesized that if context was facilitative of this mapping process that participants should award higher conventionalization ratings to novel sentences than participants in the repetition condition. In fact, there was little change in the ratings seen following comprehension with a supportive context. The failure of context to induce changes in conventionality may have resulted because the link between the contextual sentence and the probes may have been too obscure for the participants to implicitly make the connection. But on the other hand, some evidence for conventionalization was seen in the pattern of lexical decision priming, which resembled that of the varied topic Studied condition more than the Not-Studied control condition. It may be that the context was able to set up an ad hoc conceptual mapping of the figurative meaning of the vehicle, and this was adequate to influence priming on the immediately following probe word. But without the sort of practice and reinforcement that was given in the varied topic condition, this mapping existed only in working memory. Since the conventionality ratings were obtained at the end of the session, this ad hoc mapping may have been lost.

One important contribution of this dissertation was the demonstration, in both the Surveys and Experiments, that conventionality ratings appear to be reliable and valid as a direct measure of this construct: (1) There was a great deal of stability in the ratings different participants assigned to the various sentences. This identification was critically important given that two people can interpret the same metaphor very differently based on the their knowledge base and experiences; (2) in both Survey 1 and Experiment 1, sentences rated as conventional versus novel were associated with different patterns of behavior in ways largely predicted by the Career of Metaphor Hypothesis; and (3) changes in conventionality ratings with certain kinds of experience were observed in Experiment 2, again consistent with the overarching hypothesis that novel metaphors are understood via a comparison process, while metaphors with conventional vehicles are understood through a class-inclusion process.

In sum, the experiments in this dissertation have shown that the processes involved in the comprehension of novel and conventional metaphors can be dissociated, and that it is important to consider the conventionality of stimulus material used in metaphor comprehension studies. It has been shown that lexical decision primes may be a useful tool in exploring this difference. We have also shown that even fairly limited, but appropriate, grappling with figurative meanings of novel metaphoric vehicles (the Varied Topic condition) can provide, if not a career, then at least an internship, on the road to conventionalization.

Metaphors change over time both within societies and for individuals as they initially explore and then eventually conventionalize a given metaphor. Understanding metaphor comprehension has important implications for many fields (viz., education,

advertising, journalism, oral communications, etc.). The results of this dissertation can aid in the development of powerful metaphors that enable readers/listeners to make richer mappings between the domains of the topic and vehicle terms. For example knowing that variable topic encoding increases the conventionality of a metaphor, a lecturer might want to use this technique to encourage students to create richer mappings, enabling them to maximize their understanding of the comparison that is being made.

In addition to replicating Experiment 2 with a finer scale of conventionality, other future research the author hopes to engage in is the use of the data from Surveys 1 and 2 to construct a knowledge base for the Structure Mapping Engine (SME). SME is a LISP based computational model of similarity, comparison, and analogy making that is based on Gentner's Structure-Mapping Theory (Gentner, 1983; Gentner & Markman, 1997; Gentner et al., 2001). A prediction of the Structure Mapping Theory is that SME should not be as accurate at predicting literal and conventional metaphors, as it should be at predicting novel metaphors. This is based on the assumption that novel metaphors can be modeled as an extended structural mapping between domains, but that literal statements and conventional metaphors are not (Gentner et al., 2001).

To date there have been very few studies that have employed Event Related Potential (ERP) measures as a means of exploring metaphor comprehension (e.g., Coulson & Van Petten, 2002; Kazmerski, Blasko, & Dessalegn, 2003; and Pynte, Besson, Robichon, & Poli, 1996). Now that it appears there is empirical evidence that supports the notion of differential processing for novel and conventional metaphors, ERP studies using the techniques exploited in this dissertation may provide additional evidence that

separable processes are involved in the comprehension of novel and conventional metaphors.

APPENDIX A
200 SENTENCES USED IN SURVEY 1 (SENTENCE NORMS)

Literal and Figurative Meaning (LF)

A lifetime is a day.
Alligators are problems.
Dad was sore.
Life is a game.
My boss is mad.
My brother is blue.
My cat is an animal.
My companion is a pig.
My instructor is blind.
My nephew is a baby.
My orthodontist is a magician.
My sister is cool.
That fox is a vixen.
That Shriner is a clown.
The accountant is myopic.
The boxer was dancing.
The burden is heavy.
The camera is hot.
The coach is flexible.
The congressman is an angler.
The debutante was a carnivore.
The designer is crafty.
The director is a cannibal
The discovery was a treasure.
The dishwasher was boiling.
The drummer is on fire.
The entrance is tight.
The farmer was henpecked.
The fireworks are bad.
The girls are lambs.
The kid is a goat.
The matron is a mother.
The meal was nutty.
The money is dirty.
The painter stuck to his contract.
The participant is a rat.
The path is a revolution.
The prisoner is a canary.
The punishment was a shock.
The scientist had a dream.
The secretary is cold.
The speaker is a politician.
The sprinter is flying.
The suspect was clean.
The tailor was steaming.
The temple is a landmark.
The waitress is cookin'.
The wallet is light.
The witness was deaf.
Those swine are gluttons.

Literal Only Meaning (LO)

A beet is red.
A bonnet is a hat.
A bowl is a container.
A clock is a timepiece.
A collie is a dog.
A cricket is an insect.
A cyclist is an athlete.
A dagger is a weapon.
A dime is a coin.
A Ford is a car.
A fork is a utensil.
A gorilla is an ape.
A mouse is a mammal.
A pigeon is a bird.
A pinecone is a seed.
A pistol is a gun.
A portrait is a picture.
A potato is a tuber.

A radish is a vegetable.
A refrigerator is an appliance.
A sofa is furniture.
A tent is a shelter.
A wrench is a tool.
An eggplant is purple.
An emerald is a gem.
Baseball is a sport.
Bracelets are jewelry.
Cereal is a breakfast.
Gasoline is a fuel.
Mars is a planet.
Milk is a drink.
My father is a preacher.
Paprika is a spice.
Penicillin is a medicine.
Pianos are instruments.
Pie is a dessert.
Pork is a meat.
Sandals are shoes.
Schematics are drawings.
Syrup is a liquid.
The boat is aluminum.
The jacket was corduroy.
The library is a building.
The maple is a tree.
The moon is bright.
The rain is wet.
The spleen is an organ.
The stove is warm.
The watch is an antique.
Tuesday is a weekday.

Figurative Only Meaning (FO)

A billboard is a wart.
A brain is a warehouse.
A cigar is a skunk.
A dilemma is a cactus.
A forest is a harmonica.
A rumor is a virus.
A storm is a song.
Alcohol is a crutch.
An audition is a door.
Ballerinas are swans.
Children are sponges.
Clouds are puffballs.
Creativity is a toaster.
Deceit is an ambush.
Detectives are ferrets.
Divorce is a nightmare.
Doubt is a net.
Dust is a blanket.
Education is a lantern.
Evolution is a lottery.
Faith is an anchor.
Freedom is truth.
Genes are blueprints.
His bicycle was a rocket.
History is a mirror.
Ideas are diamonds.
Knowledge is a fountain.
London is a beehive.
Love is an antidote.
Marriage is an alloy.
My Chevy is a Cadillac.
My grandmother is a peach.
My manuscript is a rag.
My roommate is a crab.
My surgeon is a butcher.
Rage is a volcano.
That senator is a pawn.
The argument is fishy.
The blackmailer is a leech.
The family is a foundation.
The lawyer was a shark.
The potter is the clay.
The professor is an encyclopedia.
The rationale was opaque.
The sea is a desert.
The stars are a map.
The tongue is a bayonet.
The truck is a lemon.
The words were poison.
Time is a river.

Nonsense (Not Literal nor Figurative) (N)

A bus is a caterpillar.
A dynasty is a play.
A meadow is a harp.
A mushroom is a gnat.
A pizza is a couch.
A rabbit is a yoyo.
A rainbow is a flower.
A starfish is a towel.
A story is a necklace.
A zebra is a xylophone.
Aches are streams.
Adoration is lightning.
An envelope is a jockey.
An island is a cork.
Birthdays are bagels.
Broccoli is a consultant.
Credit is a lubricant.
Discipline is fertilizer.
Dusk is a castle.
Enthusiasm is a fog.
Flight is a crystal.
Genius is an eagle.
Humiliation is a curtain.
Humor is a slave.
Music is death.
My nurse is a pineapple.
Mysticism is a leaf.
Noon is a dolphin.
Popcorn is a casserole.
Scissors are earmuffs.
Silence is an apron.
Sleep is a snowstorm.
Socks are olives.
That popcicle is a girdle.
The army is a plague.
The concert is a gear.
The cucumber is a wheel.
The glass was a cannonball.
The hymn is a turnip.
The nachos were closets.
The puddle is a mango.
The squash is sandpaper.
The taco was an earthquake.
The tractor is a deputy.
The sandwich is a trombone.
The sky is a parliament.
The wind is an armadillo.
Waves are octopi.
Wisdom is a foreigner.
Zucchini is a drill.

APPENDIX B
QUESTIONS ASKED PARTICIPANTS IN SURVEY 1

Comprehensibility:

A sentence can have many meanings. Rate how easy or difficult it is for you to understand the first meaning you thought of when you read the sentence above.

1 = it is very difficult to understand this sentence
2 = it is somewhat difficult to understand this sentence
3 = neutral or not sure
4 = it is somewhat easy to understand this sentence
5 = it is very easy to understand this sentence

Literalness:

Does this sentence have a literal meaning? Rate the sentence on how literal you interpret its meaning to be.

1 = no literal meaning
2 = low literal meaning
3 = neutral or not sure
4 = somewhat literal meaning
5 = very literal meaning

Figurativeness:

Does this sentence have a non-literal or figurative meaning? Rate the sentence on how figurative you interpret its meaning to be.

1 = no figurative meaning
2 = low figurative meaning
3 = neutral or not sure
4 = somewhat figurative meaning
5 = very figurative meaning

Familiarity:

How commonly or frequently have you read or heard the comparison being made between the first noun and the last word in the sentence?

1 = never
2 = uncommon
3 = neutral or not sure
4 = common
5 = very common

Figurative Interpretability:

Rate how easily you 'get the point' of the figurative meaning in the sentence above.

1 = not at all
2 = barely
3 = neutral or not sure
4 = somewhat
5 = very well

Aptness:

How apt is this sentence as a metaphor? In other words rate how good a metaphor this sentence makes.

1 = very bad metaphor
2 = bad metaphor
3 = neutral or not sure
4 = good metaphor
5 = very good metaphor

Similarity:

How similar in meaning are the first noun and the last word in the sentence?

1 = not similar at all
2 = unsimilar in meaning
3 = neutral or not sure
4 = similar in meaning
5 = very similar in meaning

Meaning:

In your own words type up to three distinct meaning(s) of this sentence. Indicate in the box following the sentence if your interpretation is a literal or a figurative one. Remember to type the first meaning that comes to mind in the 'white' box, the second meaning that comes to mind in the 'yellow' box, and the third meaning in the 'blue' box. If the sentence is meaningless, type 'meaningless'.

Attributes:

For your FIGURATIVE interpretations of the sentence, please type AS MANY attributes or qualities (NOT ASSOCIATIONS) of the last word in the sentence that you can. Try to list those attributes of the last word in the sentence that you applied to the first noun in the sentence to make it meaningful. Following each attribute is a number box - use this to assign a ranking of: '1' to the most important, '2' to the next most important, etc.

APPENDIX C
CATEGORIZATION AND AVERAGE RATINGS OF 200 SENTENCES USED IN SURVEY 1

This appendix lists the stimulus sentences and the dimensions they were rated on in Survey 1. Standard deviations for each dimension are noted parenthetically. The following legend is provided for you convenience:

LEGEND

Type — Sentence Type - refers to the category the sentence belongs to based on the literalness and figurativeness ratings averaged across participants.

- LO — Literal only meaning (Literalness > 3.0)
- LF — Literal and Figurative meanings (Literalness and Figurativeness > 3.0)
- FO — Figurative only (Figurativeness > 3.0)
- N — Nonsense (Literal and Figurativeness < 3.0)

Conv — Sentence conventionality - determined by Pilot Conventionality ratings.

Attrib — The number of attributes indicated by participants as being associated with the figurative meaning of the sentence.

Comp — Sentence Comprehensibility

Lit — Literalness

Fig — Figurativeness

FI — Figurative Interpretability

Aptness — How good or apt the metaphor is.

Fam — Familiarity

Sim — Similarity between the topic and vehicle terms

Sentence	Type	Conv	Attrib	Comp	Lit	Fig	FI	Aptness	Fam	Sim
A beet is red.	LO	4.0		4.95 (0.22)	4.80 (0.89)	1.05 (0.22)	4.10 (1.12)	3.70 (1.13)	4.70 (0.57)	2.00 (1.21)
A billboard is a wart.	FO	2.2	63	2.40 (1.57)	1.60 (0.94)	3.65 (1.35)	3.00 (1.41)	2.65 (1.09)	1.10 (0.31)	1.35 (0.59)
A bonnet is a hat.	LO	2.0		5.00 (0.00)	4.80 (0.89)	1.15 (0.67)	4.21 (1.08)	3.37 (1.16)	4.30 (0.92)	2.11 (1.37)
A bowl is a container.	LO	1.4		4.75 (0.91)	4.85 (0.37)	1.35 (0.81)	3.65 (1.63)	3.40 (1.43)	4.25 (1.07)	2.05 (1.39)
A brain is a warehouse.	FO	3.4	51	4.75 (0.44)	2.75 (1.37)	4.50 (0.83)	4.60 (0.68)	4.40 (0.68)	3.50 (1.10)	2.45 (1.28)
A bus is a caterpillar.	FO	1.4	51	3.05 (1.28)	1.55 (0.94)	3.45 (1.43)	3.10 (0.97)	2.75 (0.91)	1.20 (0.41)	1.75 (0.79)
A cigar is a skunk.	FO	3.0	57	3.00 (1.72)	1.55 (1.23)	3.70 (1.69)	4.10 (1.45)	3.55 (1.32)	1.85 (0.88)	1.75 (1.07)
A clock is a timepiece.	LO	2.2		4.75 (0.55)	4.80 (0.52)	1.80 (1.15)	4.00 (1.12)	3.40 (1.23)	3.90 (1.29)	3.05 (1.61)
A collie is a dog.	LO	4.2		4.95 (0.22)	5.00 (0.00)	1.30 (0.92)	4.28 (0.89)	3.83 (0.92)	4.60 (1.17)	2.22 (1.17)
A cricket is an insect.	LO	3.2		5.00 (0.00)	5.00 (0.00)	1.07 (0.22)	3.90 (1.24)	3.55 (1.33)	4.68 (0.75)	2.35 (1.22)
A cyclist is an athlete.	LO	2.0		5.00 (0.00)	4.95 (0.22)	1.30 (0.72)	3.68 (1.37)	3.26 (1.42)	4.65 (0.58)	2.68 (1.65)
A dagger is a weapon.	LO	3.8		4.85 (0.49)	4.90 (0.31)	1.50 (0.95)	3.84 (1.30)	3.68 (1.45)	4.35 (0.67)	2.68 (1.45)
A dilemma is a cactus.	FO	2.4	66	2.90 (1.48)	1.70 (1.13)	4.05 (1.19)	3.45 (1.19)	3.20 (1.28)	1.80 (1.15)	2.05 (0.89)
A dime is a coin.	LO	2.4		5.00 (0.00)	4.90 (0.45)	1.15 (0.67)	3.45 (1.19)	3.30 (1.53)	4.95 (0.22)	2.25 (1.37)
A dynasty is a play.	FO	3.0	23	2.40 (1.43)	2.30 (1.22)	3.15 (1.14)	2.60 (1.39)	2.75 (1.48)	1.45 (0.83)	1.50 (0.69)
A Ford is a car.	LO	2.4		4.95 (0.22)	4.95 (0.22)	1.20 (0.41)	3.20 (1.44)	2.90 (1.25)	4.80 (0.52)	2.50 (1.47)
A forest is a harmonica.	FO	1.2	61	2.85 (1.50)	1.45 (0.89)	3.90 (1.29)	2.95 (1.61)	2.55 (1.50)	1.30 (0.57)	1.35 (0.49)
A fork is a utensil.	LO	1.4		5.00 (0.00)	4.80 (0.89)	1.25 (0.91)	3.90 (1.33)	3.45 (1.47)	4.95 (0.22)	2.40 (1.47)
A gorilla is an ape.	LO	2.8		5.00 (0.00)	4.50 (1.28)	1.60 (1.14)	3.90 (1.21)	3.10 (1.07)	4.70 (0.47)	3.15 (1.60)
A lifetime is a day.	FO	2.4	41	3.80 (1.24)	2.70 (1.56)	4.10 (1.45)	3.75 (1.16)	3.55 (1.05)	2.60 (1.35)	2.05 (1.00)
A meadow is a harp.	FO	1.4	58	2.65 (1.46)	1.80 (1.15)	3.40 (1.39)	2.85 (1.42)	2.95 (1.47)	1.60 (0.99)	2.00 (1.03)
A mouse is a mammal.	LO	1.2		4.40 (1.35)	4.75 (0.91)	1.00 (0.00)	3.26 (1.63)	3.00 (1.29)	4.05 (1.32)	1.95 (1.27)
A mushroom is a gnat.	N	2.2		1.60 (0.88)	1.45 (1.05)	2.55 (1.61)	2.90 (1.62)	2.90 (1.33)	1.20 (0.62)	1.55 (0.89)
A pigeon is a bird.	LO	3.0		5.00 (0.00)	5.00 (0.00)	1.30 (0.66)	3.80 (1.20)	3.15 (1.35)	5.00 (0.00)	2.60 (1.50)
A pinecone is a seed.	LO	3.2		4.60 (0.82)	4.55 (1.10)	1.65 (1.14)	3.45 (1.23)	2.85 (1.27)	4.25 (0.97)	2.75 (1.37)
A pistol is a gun.	LO	2.2		5.00 (0.00)	4.60 (1.23)	1.20 (0.62)	3.10 (1.25)	2.75 (1.29)	4.95 (0.22)	2.10 (1.48)
A pizza is a couch.	N	1.2		1.95 (1.28)	1.35 (0.75)	2.65 (1.50)	2.58 (1.35)	2.26 (1.41)	1.00 (0.00)	1.68 (0.95)
A portrait is a picture.	LO	3.0		4.85 (0.37)	4.95 (0.22)	1.90 (1.41)	4.00 (1.03)	3.10 (1.17)	4.65 (0.59)	3.05 (1.43)
A potato is a tuber.	N	1.0		1.80 (1.40)	2.15 (1.35)	2.25 (1.21)	2.30 (1.30)	2.65 (1.09)	1.50 (1.15)	1.75 (0.91)
A rabbit is a yoyo.	FO	3.6	41	2.75 (1.41)	2.05 (1.43)	3.70 (1.42)	3.10 (1.29)	2.95 (1.19)	1.65 (0.88)	1.80 (1.01)
A radish is a vegetable.	LO	3.6		4.95 (0.22)	4.95 (0.22)	1.55 (1.28)	3.68 (1.29)	3.16 (1.21)	4.40 (0.88)	2.32 (1.53)
A rainbow is a flower.	FO	3.8	49	2.65 (1.39)	1.40 (0.75)	4.20 (0.83)	2.85 (1.31)	2.80 (1.20)	1.60 (0.88)	2.15 (1.09)
A refrigerator is an appliance.	LO	1.6		5.00 (0.00)	4.95 (0.22)	1.30 (0.80)	2.85 (1.65)	2.80 (1.69)	4.80 (0.52)	2.20 (1.28)
A rumor is a virus.	FO	4.4	69	4.20 (1.15)	2.55 (1.57)	4.65 (0.81)	4.55 (1.00)	4.45 (0.76)	3.05 (1.32)	2.50 (1.05)
A sofa is furniture.	LO	2.0		5.00 (0.00)	5.00 (0.00)	1.10 (0.45)	3.60 (1.39)	3.30 (1.22)	4.85 (0.37)	2.70 (1.34)

Sentence	Type	Conv	Attrib	Comp	Lit	Fig	FI	Aptness	Fam	Sim
A starfish is a towel.	N	1.6		2.37 (1.42)	1.89 (1.05)	2.68 (1.49)	2.67 (1.33)	2.06 (1.26)	1.16 (0.50)	1.44 (0.78)
A storm is a song.	FO	3.6	56	3.05 (1.47)	1.75 (1.12)	3.85 (1.46)	3.55 (1.23)	3.35 (1.39)	1.70 (0.92)	2.10 (1.02)
A story is a necklace.	FO	2.2	48	2.50 (1.28)	1.60 (0.94)	4.00 (1.03)	2.85 (1.53)	2.90 (1.45)	1.20 (0.52)	1.30 (0.47)
A tent is a shelter.	LO	3.2		4.70 (0.98)	4.85 (0.67)	1.70 (0.92)	3.00 (1.41)	2.55 (1.19)	4.35 (0.67)	2.50 (1.61)
A wrench is a tool.	LO	3.4		4.95 (0.22)	5.00 (0.00)	1.20 (0.52)	4.05 (1.43)	3.80 (1.40)	4.70 (0.80)	3.15 (1.50)
A zebra is a xylophone.	N	1.0		2.15 (1.42)	1.45 (1.05)	2.85 (1.79)	3.15 (1.46)	2.85 (1.31)	1.40 (0.99)	2.10 (1.17)
Aches are streams.	N	2.6		1.75 (1.02)	1.65 (0.93)	2.65 (1.50)	3.25 (1.62)	3.05 (1.19)	1.05 (0.22)	1.60 (0.75)
Adoration is lightning.	FO	3.4	43	1.85 (0.99)	1.50 (0.89)	3.65 (1.53)	2.55 (1.10)	3.15 (1.24)	1.25 (0.57)	1.55 (0.68)
Alcohol is a crutch.	FO	4.6	48	4.40 (0.82)	2.00 (1.30)	4.50 (0.69)	4.50 (0.69)	4.25 (0.91)	3.40 (1.35)	1.75 (0.91)
Alligators are problems.	LO	2.0		4.50 (0.76)	4.05 (1.23)	2.20 (1.24)	3.68 (1.38)	2.89 (1.24)	3.55 (0.94)	2.05 (1.13)
An audition is a door.	FO	4.2	51	3.25 (1.59)	1.50 (0.95)	4.30 (1.22)	4.20 (1.20)	4.10 (0.91)	2.40 (1.23)	1.90 (1.12)
An eggplant is purple.	LO	1.6		4.85 (0.49)	4.90 (0.45)	1.10 (0.45)	3.33 (1.41)	2.94 (1.51)	4.70 (0.57)	1.78 (0.88)
An emerald is a gem.	LO	4.2		5.00 (0.00)	4.90 (0.45)	2.10 (1.25)	4.11 (1.15)	3.42 (1.12)	4.35 (0.93)	3.26 (1.56)
An envelope is a jockey.	FO	1.8	20	1.65 (1.18)	1.70 (1.08)	3.35 (1.35)	2.15 (1.35)	2.70 (1.49)	1.15 (0.67)	1.50 (0.95)
An island is a cork.	N	2.6		1.65 (1.09)	1.45 (0.83)	2.90 (1.68)	3.15 (1.50)	3.10 (1.25)	1.35 (0.93)	2.15 (1.18)
Ballerinas are swans.	FO	3.2	80	4.45 (1.10)	1.55 (0.89)	4.80 (0.70)	4.55 (1.10)	4.20 (1.24)	3.70 (1.22)	2.75 (1.16)
Baseball is a sport.	LO	2.8		5.00 (0.00)	4.70 (0.92)	1.55 (1.00)	4.00 (1.33)	3.42 (1.02)	4.70 (0.80)	3.05 (1.27)
Birthdays are bagels.	N	1.0		1.50 (1.00)	1.55 (1.15)	2.70 (1.69)	2.90 (1.80)	2.80 (1.58)	1.20 (0.70)	1.75 (1.07)
Bracelets are jewelry.	LO	1.6		5.00 (0.00)	5.00 (0.00)	1.05 (0.22)	3.45 (1.57)	3.15 (1.27)	4.80 (0.52)	1.75 (0.97)
Broccoli is a consultant.	N	1.2		2.05 (1.19)	2.30 (1.49)	2.55 (1.47)	2.37 (1.21)	2.21 (1.32)	1.00 (0.00)	1.58 (0.90)
Cereal is a breakfast.	LO	1.4		4.70 (0.73)	4.45 (1.36)	1.50 (1.00)	3.58 (1.57)	3.11 (1.52)	4.40 (1.10)	2.58 (1.39)
Children are sponges.	FO	4.0	76	4.60 (0.60)	1.50 (0.89)	4.65 (0.81)	4.65 (0.75)	4.35 (0.99)	3.75 (1.52)	1.60 (0.99)
Clouds are puffballs.	LF	1.6	50	4.60 (0.60)	3.15 (1.42)	3.60 (1.47)	4.45 (0.69)	4.20 (0.77)	3.70 (1.13)	3.20 (1.06)
Creativity is a toaster.	FO	1.2	31	2.25 (1.33)	1.45 (1.05)	3.80 (1.54)	2.70 (1.49)	2.60 (1.27)	1.35 (0.52)	1.35 (0.81)
Credit is a lubricant.	FO	2.2	35	3.00 (1.38)	1.85 (1.14)	3.60 (1.50)	3.60 (1.14)	3.20 (1.15)	1.60 (0.82)	1.75 (1.21)
Dad was sore.	LO	2.8		4.50 (0.95)	4.40 (1.05)	2.55 (1.50)	3.55 (1.19)	2.85 (1.14)	3.60 (1.19)	1.40 (0.75)
Deceit is an ambush.	FO	3.2	40	3.40 (1.14)	2.40 (1.39)	3.75 (1.25)	3.40 (1.27)	3.40 (1.10)	1.80 (0.83)	2.60 (1.31)
Detectives are ferrets.	FO	2.8	57	3.25 (1.33)	1.80 (1.24)	3.55 (1.61)	3.26 (1.24)	3.32 (1.06)	1.95 (1.19)	1.89 (0.94)
Discipline is fertilizer.	FO	3.0	51	3.95 (1.15)	1.55 (0.83)	4.45 (0.60)	3.80 (1.24)	3.70 (0.98)	1.85 (1.09)	2.05 (1.28)
Divorce is a nightmare.	FO	4.6	69	4.95 (0.22)	2.95 (1.50)	4.60 (0.88)	4.85 (0.67)	4.65 (0.75)	4.75 (0.44)	2.50 (1.32)
Doubt is a net.	FO	3.8	60	2.80 (1.24)	1.40 (0.60)	4.35 (1.14)	3.45 (1.05)	3.35 (1.19)	1.65 (0.75)	1.95 (1.21)
Dusk is a castle.	FO	3.2	27	1.95 (1.06)	2.40 (1.50)	3.30 (1.18)	2.35 (1.37)	2.55 (1.10)	1.30 (0.46)	1.75 (0.91)
Dust is a blanket.	FO	3.4	58	3.60 (1.27)	2.40 (1.27)	4.25 (0.79)	3.50 (1.28)	3.55 (1.23)	3.05 (1.23)	2.00 (1.12)
Education is a lantern.	FO	3.0	56	3.60 (1.31)	1.45 (1.00)	4.50 (0.76)	4.15 (1.09)	4.10 (0.97)	2.50 (1.40)	2.10 (1.17)
Enthusiasm is a fog.	FO	3.6	42	2.95 (1.28)	1.75 (1.07)	3.55 (1.15)	2.90 (1.17)	2.55 (1.23)	1.80 (1.01)	1.50 (0.89)

Sentence	Type	Conv	Attrib	Comp	Lit	Fig	FI	Aptness	Fam	Sim
Evolution is a lottery.	FO	4.0	39	3.75 (1.29)	2.20 (1.28)	4.35 (0.93)	3.55 (1.36)	3.45 (1.23)	2.05 (1.28)	1.95 (1.15)
Faith is an anchor.	FO	4.0	75	4.25 (0.97)	1.70 (1.08)	4.75 (0.44)	4.60 (0.50)	4.35 (0.75)	3.40 (1.47)	1.95 (1.10)
Flight is a crystal.	FO	2.2	48	1.95 (1.19)	2.00 (1.08)	3.25 (1.33)	2.21 (1.36)	2.63 (1.46)	1.15 (0.67)	1.63 (1.07)
Freedom is truth.	FO	2.8	49	4.20 (0.83)	2.65 (1.46)	4.25 (1.12)	4.00 (1.03)	3.90 (0.79)	3.85 (0.75)	3.00 (1.12)
Gasoline is a fuel.	LO	3.6		5.00 (0.00)	4.95 (0.22)	1.55 (1.28)	3.45 (1.54)	3.40 (1.31)	4.85 (0.49)	2.85 (1.60)
Genes are blueprints.	LF	4.2	40	4.80 (0.41)	3.20 (1.61)	3.45 (1.39)	4.55 (0.76)	4.55 (0.76)	4.70 (0.57)	3.10 (1.41)
Genius is an eagle.	FO	2.8	58	2.85 (1.23)	1.70 (1.08)	3.75 (1.25)	3.20 (1.36)	2.85 (1.14)	1.80 (1.01)	1.85 (1.09)
His bicycle was a rocket.	FO	4.0	59	4.60 (0.75)	2.55 (1.47)	4.60 (0.94)	4.50 (1.00)	4.30 (0.80)	2.85 (1.27)	2.10 (0.85)
History is a mirror.	FO	3.2	44	3.80 (1.01)	1.95 (1.43)	4.50 (0.61)	4.15 (0.81)	4.10 (0.97)	2.70 (1.17)	1.55 (1.00)
Humiliation is a curtain.	FO	3.2	54	3.10 (1.17)	1.80 (1.24)	4.05 (1.00)	3.25 (0.97)	3.25 (1.16)	2.05 (1.10)	1.55 (1.00)
Humor is a slave.	FO	3.6	53	2.75 (1.33)	1.85 (1.23)	4.05 (1.10)	3.05 (1.23)	3.05 (1.00)	1.25 (0.44)	1.35 (0.49)
Ideas are diamonds.	FO	3.2	85	4.10 (1.12)	1.50 (0.95)	4.70 (0.73)	4.30 (1.13)	3.95 (1.23)	2.10 (0.85)	1.75 (0.97)
Knowledge is a fountain.	FO	4.0	84	4.50 (0.61)	2.05 (1.50)	4.60 (0.75)	4.35 (0.75)	4.15 (0.88)	3.95 (1.05)	2.00 (0.97)
Life is a game.	LF	3.8	88	4.55 (0.69)	3.25 (1.41)	4.35 (1.23)	4.55 (0.76)	4.50 (0.61)	4.30 (0.80)	2.15 (0.99)
London is a beehive.	FO	3.6	66	3.70 (1.42)	1.55 (0.69)	4.30 (1.03)	3.85 (1.46)	3.95 (1.28)	2.20 (1.44)	1.65 (0.75)
Love is an antidote.	FO	3.0	45	4.10 (1.17)	2.75 (1.48)	4.30 (0.86)	3.90 (1.17)	4.05 (0.89)	2.90 (1.33)	2.25 (1.16)
Marriage is an alloy.	FO	1.8	44	2.95 (1.47)	1.80 (0.83)	4.00 (1.26)	3.45 (1.43)	3.45 (1.05)	1.70 (1.13)	2.25 (0.91)
Mars is a planet.	LO	1.4		5.00 (0.00)	4.60 (1.23)	1.00 (0.00)	3.65 (1.37)	3.24 (1.25)	4.95 (0.22)	2.29 (1.31)
Milk is a drink.	LO	2.4		4.80 (0.89)	5.00 (0.00)	1.15 (0.49)	3.72 (1.67)	3.56 (1.46)	4.85 (0.67)	2.44 (1.65)
Music is death.	FO	3.2	57	3.00 (1.41)	1.55 (1.10)	4.50 (1.05)	3.15 (1.27)	3.25 (0.85)	1.60 (0.60)	1.45 (0.60)
My boss is mad.	LO	3.6		4.95 (0.22)	4.85 (0.37)	2.95 (1.28)	4.65 (0.75)	3.30 (1.13)	4.05 (1.05)	2.00 (0.73)
My brother is blue.	LF	4.4	68	4.40 (0.88)	3.15 (1.50)	4.40 (1.10)	4.75 (0.72)	4.15 (0.99)	3.85 (1.04)	1.55 (0.83)
My cat is an animal.	LO	3.6		5.00 (0.00)	4.85 (0.49)	2.65 (1.57)	4.25 (0.97)	3.80 (0.95)	4.65 (0.81)	3.30 (1.30)
My Chevy is a Cadillac.	LO	3.2		4.05 (1.32)	3.90 (1.45)	2.50 (1.43)	3.58 (1.39)	2.68 (1.20)	3.20 (1.47)	3.05 (1.31)
My companion is a pig.	LF	4.4	98	4.50 (0.83)	4.10 (0.97)	4.30 (0.86)	4.50 (0.69)	4.25 (0.97)	3.30 (1.03)	2.00 (0.73)
My father is a preacher.	LF	3.2	72	4.85 (0.37)	4.90 (0.31)	4.05 (1.15)	4.63 (0.76)	4.11 (0.74)	3.80 (1.15)	2.95 (1.08)
My grandmother is a peach.	FO	4.0	73	3.85 (1.09)	1.65 (0.99)	4.40 (0.88)	3.85 (1.04)	3.70 (1.03)	2.70 (1.38)	1.70 (0.66)
My instructor is blind.	LF	4.0	55	4.65 (0.93)	4.70 (0.57)	3.90 (1.21)	4.63 (0.76)	3.74 (1.28)	3.10 (1.21)	2.05 (0.78)
My manuscript is a rag.	FO	2.2	68	3.30 (1.13)	2.45 (1.39)	4.10 (0.91)	3.70 (1.08)	3.05 (1.15)	1.90 (1.12)	1.60 (0.68)
My nephew is a baby.	LF	3.8	53	4.90 (0.31)	4.80 (0.52)	3.10 (1.63)	4.20 (1.15)	3.70 (1.13)	4.20 (1.11)	3.10 (1.17)
My nurse is a pineapple.	FO	1.2	58	2.40 (1.47)	1.35 (0.75)	3.65 (1.35)	2.50 (1.43)	2.70 (1.22)	1.40 (0.68)	1.15 (0.49)
My orthodontist is a magician.	LF	3.4	68	4.55 (0.94)	3.10 (1.55)	4.55 (0.83)	4.55 (0.94)	4.20 (0.95)	2.35 (1.35)	1.50 (0.61)
My roommate is a crab.	FO	4.0	66	4.50 (0.95)	2.90 (1.62)	4.85 (0.37)	4.65 (0.59)	4.45 (0.60)	3.55 (1.19)	1.70 (0.73)
My sister is cool.	LF	4.4	59	4.90 (0.31)	4.00 (1.21)	3.75 (1.52)	4.89 (0.46)	3.26 (1.24)	4.25 (1.21)	2.05 (1.22)
My surgeon is a butcher.	LF	4.0	54	4.70 (0.57)	3.40 (1.50)	4.45 (0.94)	4.60 (0.60)	4.35 (0.81)	3.50 (1.10)	3.20 (1.01)

Sentence	Type	Conv	Attrib	Comp	Lit	Fig	FI	Aptness	Fam	Sim
Mysticism is a leaf.	FO	1.8	25	2.05 (1.23)	2.25 (1.45)	3.35 (1.53)	2.40 (1.47)	2.55 (1.50)	1.05 (0.22)	1.25 (0.55)
Noon is a dolphin.	N	1.6		1.95 (1.28)	1.90 (1.45)	2.75 (1.55)	2.68 (1.42)	2.16 (1.17)	1.25 (0.72)	1.47 (0.77)
Paprika is a spice.	LO	3.4		4.90 (0.45)	4.90 (0.45)	1.50 (1.00)	3.25 (1.55)	2.65 (1.27)	4.55 (1.00)	2.65 (1.79)
Penicillin is a medicine.	LO	3.8		4.95 (0.22)	5.00 (0.00)	1.10 (0.31)	4.05 (1.08)	3.79 (1.13)	4.95 (0.22)	2.42 (1.22)
Pianos are instruments.	LO	2.2		5.00 (0.00)	4.95 (0.22)	1.35 (0.93)	3.94 (1.11)	3.44 (1.34)	4.95 (0.22)	2.50 (1.47)
Pie is a dessert.	LO	2.8		4.80 (0.89)	4.60 (1.23)	1.65 (1.27)	3.79 (1.40)	3.26 (1.19)	4.55 (1.10)	2.37 (1.42)
Popcorn is a casserole.	N	1.0		1.70 (1.13)	2.45 (1.67)	1.70 (1.30)	3.75 (1.55)	2.95 (1.36)	1.10 (0.31)	2.00 (1.08)
Pork is a meat.	LO	2.2		5.00 (0.00)	5.00 (0.00)	1.15 (0.49)	2.72 (1.21)	2.72 (1.47)	4.60 (0.94)	2.06 (1.26)
Rage is a volcano.	FO	3.2	75	4.20 (1.01)	1.95 (1.33)	4.45 (1.15)	4.40 (0.88)	4.30 (0.98)	2.80 (1.18)	2.90 (1.47)
Sandals are shoes.	LO	2.0		4.75 (0.79)	5.00 (0.00)	1.20 (0.70)	4.05 (1.28)	3.75 (1.07)	4.45 (1.10)	2.45 (1.32)
Schematics are drawings.	LO	1.6		4.05 (1.05)	4.15 (1.23)	1.90 (1.07)	3.65 (1.14)	3.20 (1.15)	3.35 (1.35)	2.75 (1.21)
Scissors are earmuffs.	FO	1.2	17	1.80 (1.44)	1.10 (0.45)	3.05 (1.70)	2.70 (1.53)	2.15 (1.35)	1.00 (0.00)	1.35 (0.67)
Silence is an apron.	FO	1.6	29	2.60 (1.10)	1.60 (0.94)	3.55 (1.32)	3.00 (1.21)	2.75 (1.12)	1.55 (0.89)	1.65 (0.93)
Sleep is a snowstorm.	FO	2.6	41	2.20 (1.16)	1.45 (0.89)	3.70 (1.44)	2.35 (1.36)	2.60 (1.44)	1.20 (0.66)	1.75 (1.23)
Socks are olives.	FO	1.2	30	1.85 (1.53)	1.30 (0.92)	3.15 (1.66)	2.45 (1.50)	2.20 (1.28)	1.00 (0.00)	1.40 (0.75)
Syrup is a liquid.	LO	2.2		5.00 (0.00)	4.95 (0.22)	1.15 (0.67)	3.32 (1.53)	2.79 (1.65)	4.05 (1.05)	1.58 (0.90)
That fox is a vixen.	LF	3.0	30	4.00 (1.03)	3.35 (1.31)	3.55 (1.47)	3.90 (0.97)	3.70 (0.98)	2.75 (1.29)	2.80 (1.28)
That popcicle is a girdle.	FO	1.6	12	1.60 (1.10)	1.70 (1.17)	3.50 (1.28)	1.85 (1.18)	2.15 (1.18)	1.25 (0.91)	1.65 (0.88)
That senator is a pawn.	FO	4.4	44	3.65 (1.35)	2.30 (1.26)	4.15 (0.93)	3.80 (1.24)	3.95 (0.83)	2.90 (1.41)	2.10 (0.79)
That Shriner is a clown.	LF	4.0	54	3.45 (1.43)	3.60 (1.35)	3.60 (1.19)	3.55 (1.32)	3.45 (1.10)	2.15 (1.14)	2.25 (0.91)
The accountant is myopic.	LO	2.0		2.85 (1.53)	3.90 (1.12)	2.40 (1.05)	3.15 (1.39)	2.75 (1.41)	2.15 (1.14)	2.10 (1.07)
The argument is fishy.	FO	4.6	48	4.35 (0.88)	2.50 (1.50)	4.05 (1.23)	4.45 (0.83)	3.90 (1.21)	3.40 (1.27)	1.80 (0.83)
The army is a plague.	FO	4.0	75	3.45 (0.94)	2.15 (1.39)	4.15 (1.18)	4.25 (0.79)	3.65 (0.99)	2.30 (0.73)	1.75 (0.85)
The blackmailer is a leech.	FO	3.2	70	4.40 (0.94)	1.90 (1.17)	4.70 (0.57)	4.60 (0.60)	4.40 (0.82)	2.90 (1.33)	2.30 (1.17)
The boat is aluminum.	LO	1.0		4.70 (0.73)	4.80 (0.52)	1.70 (1.13)	3.55 (1.23)	3.00 (1.34)	3.40 (1.31)	2.25 (1.45)
The boxer was dancing.	LF	3.0	55	4.65 (0.49)	4.40 (0.94)	3.30 (1.38)	4.20 (1.11)	3.40 (1.19)	3.40 (1.19)	2.05 (1.00)
The burden is heavy.	LF	3.2	58	4.70 (0.47)	3.10 (1.59)	4.05 (1.43)	4.40 (1.14)	4.05 (1.23)	4.45 (0.94)	3.10 (1.48)
The camera is hot.	LF	3.8	66	4.20 (0.95)	3.80 (1.24)	3.80 (1.20)	3.90 (0.97)	3.40 (0.88)	2.80 (1.06)	1.85 (0.93)
The coach is flexible.	LF	3.0	51	4.75 (0.55)	4.50 (0.69)	3.50 (1.36)	4.40 (0.94)	3.45 (1.28)	3.50 (1.19)	2.10 (1.02)
The concert is a gear.	N	2.0		1.10 (0.37)	1.75 (1.14)	2.50 (1.57)	2.50 (1.35)	2.20 (1.26)	1.10 (0.31)	1.80 (1.23)
The congressman is an angler.	LO	1.8		3.00 (1.52)	3.40 (1.67)	3.00 (1.38)	3.30 (1.22)	3.00 (1.08)	1.85 (1.27)	2.15 (0.93)
The cucumber is a wheel.	N	2.6		2.20 (1.24)	1.75 (1.02)	2.65 (1.50)	2.05 (1.05)	2.10 (1.17)	1.10 (0.31)	1.60 (0.82)
The debutante was a carnivore.	LF	2.4	35	3.45 (1.28)	3.55 (1.36)	3.45 (1.36)	3.35 (1.31)	3.05 (1.23)	1.60 (0.99)	1.95 (1.00)
The designer is crafty.	LO	3.6		4.85 (0.49)	4.25 (1.25)	2.65 (1.57)	4.11 (1.41)	3.68 (1.28)	3.90 (1.07)	3.58 (1.33)
The director is a cannibal	LF	2.2	61	3.95 (0.89)	3.15 (1.66)	4.00 (1.30)	4.05 (0.94)	3.55 (1.15)	1.80 (0.83)	1.45 (0.51)

Sentence	Type	Conv	Attrib	Comp	Lit	Fig	FI	Aptness	Fam	Sim
The discovery was a treasure.	LF	3.6	76	4.40 (0.99)	4.60 (0.75)	3.85 (1.39)	4.60 (0.75)	4.05 (1.10)	3.90 (0.97)	3.20 (1.11)
The dishwasher was boiling.	N	3.4		3.30 (1.42)	2.80 (1.51)	2.65 (1.52)	3.00 (1.15)	2.70 (1.16)	1.75 (0.99)	1.70 (1.11)
The drummer is on fire.	LF	4.0	58	4.70 (0.47)	3.60 (1.54)	4.40 (0.99)	4.40 (0.82)	3.95 (1.19)	3.75 (1.29)	1.65 (0.67)
The entrance is tight.	LF	2.6	40	4.65 (0.49)	3.55 (1.47)	3.55 (1.67)	4.11 (1.15)	3.47 (1.02)	3.50 (1.28)	2.00 (1.11)
The family is a foundation.	LF	4.0	83	4.40 (0.82)	3.05 (1.47)	4.30 (1.13)	4.35 (0.88)	4.15 (1.04)	4.00 (1.21)	2.95 (1.28)
The farmer was henpecked.	LF	3.4	23	3.20 (1.40)	3.80 (1.06)	3.35 (1.39)	3.10 (1.33)	3.00 (1.30)	1.95 (1.19)	2.05 (0.89)
The fireworks are bad.	LO	3.0		4.75 (0.44)	4.35 (1.14)	2.75 (1.52)	4.05 (0.91)	2.89 (0.99)	3.10 (1.25)	2.21 (1.27)
The girls are lambs.	FO	3.0	67	3.10 (1.52)	1.95 (1.15)	3.70 (1.34)	3.75 (1.12)	3.40 (1.05)	2.20 (1.32)	2.35 (1.27)
The glass was a cannonball.	FO	2.4	29	1.90 (0.79)	2.00 (1.08)	3.45 (1.10)	2.25 (1.16)	2.40 (1.19)	1.20 (0.52)	1.50 (0.83)
The hymn is a turnip.	FO	1.2	26	1.75 (1.02)	2.00 (1.26)	3.05 (1.50)	2.50 (1.36)	2.60 (1.39)	1.25 (0.79)	1.90 (1.07)
The jacket was corduroy.	LO	1.0		4.95 (0.22)	4.85 (0.49)	1.50 (0.89)	3.15 (1.57)	2.85 (1.31)	4.20 (1.11)	2.20 (1.28)
The kid is a goat.	LF	2.8	43	3.60 (1.27)	3.35 (1.69)	3.45 (1.61)	3.20 (1.28)	2.75 (1.25)	2.40 (1.39)	2.55 (1.64)
The lawyer was a shark.	FO	4.2	73	4.60 (0.75)	1.95 (1.43)	4.75 (0.44)	4.65 (0.49)	4.60 (0.50)	3.70 (1.17)	2.20 (1.20)
The library is a building.	LO	2.0		5.00 (0.00)	5.00 (0.00)	1.20 (0.52)	3.63 (0.96)	3.53 (0.90)	4.75 (0.44)	2.00 (1.11)
The maple is a tree.	LO	2.4		4.75 (0.91)	4.65 (0.99)	1.40 (0.75)	3.95 (1.23)	3.40 (1.19)	4.50 (0.95)	2.75 (1.48)
The matron is a mother.	LO	2.8		3.70 (1.03)	3.95 (1.10)	2.45 (1.10)	3.05 (1.10)	3.10 (0.97)	2.75 (1.29)	3.15 (1.04)
The meal was nutty.	LF	3.6	48	3.95 (1.39)	3.95 (1.10)	3.15 (1.46)	3.60 (1.27)	2.70 (1.22)	2.10 (0.91)	2.20 (0.83)
The money is dirty.	LF	3.8	58	4.70 (0.47)	4.25 (1.07)	4.25 (1.33)	4.50 (1.15)	4.30 (0.98)	4.05 (0.83)	1.75 (0.97)
The moon is bright.	LO	3.8		5.00 (0.00)	4.95 (0.22)	1.85 (0.99)	3.40 (1.60)	2.70 (1.45)	4.70 (0.47)	1.95 (0.94)
The nachos were closets.	N	1.8		1.60 (1.27)	1.20 (0.52)	2.95 (1.70)	2.95 (1.57)	2.30 (1.42)	1.00 (0.00)	1.30 (0.57)
The painter stuck to his contract.	LF	3.2	47	4.70 (0.73)	4.00 (1.30)	3.25 (1.68)	4.20 (1.15)	3.45 (1.32)	3.90 (1.12)	2.00 (1.12)
The participant is a rat.	LF	3.6	81	4.65 (0.81)	4.05 (1.50)	4.10 (1.29)	4.45 (1.10)	4.15 (1.09)	3.90 (1.07)	2.05 (1.23)
The path is a revolution.	FO	3.2	51	3.80 (0.89)	2.35 (1.39)	4.20 (1.01)	3.80 (0.77)	3.70 (0.73)	2.70 (0.98)	1.80 (0.77)
The potter is the clay.	FO	2.4	49	2.30 (1.08)	1.95 (1.05)	3.70 (1.30)	2.95 (1.00)	2.95 (1.15)	2.00 (1.17)	2.95 (1.05)
The prisoner is a canary.	LF	2.4	52	3.55 (1.54)	3.30 (1.53)	3.95 (1.19)	3.80 (1.20)	3.35 (1.09)	2.55 (1.19)	2.05 (1.00)
The professor is an encyclopedia.	FO	4.0	68	4.20 (1.15)	2.35 (1.57)	4.65 (0.75)	4.65 (0.67)	4.60 (0.50)	3.60 (0.99)	2.95 (0.89)
The puddle is a mango.	N	1.2		1.80 (0.95)	2.00 (1.21)	2.70 (1.42)	2.35 (1.35)	2.15 (1.14)	1.10 (0.31)	1.50 (0.83)
The punishment was a shock.	LF	3.2	38	4.95 (0.22)	4.70 (0.71)	3.45 (1.60)	4.11 (1.12)	3.63 (1.22)	3.95 (1.02)	2.95 (1.36)
The rain is wet.	LO	2.6		4.75 (0.91)	4.90 (0.45)	1.70 (1.22)	4.21 (1.08)	3.47 (1.39)	4.35 (1.09)	2.95 (1.65)
The rationale was opaque.	FO	4.0	39	2.60 (1.31)	2.50 (1.47)	3.45 (1.36)	3.05 (1.32)	3.35 (1.09)	1.70 (0.80)	1.75 (0.72)
The sandwich is a trombone.	FO	1.0	11	2.00 (1.34)	1.50 (1.10)	3.20 (1.64)	2.80 (1.64)	2.70 (1.45)	1.05 (0.22)	1.10 (0.31)
The scientist had a dream.	LO	3.8		4.90 (0.31)	4.85 (0.49)	2.40 (1.57)	3.89 (1.33)	2.84 (1.26)	3.05 (1.50)	1.58 (0.84)
The sea is a desert.	FO	3.4	72	3.80 (1.20)	1.85 (0.99)	3.95 (1.19)	3.75 (1.25)	3.80 (0.95)	2.40 (1.10)	1.70 (0.86)
The secretary is cold.	LO	4.0		4.50 (0.76)	4.10 (1.12)	1.24 (0.87)	4.35 (0.81)	3.80 (1.15)	2.95 (1.10)	1.65 (0.88)
The sky is a parliament.	FO	1.4	10	2.00 (1.26)	1.95 (1.23)	3.25 (1.62)	2.50 (1.47)	2.60 (1.35)	1.30 (0.80)	2.15 (1.35)

Sentence	Type	Conv	Attrib	Comp	Lit	Fig	FI	Aptness	Fam	Sim
The speaker is a politician.	LO	2.8		4.90 (0.31)	4.75 (0.44)	2.55 (1.54)	4.00 (1.17)	3.90 (1.12)	3.60 (1.47)	2.60 (1.19)
The spleen is an organ.	LO	1.8		4.75 (0.91)	4.70 (0.80)	1.65 (1.27)	3.67 (1.14)	2.83 (1.04)	4.20 (1.32)	1.83 (1.15)
The sprinter is flying.	FO	3.6	68	4.90 (0.31)	2.70 (1.59)	4.40 (1.14)	4.80 (0.41)	4.55 (0.69)	4.15 (1.18)	2.65 (1.27)
The squash is sandpaper.	FO	2.6	54	2.05 (1.47)	2.20 (1.44)	3.70 (1.38)	2.40 (1.43)	2.65 (1.35)	1.10 (0.31)	1.50 (0.83)
The stars are a map.	LF	3.8	41	4.10 (1.02)	3.30 (1.53)	3.60 (1.26)	3.90 (0.76)	3.40 (1.15)	3.25 (1.38)	2.30 (1.00)
The stove is warm.	LO	3.2		4.95 (0.22)	4.75 (0.91)	1.10 (0.31)	3.39 (1.42)	3.28 (1.27)	4.90 (0.31)	2.00 (1.08)
The suspect was clean.	LF	3.2	59	4.70 (0.57)	4.30 (0.92)	3.85 (1.18)	4.55 (0.69)	3.70 (1.26)	4.50 (0.61)	1.90 (1.07)
The taco was an earthquake.	FO	2.2	55	2.25 (1.41)	1.85 (1.39)	4.05 (1.05)	2.75 (1.55)	2.25 (1.07)	1.10 (0.31)	1.20 (0.52)
The tailor was steaming.	LF	3.6	57	3.70 (1.38)	3.30 (1.38)	3.95 (1.15)	4.25 (0.97)	3.65 (1.23)	3.30 (1.26)	2.30 (1.26)
The temple is a landmark.	LO	3.4		4.70 (0.73)	4.60 (0.75)	2.00 (1.21)	3.63 (1.61)	3.26 (1.19)	4.00 (0.79)	2.32 (1.29)
The tongue is a bayonet.	FO	2.0	53	3.35 (1.39)	2.05 (0.94)	4.15 (0.99)	3.55 (1.23)	3.85 (1.04)	1.90 (1.12)	2.25 (0.91)
The tractor is a deputy.	FO	1.6	27	1.50 (0.61)	1.40 (1.00)	3.30 (1.80)	2.45 (1.43)	2.55 (1.47)	1.05 (0.22)	1.65 (0.92)
The truck is a lemon.	FO	4.4	49	3.65 (1.73)	1.45 (0.83)	4.40 (1.10)	3.65 (1.73)	3.45 (1.67)	3.50 (1.76)	1.30 (0.57)
The waitress is cookin'.	LF	3.2	48	4.55 (0.76)	4.35 (1.23)	3.55 (1.50)	4.45 (0.94)	3.25 (1.21)	2.80 (1.36)	2.15 (1.18)
The wallet is light.	LF	2.8	38	4.80 (0.41)	4.45 (0.94)	3.05 (1.64)	3.75 (1.29)	2.90 (1.25)	3.65 (0.93)	1.45 (0.60)
The watch is an antique.	LO	3.4		4.95 (0.22)	4.90 (0.31)	1.95 (1.28)	4.55 (0.94)	4.00 (1.34)	4.30 (0.98)	2.45 (1.28)
The wind is an armadillo.	FO	1.0	23	1.70 (0.92)	1.45 (0.89)	3.10 (1.68)	2.16 (1.21)	2.53 (1.31)	1.00 (0.00)	1.11 (0.32)
The witness was deaf.	LO	3.0		4.60 (1.10)	4.90 (0.31)	2.20 (1.44)	3.70 (1.22)	3.45 (1.10)	2.85 (1.31)	1.95 (0.89)
The words were poison.	FO	3.6	97	4.70 (0.47)	1.70 (1.17)	4.85 (0.37)	4.95 (0.22)	4.50 (0.69)	3.85 (0.88)	1.50 (1.00)
Those swine are gluttons.	LF	2.2	34	3.20 (1.47)	3.65 (1.31)	3.15 (1.35)	3.55 (1.43)	3.80 (1.01)	2.55 (1.57)	3.35 (1.04)
Time is a river.	FO	3.4	70	4.40 (0.68)	1.65 (1.09)	4.70 (0.73)	4.45 (0.60)	4.50 (0.61)	3.50 (1.10)	2.10 (1.02)
Tuesday is a weekday.	LO	1.0		4.95 (0.22)	5.00 (0.00)	1.00 (0.00)	4.05 (0.91)	3.42 (1.17)	4.75 (0.91)	2.11 (0.94)
Waves are octopi.	N	1.6		1.80 (1.15)	1.85 (1.18)	2.80 (1.32)	2.50 (1.50)	2.30 (1.03)	1.00 (0.00)	1.95 (1.00)
Wisdom is a foreigner.	FO	3.0	37	2.90 (1.21)	2.30 (1.30)	4.15 (1.09)	3.35 (1.23)	3.45 (1.10)	1.95 (0.94)	1.80 (0.70)
Zucchini is a drill.	FO	2.2	25	1.80 (0.95)	1.80 (1.11)	3.15 (1.50)	2.35 (1.39)	2.15 (1.23)	1.05 (0.22)	1.25 (0.55)

APPENDIX D
400 WORDS USED AS STIMULUS MATERIAL IN SURVEY 2

accountant
aches
adoration
alcohol
alligators
alloy
aluminum
ambush
anchor
angler
animal
antidote
antique
ape
appliance
apron
argument
armadillo
army
athlete
audition
baby
bad
bagel
ballerinas
baseball
bayonet
beehive
beet
bicycle
billboard
bird
birthday
blackmailer
blanket
blind
blue
blueprints
boat
boiling
bonnet
boss
bowl
boxer
bracelets
brain
breakfast
bright
broccoli
brother
building
burden
bus
butcher
cactus
Cadillac
camera
canary
cannibal
cannonball
car
carnivore
casserole
castle
cat
caterpillar
cereal
Chevy
children
cigar
clay
clean
clock
closets
clouds
clown
coach
coin
cold
collie
companion
concert
consultant
container
contract
cookin'
cool
corduroy
cork
couch
crab
crafty
creativity
credit
cricket
crutch
crystal
cucumber
curtain
cyclist
Dad
dagger
dancing
day
deaf
death
debutante
deceit
deputy
desert
designer
dessert
detectives
diamonds
dilemma
dime
director
dirty
discipline
discovery
dishwasher
divorce
dog
dolphin
door
doubt
drawings
dream
drill
drink
drummer
dusk
dust
dynasty
eagle
earmuffs
earthquake
education
eggplant
emerald
encyclopedia
enthusiasm
entrance
envelope
evolution
faith
family
farmer
father
ferrets
fertilizer
fire
fireworks
fishy
flexible
flight
flower
flying
fog
Ford
foreigner
forest
fork
foundation
fountain
fox
freedom
fuel
furniture
game
gasoline
gear
gem
genes
genius
girdle
girls
glass
gluttons
gnat
goat
gorilla
grandmother
gun
harmonica
harp
hat
heavy
henpecked
history

hot
humiliation
humor
hymn
ideas
insect
instructor
instruments
island
jacket
jewelry
jockey
kid
knowledge
lambs
landmark
lantern
lawyer
leaf
leech
lemon
library
Life
lifetime
light
lightening
liquid
London
lottery
love
lubricant
mad
magician
mammal
mango
manuscript
map
maple
marriage
Mars
matron
meadow
meal
meat
medicine
milk
mirror
money
moon
mother
mouse
mushroom
music
myopic
mysticism
nachos
necklace
net
nightmare
noon
nurse
nutty
octopi
olives
opaque
organ
orthodontist
painter
paprika
parliament
participant
path
pawn
peach
penicillin
piano
picture
pie
pig
pigeon
pineapple
pinecone
pistol
pizza
plague
planet
play
poison
politician
popcicle
popcorn
pork
portrait
potato
potter
preacher
prisoner
problem
professor
puddle
puffballs
punishment
purple
rabbit
radish
rag
rage
rain
rainbow
rat
rationale
red
refrigerator
revolution
river
rocket
roommate
rumor
sandals
sandpaper
sandwich
schematics
scientist
scissors
sea
secretary
seed
senator
shark
shelter
shock
shoes
Shriner
silence
sister
skunk
sky
slave
sleep
snowstorm
socks
sofa
song
sore
speaker
spice
spleen
sponges
sport
sprinter
squash
starfish
stars
steaming
stock
storm
story
stove
streams
stuck
surgeon
suspect
swan
swine
syrup
taco
tailor
temple
tent
tight
time
timepiece
toaster
tongue
tool
towel
tractor
treasure
tree
trombone
truck
truth
tuber
Tuesday
turnip
utensil
vegetable
virus
vixen
volcano
waitress
wallet
warehouse
warm
wart
watch
waves
weapon
weekday
wet
wheel
wind
wisdom
witness
words
wrench
xylophone
yoyo
zebra
zucchini

APPENDIX E
QUESTIONS/INSTRUCTIONS ASKED OF PARTICIPANTS IN SURVEY 2

Rate the word above in terms of its abstractness/concreteness

1 = very abstract
2 = somewhat abstract
3 = neutral or not sure
4 = somewhat concrete
5 = very concrete

How commonly or frequently have you read or heard the word above?

1 = I've never seen this word before
2 = somewhat infrequently
3 = moderately or not sure
4 = somewhat frequently
5 = very frequently

Please type at least three qualities or attributes (NOT ASSOCIATIONS) that are related to the displayed word in the boxes that are provided.

Next to each box is a number – use this to assign a rating of:
"1" to the most important,
"2" to the next most important, etc.

APPENDIX F

96 SENTENCES AND ATTRIBUTE PROBES USED IN EXPERIMENT 1

Novel Sentences	**High Salient**	**Low Salient**	**Unrelated**
A billboard is a wart.	misplaced	nuisance	betrayed
A bus is a caterpillar.	creeping	inches	alluring
A dilemma is a cactus.	painful	bothersome	looming
A forest is a harmonica.	noise	unity	crafty
A lifetime is a day.	short	existence	elegant
A meadow is a harp.	tranquil	content	reveal
A story is a necklace.	sequence	connected	offensive
An envelope is a jockey.	deliver	quick	mature
Clouds are puffballs.	shapes	wispy	guide
Creativity is a toaster.	thought	activity	persistent
Credit is a lubricant.	ease	manage	bright
Flight is a crystal.	exciting	worth	complex
Marriage is an alloy.	combine	durable	poised
My manuscript is a rag.	messy	horrible	electricity
My nurse is a pineapple.	sour	callous	bathes
Mysticism is a leaf.	healthy	new	keen
Scissors are earmuffs.	comfort	alter	chew
Silence is an apron.	protect	barrier	disappear
Sleep is a snowstorm.	forget	overpowering	impressive
Socks are olives.	dirty	insulated	drama
That popcicle is a girdle.	slimmer	smaller	geometry
The debutante was a carnivore.	hungry	spiteful	silver
The director is a cannibal.	scary	unfeeling	potato
The glass was a cannonball.	crashing	piercing	innocent
The hymn is a turnip.	tasteful	unprepared	precision
The potter is the clay.	potential	overworked	straining
The prisoner is a canary.	caged	nervous	wooden
The sandwich is a trombone.	huge	complicated	levitate
The sky is a parliament.	under	override	furry
The squash is sandpaper.	bumpy	grainy	quantity
The taco was an earthquake.	churning	wild	foggy
The tongue is a bayonet.	harsh	injure	stiff
The tractor is a deputy.	badge	work	zap
The wind is an armadillo.	calm	steady	drain
Those swine are gluttons.	obese	rotund	dodge
Zucchini is a drill.	routine	rut	empty

Conventional Sentences	High Salient	Low Salient	Unrelated
A rabbit is a yoyo.	tricks	playful	monster
A rainbow is a flower.	beautiful	awesome	intimidating
A rumor is a virus.	invisible	vicious	knowledge
A storm is a song.	rhythm	clamor	nurture
Alcohol is a crutch.	support	stumble	valuable
An audition is a door.	beginning	confidence	melancholy
Children are sponges.	absorb	learn	vision
Divorce is a nightmare.	distress	hopeless	tiresome
Doubt is a net.	trap	dangerous	repeat
Enthusiasm is a fog.	opaque	blinding	skill
Evolution is a lottery.	luck	uncertain	annoying
Faith is an anchor.	unmoving	unwavering	untrustable
Genes are blueprints.	layout	grid	slick
His bicycle was a rocket.	explore	roaring	potion
Humor is a slave.	control	dependent	effort
Knowledge is a fountain.	continuous	journey	emotional
Life is a game.	goal	pleasant	unexpected
London is a beehive.	busy	hectic	admire
My companion is a pig.	sloppy	gross	remedy
My grandmother is a peach.	kind	warm	dark
My orthodontist is a magician.	incredible	respected	democracy
My roommate is a crab.	mean	unhappy	fresh
My surgeon is a butcher.	careless	rough	upset
That senator is a pawn.	insignificant	weak	hanging
That Shriner is a clown.	crazy	laughing	bottle
The army is a plague.	destructive	unhealthy	combustion
The discovery was a treasure.	important	extraordinary	production
The family is a foundation.	strong	backbone	predict
The lawyer was a shark.	ruthless	stealthy	weird
The participant is a rat.	manipulate	selfish	triangular
The professor is an encyclopedia.	brilliant	old	canvas
The sea is a desert.	endless	space	glide
The stars are a map.	guidance	lost	idiotic
The truck is a lemon.	bad	broken	naïve
The words were poison.	hurtful	unpleasant	unknown
Time is a river.	slow	moving	rude

Filler Sentences	**Nonword**
A cigar is a skunk.	yelms
A dynasty is a play.	fiunds
Ballerinas are swans.	pinture
Deceit is an ambush.	coggert
Detectives are ferrets.	slexible
Discipline is fertilizer.	metion
Dusk is a castle.	browth
Education is a lantern.	gretch
Freedom is truth.	smelk
Genius is an eagle.	galued
History is a mirror.	vising
Humiliation is a curtain.	sneldorm
Ideas are diamonds.	welcode
Love is an antidote.	warmted
Music is death.	bledac
Rage is a volcano.	scallow
That fox is a vixen.	blarking
The blackmailer is a leech.	strast
The boxer was dancing.	predicle
The coach is flexible.	applisant
The girls are lambs.	minolant
The kid is a goat.	anairy
The wallet is light.	leacro
Wisdom is a foreigner.	experent

APPENDIX G
CONVENTIONALITY RATINGS

This appendix lists the stimulus sentences and the conventionality ratings they received in Experiments 1 and 2. Standard deviations are noted parenthetically. The following legend is provided for you convenience:

LEGEND

Type	Sentence Type - refers to the category the sentence belongs to based on the conventionality ratings averaged across participants in Experiment 1 and the three not-studied conditions of Experiment 2. N Novel - ratings less than 3.0 C Conventional - ratings greater than 3.0
Convent	Grand Mean of sentence conventionality ratings averaged across participants in Experiment 1 and the three not-studied conditions of Experiment 2.
Exp 1	Conventionality ratings averaged across participants in Experiment 1.
Rep S	Conventionality ratings of studied sentences averaged across participants in condition 1 (Repetition) of Experiment 2.
Rep NS	Conventionality ratings of not-studied sentences averaged across participants in condition 1 (Repetition) of Experiment 2.
Varied S	Conventionality ratings of studied sentences averaged across participants in condition 2 (Varied Topics) of Experiment 2.
Varied NS	Conventionality ratings of not-studied sentences averaged across participants in condition 2 (Varied Topics) of Experiment 2.
Context S	Conventionality ratings of studied sentences averaged across participants in condition 3 (Context) of Experiment 2.
Context NS	Conventionality ratings of not-studied sentences averaged across participants in condition 3 (Context) of Experiment 2.

Conventionality Ratings for Novel Metaphors used in Experiment 2

Sentence	**Type**	**Convent**	**Exp 1**	**Rep S**	**Rep NS**	**Varied S**	**Varied NS**	**Context S**	**Context NS**
The sandwich is a trombone.	1.42	N	1.40 (0.66)	1.56 (0.89)	1.44 (0.72)	1.36 (0.68)	1.47 (0.79)	1.29 (0.59)	1.38 (0.75)
Socks are olives.	1.44	N	1.29 (0.58)	1.47 (0.81)	1.67 (1.07)	1.47 (0.81)	1.40 (0.69)	1.42 (0.69)	1.47 (0.94)
The wind is an armadillo.	1.44	N	1.43 (0.82)	1.44 (0.72)	1.53 (0.81)	1.36 (0.68)	1.42 (0.87)	1.40 (0.84)	1.40 (0.91)
My nurse is a pineapple.	1.50	N	1.35 (0.68)	1.64 (0.98)	1.49 (0.87)	1.64 (1.13)	1.58 (0.94)	1.42 (0.89)	1.64 (1.00)
The hymn is a turnip.	1.52	N	1.49 (0.86)	1.53 (0.92)	1.67 (0.98)	1.49 (0.87)	1.47 (0.81)	1.49 (0.79)	1.44 (0.87)
A forest is a harmonica.	1.64	N	1.65 (0.85)	1.58 (1.01)	1.64 (0.98)	1.78 (1.11)	1.51 (0.84)	1.67 (0.83)	1.76 (0.98)
Scissors are earmuffs.	1.65	N	1.56 (0.86)	1.49 (0.87)	1.82 (1.07)	1.89 (0.98)	1.67 (1.00)	1.44 (0.78)	1.58 (0.89)
Creativity is a toaster.	1.68	N	1.67 (0.95)	1.67 (0.88)	1.84 (1.02)	1.62 (0.94)	1.76 (0.96)	1.82 (1.09)	1.47 (0.73)
The sky is a parliament.	1.82	N	1.62 (0.85)	2.07 (1.19)	1.89 (1.03)	1.91 (1.00)	1.89 (1.03)	1.56 (0.84)	1.98 (1.01)
A story is a necklace.	1.88	N	1.78 (1.02)	2.36 (1.17)	2.09 (1.16)	1.84 (1.04)	1.89 (1.07)	1.98 (0.92)	1.80 (0.97)
That popcicle is a girdle.	1.90	N	2.10 (0.95)	1.67 (0.88)	1.80 (0.87)	1.69 (0.92)	1.87 (0.94)	2.00 (1.04)	1.78 (0.90)
The tractor is a deputy.	1.93	N	1.87 (0.94)	1.73 (0.96)	1.93 (1.03)	1.98 (1.01)	2.02 (1.01)	1.87 (1.08)	1.93 (0.94)
Silence is an apron.	2.04	N	2.11 (1.19)	2.47 (1.12)	1.98 (0.99)	2.62 (1.13)	2.00 (1.02)	2.44 (1.27)	2.04 (1.11)
Mysticism is a leaf.	2.09	N	1.94 (0.93)	2.31 (1.22)	2.44 (1.20)	1.98 (1.03)	2.04 (1.09)	2.13 (1.06)	2.00 (1.07)
A bus is a caterpillar.	2.14	N	2.05 (1.18)	2.38 (1.19)	2.47 (1.27)	2.31 (1.16)	2.07 (1.14)	2.38 (1.13)	2.00 (1.07)
The tongue is a bayonet.	2.27	N	2.17 (1.26)	2.80 (1.29)	2.49 (1.20)	2.67 (1.22)	2.33 (1.15)	2.36 (1.21)	2.13 (1.25)
The glass was a cannonball.	2.31	N	2.22 (1.07)	2.07 (1.03)	2.33 (0.95)	3.07 (1.21)	2.29 (1.14)	2.36 (1.21)	2.44 (1.16)
Clouds are puffballs.	2.33	N	2.22 (1.18)	2.16 (1.13)	2.27 (1.16)	2.53 (1.39)	2.47 (1.18)	2.31 (1.14)	2.40 (1.29)
A lifetime is a day.	2.35	N	2.22 (1.30)	2.84 (1.30)	2.64 (1.15)	2.49 (1.14)	2.27 (1.29)	2.58 (1.16)	2.33 (1.11)
A meadow is a harp.	2.37	N	2.44 (1.19)	2.27 (1.25)	2.47 (1.24)	2.51 (1.20)	2.27 (1.12)	2.31 (1.06)	2.29 (1.06)
Marriage is an alloy.	2.44	N	2.30 (1.04)	2.78 (1.22)	2.40 (1.16)	2.56 (1.12)	2.64 (1.00)	2.56 (1.25)	2.47 (0.99)
A dilemma is a cactus.	2.49	N	2.43 (1.16)	2.67 (1.17)	2.71 (1.12)	2.44 (1.29)	2.62 (1.35)	2.16 (1.13)	2.22 (1.17)
A billboard is a wart.	2.51	N	2.45 (1.05)	2.40 (1.10)	2.69 (1.22)	2.91 (1.18)	2.53 (1.18)	2.67 (1.22)	2.38 (1.13)
Sleep is a snowstorm.	2.51	N	2.65 (1.21)	2.69 (1.14)	2.62 (1.07)	2.76 (1.19)	2.24 (1.28)	2.42 (1.03)	2.47 (1.20)
The prisoner is a canary.	2.54	N	2.48 (1.16)	3.07 (1.12)	2.64 (1.17)	3.11 (1.28)	2.60 (1.32)	2.62 (1.23)	2.44 (1.27)
Those swine are gluttons.	2.56	N	2.56 (1.34)	2.60 (1.30)	2.49 (1.10)	2.76 (1.35)	2.76 (1.26)	2.53 (1.34)	2.44 (1.20)
Credit is a lubricant.	2.63	N	2.57 (1.23)	2.96 (1.19)	2.62 (1.05)	2.89 (1.35)	2.69 (1.16)	2.84 (1.19)	2.64 (1.26)
The taco was an earthquake.	2.76	N	2.67 (1.32)	2.71 (1.01)	3.09 (1.00)	2.84 (1.19)	2.71 (1.10)	2.71 (1.08)	2.60 (1.01)
The potter is the clay.	2.79	N	2.81 (1.28)	2.98 (1.12)	2.89 (1.27)	2.71 (1.18)	2.71 (1.22)	2.82 (1.25)	2.73 (1.05)
London is a beehive.	2.94	N	2.97 (1.36)	3.29 (1.16)	3.24 (1.09)	3.58 (1.16)	2.96 (1.31)	2.71 (1.20)	2.60 (1.16)

Conventionality Ratings for Conventional Metaphors used in Experiment 2

Sentence	Type	Convent	Exp 1	Rep S	Rep NS	Varied S	Varied NS	Context S	Context NS
My grandmother is a peach.	3.30	C	3.44 (1.12)	3.60 (1.05)	3.40 (1.14)	3.42 (1.16)	3.20 (1.22)	3.16 (1.24)	3.09 (1.35)
The stars are a map.	3.34	C	3.33 (1.18)	3.47 (0.99)	3.47 (1.06)	3.67 (0.90)	3.40 (1.10)	3.07 (1.30)	3.16 (1.17)
Enthusiasm is a fog.	3.38	C	3.24 (1.19)	3.51 (0.99)	3.49 (0.92)	3.67 (0.98)	3.38 (0.96)	3.20 (0.92)	3.47 (1.20)
An audition is a door.	3.44	C	3.22 (1.33)	3.51 (1.27)	3.53 (1.22)	3.76 (0.98)	3.47 (1.22)	3.51 (1.20)	3.64 (1.19)
Genes are blueprints.	3.47	C	3.13 (1.35)	3.38 (1.09)	3.62 (1.07)	3.93 (0.96)	3.49 (1.08)	3.49 (1.18)	3.80 (0.89)
The professor is an encyclopedia.	3.52	C	3.38 (1.26)	3.62 (0.94)	3.56 (1.18)	4.11 (0.91)	3.71 (0.92)	3.29 (1.20)	3.47 (1.16)
The sea is a desert.	3.54	C	3.60 (1.02)	3.67 (1.00)	3.71 (0.84)	3.60 (1.19)	3.33 (1.09)	3.84 (0.88)	3.47 (1.22)
Time is a river.	3.62	C	3.73 (1.00)	3.87 (0.84)	3.58 (1.08)	3.47 (1.24)	3.53 (0.94)	3.64 (0.96)	3.60 (1.21)
His bicycle was a rocket.	3.64	C	3.65 (1.06)	3.78 (0.90)	3.58 (0.97)	4.24 (0.71)	3.67 (0.83)	3.64 (1.09)	3.67 (0.85)
My orthodontist is a magician.	3.65	C	3.60 (1.16)	3.76 (1.05)	3.67 (1.17)	3.93 (1.05)	3.64 (0.98)	3.71 (0.94)	3.69 (1.06)
Knowledge is a fountain.	3.71	C	3.59 (0.98)	4.09 (0.76)	3.80 (0.89)	3.56 (1.06)	3.73 (1.10)	3.80 (1.06)	3.76 (0.86)
The truck is a lemon.	3.75	C	3.84 (0.97)	3.67 (1.15)	3.73 (1.32)	3.82 (1.42)	3.82 (1.17)	3.87 (1.01)	3.56 (1.24)
Doubt is a net.	3.77	C	3.87 (0.99)	3.51 (0.97)	3.69 (0.90)	3.49 (1.08)	3.73 (0.89)	3.58 (0.94)	3.73 (0.99)
Children are sponges.	3.81	C	3.84 (1.07)	4.09 (0.90)	3.73 (0.96)	4.07 (0.99)	3.80 (1.22)	4.07 (1.01)	3.84 (1.07)
My surgeon is a butcher.	3.84	C	3.84 (1.14)	3.82 (1.07)	3.89 (0.96)	4.18 (0.75)	4.00 (0.90)	3.84 (0.80)	3.62 (1.07)
The army is a plague.	3.86	C	3.86 (1.03)	4.00 (0.83)	3.82 (1.07)	4.04 (0.95)	4.11 (0.78)	3.91 (0.90)	3.67 (0.88)
My roommate is a crab.	3.91	C	3.95 (0.96)	3.78 (1.15)	4.04 (1.02)	4.40 (0.75)	3.82 (1.27)	3.80 (1.10)	3.80 (1.06)
Alcohol is a crutch.	3.91	C	3.86 (0.95)	4.27 (0.65)	4.00 (0.98)	3.91 (0.90)	3.84 (0.95)	3.73 (1.07)	3.96 (0.90)
That senator is a pawn.	3.92	C	4.03 (1.05)	3.73 (1.12)	3.78 (1.22)	4.09 (0.92)	3.93 (1.05)	3.96 (0.98)	3.91 (1.12)
Faith is an anchor.	3.94	C	3.89 (0.97)	4.09 (0.85)	4.02 (0.92)	4.16 (0.85)	3.91 (0.82)	4.09 (0.87)	3.98 (0.89)
The family is a foundation.	3.95	C	3.79 (1.12)	4.00 (0.90)	3.98 (0.89)	4.22 (0.77)	4.00 (1.00)	3.76 (1.07)	4.11 (0.86)
Life is a game.	3.96	C	3.81 (0.91)	4.29 (0.82)	4.18 (0.86)	4.24 (0.86)	4.02 (0.92)	4.18 (0.89)	3.91 (1.04)
The lawyer was a shark.	3.97	C	3.83 (1.17)	4.22 (0.82)	4.11 (0.71)	4.09 (0.95)	3.87 (1.04)	3.98 (0.89)	4.16 (1.02)
Divorce is a nightmare.	4.03	C	4.00 (0.93)	4.33 (0.67)	3.98 (0.92)	4.27 (0.86)	4.22 (1.00)	4.18 (0.91)	3.93 (1.14)
That Shriner is a clown.	4.04	C	3.83 (1.20)	3.96 (0.93)	4.16 (0.90)	4.22 (0.90)	3.98 (0.87)	3.98 (1.12)	4.27 (0.75)
The words were poison.	4.10	C	4.21 (0.81)	4.22 (0.64)	4.13 (0.76)	4.29 (0.66)	4.16 (0.95)	4.29 (0.73)	3.84 (0.88)
The discovery was a treasure.	4.10	C	4.03 (0.78)	4.07 (0.81)	4.16 (0.74)	4.40 (0.65)	4.09 (0.67)	4.24 (0.68)	4.13 (0.76)
A rumor is a virus.	4.14	C	4.11 (0.79)	4.13 (0.81)	4.02 (0.72)	4.18 (0.91)	4.22 (0.70)	4.22 (0.77)	4.22 (0.67)
The participant is a rat.	4.28	C	4.17 (1.01)	4.33 (0.90)	4.29 (0.87)	4.42 (0.69)	4.49 (0.55)	4.33 (0.85)	4.20 (0.99)
My companion is a pig.	4.31	C	4.30 (0.91)	4.31 (0.85)	4.38 (0.53)	4.53 (0.59)	4.27 (0.91)	4.53 (0.89)	4.31 (0.97)

LIST OF REFERENCES

Blank, G. D. (1988). Metaphors in the lexicon. *Metaphor and Symbolic Activity, 3,* 21-36.

Blasko, D. G., & Briihl, D. S. (1997). Reading and recall of metaphorical sentences: Effects of familiarity and context. *Metaphor & Symbol, 12,* 261-285.

Blasko, D. G., & Connine, C. (1993). Effects of familiarity and aptness on metaphor processing. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 19,* 295-308.

Bortfeld, H., & McGlone, M. S. (2001). The continuum of metaphor processing. *Metaphor and Symbol, 16,* 75-86.

Bowdle, B., & Gentner, D. (1995). The career of metaphor. Paper presented at the meeting of the Psychonomics Society, Los Angeles, CA.

Bowdle, B., & Gentner, D. (1999). Metaphor comprehension: From comparison to categorization. *Proceedings of the twenty-first annual conference of the Cognitive Science Society.* Hillsdale, NJ: Erlbaum.

Brown, James I., Fishco, V. V., & Hanna, G. (1993). *Nelson-Denny Reading Test, Form H.* (Available from Riverside Publishing, 425 Spring Lake Drive, Itasca, IL 60143-2079).

Cacciari, C., & Tabossi, P. (1988). The comprehension of idioms. *Journal of Memory & Language, 27,* 668-683.

Chiappe, D. L., & Kennedy, J. M. (2001). Literal bases for metaphor and simile. *Metaphor and Symbol, 16,* 249-276.

Connine, C. M., Blasko, D., & Wang, J. (1994). Vertical similarity in spoken word recognition: Multiple lexical activation, individual differences, and the role of sentence context. *Perception & Psychophysics, 56,* 624-636.

Coulson, S., & Van Petten, C. (2002). Conceptual integration and metaphor: An event-related potential study. *Memory & Cognition, 30,* 958-968.

Cronk, B. C., Lima, S. D., & Schweigert, W. A. (1993). Idioms in sentences: Effects of frequency, literalness, and familiarity. *Journal of Psycholinguistic Research, 22,* 59-82.

Dukes, W. E., & Bevan, W. (1967). Stimulus variation and recall: The role of belongingness. *American Journal of Psychology, 80,* 309-312.

Falkenhainer, B., Forbus, K. D., & Gentner, D. (1989). The structure-mapping engine: Algorithm and examples. *Artificial Intelligence, 41,* 1-63.

Galinski, A. D., & Glucksberg, S. (2000). Inhibition of the literal: Metaphors and idioms as judgmental primes. *Social Cognition, 18,* 35-54.

Gentner, D. (1983) Structure-mapping: A theoretical framework for analogy. *Cognitive Science, 7,* 155-170.

Gentner, D., & Bowdle, B. F. (2001). Convention, form, and figurative language processing. *Metaphor and Symbol, 16,* 23-247.

Gentner, D., & Wolff, P. (1997). Alignment in the processing of metaphor. *Journal of Memory and Language, 37,* 331-355.

Gernsbacher, M. A., Keysar, B., & Robertson, R. W. (1995, November). *The role of suppression in metaphor interpretation.* Paper presented at the annual meeting of the Psychonomic Society, Los Angeles.

Gernsbacher, M. A., Keysar, B., Robertson, R. R. W., & Werner, N. K. (2001). The role of suppression and enhancement in understanding metaphors. *Journal of Memory and Language, 45,* 433-450.

Gerrig, R. J., & Healy, A. F. (1983). Dual processes in metaphor understanding: Comprehension and Appreciation. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 9,* 667-675.

Gibbs, R. W., Jr. (1994a). Figurative thought and figurative language. In M. A. Gernsbacher (Ed.), *Handbook of psycholinguistics.* (pp. 411-446). San Diego: Academic Press.

Gibbs, R. W., Jr. (1994b). *The poetics of mind: Figurative thought, language and understanding*. Chapter 5: Understanding metaphorical expressions. Cambridge, England: Cambridge University Press.

Gibbs, R. W., Jr., Bogdanovich, J. M., Sykes, J. R., & Barr, D. J. (1997). Metaphor in idiom comprehension. *Journal of Memory and Language, 37,* 141-154.

Gildea, P. & Glucksberg, S. (1983). On understanding metaphor: The role context. *Journal of Verbal Learning and Verbal Behavior, 22,* 577-590.

Giora, R. (1997). Understanding figurative and literal language: The graded salience hypothesis. *Cognitive Linguistics, 8,* 183-206.

Giora R., & Fein, O. (1999). On understanding familiar and less-familiar figurative language. *Journal of Pragmatics, 31,* 1601-1618.

Glucksberg, S. (2001). *Understanding figurative language: From metaphors to idioms*. New York: Oxford University Press.

Glucksberg, S., Gildea, P., & Bookin, H. B. (1982). On understanding nonliteral speech: Can people ignore metaphors? *Journal of Verbal Learning and Verbal Behavior, 21,* 85-98.

Glucksberg, S., & Keysar, B. (1990). Understanding metaphorical comparisons: Beyond similarity. *Psychological Review, 97,* 3-18.

Glucksberg, S., & Keysar, B. (1993). How metaphors work. In A. Ortony (Ed.), *Metaphor and thought* (2nd ed. pp. 401-424). New York: Cambridge University Press.

Glucksberg, S., McGlone, M. S., & Manfredi, D. A. (1997). Property attribution in metaphor comprehension. *Journal of Memory and Language, 36,* 50-67.

Glucksberg, S., Newsome, M. R., & Goldvarg, Y. (2001). Inhibition of the Literal: Filtering Metaphor-Irrelevant Information During Metaphor Comprehension. *Metaphor and Symbol, 16,* 277-293.

Goldvarg, Y., & Glucksberg, S. (1998). Conceptual combinations: The role of similarity. *Metaphor and Symbol, 13,* 243-255.

Katz, A., Paivio, A., Marschark, M., & Clark, J. M. (1988). Norms for 204 literary and 260 nonliterary metaphors on 10 psychological dimensions. *Metaphor and Symbolic Activity, 3.* 191-214.

Kazmerski, V., Blasko, D., & Dessalegn, B. (2003). ERP and behavioral evidence of individual differences in metaphor comprehension. *Memory and Cognition, 31,* 673-689.

Kemper, S. (1989). Priming the comprehension of metaphors. *Metaphor and Symbolic Activity, 4.* 1-17.

Keysar, B. (1989). On the functional equivalence of literal and metaphorical interpretations in discourse. *Journal of Memory and Language, 28,* 375-385.

Lakoff, G., & Johnson, M. (1980). *Metaphors we live by.* Chicago: University of Chicago Press.

Lakoff, G., & Turner, M. (1989) *More than a cool reason: A field guide to poetic metaphor*. Chicago: University of Chicago Press.

Martin, J. M. (1992). Computer understanding of conventional metaphoric language. *Cognitive Linguistics, 1,* 207-246.

McKay, M. T. (2003) *Understanding metaphor comprehension: Theoretical models and empirical findings*. Unpublished manuscript, University of Florida, Gainesville.

Medin, D. L., Goldstone, R. L., & Gentner, D. (1990). Similarity involving attributes and relations: Judgments of similarity and difference are not inverses. *Psychological Review, 100*, 254-278.

Ortony, A. (1979). Beyond literal similarity. *Psychological Review, 86,* 161-180.

Peleg, O., Giora, R., & Fein, O. (2001). Salience and context effects: Two are better than one. *Metaphor and Symbol, 16*, 173-192.

Pynte, J., Besson, M., Robichon, F-H., & Poli, J. (1996). The time-course of metaphor comprehension: An event-related potential study. *Brain and Language, 55,* 293-316.

Ratcliff, R. (1993). Methods for dealing with reaction time outliers. *Psychological Bulletin, 114*, 510-532.

Richards, I. A. (1936). *The philosophy of rhetoric*, New York: Oxford University.

Rips, L. J. (1975). Quantification and semantic memory. *Cognitive Psychology, 7,* 307-340.

Tourangeau, R., & Rips, L. (1991). Interpreting and evaluating metaphors. *Journal of Memory and Language, 30,* 452-472.

Tourangeau, R., & Sternberg, R. J., (1981). Aptness in metaphor. *Cognitive Psychology, 13,* 27-55.

Tourangeau, R., & Sternberg, R. J., (1982). Understanding and appreciating metaphors. *Cognition, 11,* 203-244.

Turner, N. E., & Katz, A. N. (1997). The availability of conventional and of literal meaning during the comprehension of proverbs. *Pragmatics and Cognition, 5,* 199-233.

Tversky, A. (1977). Features of similarity. *Psychological Review, 84*(4), 327-352.

Verbrugge, R. R., & McCarrell, N. S. (1977). Metaphoric comprehension: Studies in reminding and resembling. *Cognitive Science, 9,* 494-533.

Vu, H., Kellas, G., & Paul, S. T. (1998). Sources of sentence constraint in lexical ambiguity resolution. *Memory & Cognition, 26,* 979-1001.

Wolff, P., & Gentner, D. (1992). The time course of metaphor comprehension. *Proceedings of the fourteenth annual conference of the Cognitive Science Society,* pp. 504-509.

Wolff, P., & Gentner, D. (2000). Evidence for role-neutral initial processing of metaphors. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 26,* 529-541.

BIOGRAPHICAL SKETCH

Michael T. McKay was born in Denver, Colorado, on January 9, 1959. He attended high school at Lakewood High School and graduated in 1977. He enlisted in the Naval Reserve in June 1977, completed Basic Training at RTC San Diego in August 1977, and completed Radioman "A" School at Naval Training Center, San Diego, in September 1977.

His first duty assignment was at the Naval Communications Area Master Station (NAVCAMS MED) in Naples, Italy. There he was recognized as “Seaman of the Year” in 1979. Upon completion of his tour and his three-year active duty obligation in June 1980 he transferred to the Naval Reserves and was assigned to Commander Naval Surface Forces, Pacific – Detachment 118 (COMNAVSURFPAC DET 118) in Denver, Colorado.

After serving in the Naval Reserve for three years, getting married, and going broke (in that order), he reenlisted in the regular Navy and headed back to Naples, Italy, in 1983 to Commander, Maritime Surveillance and Reconnaissance Forces, Sixth Fleet (COMARSURVRECFORSIXTHFLT). While stationed there, he was co-assigned to the staffs of Commander Task Force Sixty Seven (CTF 67) where he served as the Leading Petty Officer in the operational Control Center. Additionally, he was co-assigned to the staff of Commander, Fleet, Air, and Mediterranean (COMFAIRMED) where he served as the Communications Deputy Assistant Chief of Staff. His most significant accomplishments included selection as the COMFAIRMED/CTF 67 combined Staff

Sailor of the Year in 1986 and completion of a Bachelor of Science degree in behavioral science from the University of La Verne in May 1987.

In May 1987, he reported on board the Sixth Fleet flagship, the USS BELKNAP (CG 26). There he worked in the communications center and was advanced to Chief Petty Officer in September 1989 and earned his Enlisted Surface Warfare Specialist (ESWS) designation prior to transferring.

In June 1990, he reported to the Naval Education and Training, Program Management Support Activity (NETPMSA), where he served as the "Radioman" rating exam writer. While there he completed a Master of Arts degree in psychology from the University of West Florida in December 1993.

In July 1993 he transferred to Naval Computer and Telecommunications Station (NAVCOMTELSTA), Puerto Rico. There he held a variety of positions including Puerto Rico were he was assigned as: Space and Electronics Warfare (SEW) Division Officer, Command Training Officer, Command Career Counselor, Communications Department Leading Chief Petty Officer, AIS/ Customer Service Division Officer, and Command Master Chief. While stationed in Puerto Rico he was promoted to both Senior and Master Chief Radioman and worked as an adjunct psychology instructor for Columbia College for four years.

In October 1996, he transferred to the Information Systems Officer School in Newport Rhode Island. There he taught entry-level communications courses to Ensigns who were destined to become shipboard Communications Officers. Additionally, he provided instruction to Department Heads, Prospective Executive and Commanding Officers, Chaplains, and JAG officers. While there he earned designation as a Master

Training Specialist. He retired with 23 years of Naval service on September 1, 2000, and is authorized to wear the Navy Commendation Medal (5), Navy Achievement Medal (2), Naval Unit Commendation (2), Meritorious Unit Commendation (2), Battle Efficiency (2), Good Conduct (4), National Defense Service Medal, Sea Service Deployment Ribbon (3), Overseas Service Ribbon (11), and the Rifle Marksman Ribbon.

In August 2000 he began graduate school in the Cognitive and Sensory Processes area of the Department of Psychology at the University of Florida. He was awarded the Allyn & Bacon/Longman Outstanding Graduate Student Teaching Award for 2002. In April 2004 he was awarded the E. Porter Horne Memorial Scholarship for outstanding graduate student pursuing the study of Sensory Processes, Perception and/or Cognitive Psychology and the Henry C. and Audrey S. Shumacher Fellowship Fund recognizing outstanding doctoral dissertation work in the department of psychology.

Michael has accepted a position as Assistant Professor of Human Psychophysiology at Randolph-Macon College in Ashland, Virginia. He is married to his always loving, highly supportive, and extremely tolerant wife Cathy. They have two incredible sons: Jerome, age twenty-five, and Sean, age twenty-one.