HOW METAPHOR SHAPES THOUGHT

A DISSERTATION

SUBMITTED TO THE DEPARTMENT OF PSYCHOLOGY

AND THE COMMITTEE ON GRADUATE STUDIES

OF STANFORD UNIVERSITY

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS

FOR THE DEGREE OF

DOCTOR OF PHILOSOPHY

Paul H. Thibodeau

August 2012

This dissertation is online at: http://purl.stanford.edu/vh836qp4802

I certify that I have read this dissertation and that, in my opinion, it is fully adequate in scope and quality as a dissertation for the degree of Doctor of Philosophy.

Lera Boroditsky, Primary Adviser

I certify that I have read this dissertation and that, in my opinion, it is fully adequate in scope and quality as a dissertation for the degree of Doctor of Philosophy.

James McClelland

I certify that I have read this dissertation and that, in my opinion, it is fully adequate in scope and quality as a dissertation for the degree of Doctor of Philosophy.

Samuel McClure

Approved for the Stanford University Committee on Graduate Studies.

Patricia J. Gumport, Vice Provost Graduate Education

This signature page was generated electronically upon submission of this dissertation in electronic format. An original signed hard copy of the signature page is on file in University Archives.

Abstract

The way we talk about complex and abstract ideas is suffused with metaphor. In a series of behavioral experiments and computational simulations, I explore how these metaphors influence the way that we reason about complex issues. I find that even the subtlest instantiation of a metaphor (via a single word) can have a powerful influence on how people attempt to solve social problems like crime and how they gather information to make ‘‘well-informed’’ decisions. Interestingly, I find that the influence of the metaphorical framing effect is covert: people do not recognize metaphors as influential in their decisions; instead they point to more ‘‘substantive’’ (often numerical) information as the motivation for their problem-solving decision. I also investigate the underlying mechanisms and contextual factors that support and influence how this process unfolds. Metaphors in language appear to instantiate frame-consistent knowledge structures and invite structurally consistent inferences. Far from being mere rhetorical flourishes, metaphors have profound influences on how we conceptualize and act with respect to important societal issues.

Acknowledgments

Graduate school is a journey and mine has been illuminated by many guides and fellow travelers. I sincerely appreciate all of the support they have given me along the way. This dissertation would not have been possible without their help.

I have had the pleasure of working with a number of inspiring mentors and collaborators at Stanford. I am particularly indebted to my advisor, Lera Boroditsky, who graciously accommodated my intellectual (and geographic) wanderings over the past five years, steering me through obstacles, and posing intellectual challenges. I would not be the scientist I am today without the mentorship that Lera provided.

I am grateful to Jay McClelland for his continued support and guidance throughout my graduate career. I owe many thanks to my committee for their questions and advice: Eve Clark, Herb Clark, and Sam McClure.

Ewart Thomas' wisdom and smile have been a strong positive force in my graduate experience. He has always been available to discuss ideas and to give feedback and perspective. Benoit Monin, Michael Ramscar, and Sam McClure have been mentors and friends.

Frank Durgin, advisor, mentor, friend, warrants pages of acknowledgments. Without him, I certainly would not be where I am. Dan Grodner and Les Sikos have been welcoming and engaging colleagues and collaborators at Swarthmore.

One of the best parts about being a graduate student at Stanford is being surrounded by interesting and amicable graduate students and postdocs. I became close friends and developed productive collaborations with many of my peers and look forward to continuing these relationships for years to come.

Caitlin Fausey helped to make sure this dissertation got written by reading and giving feedback on multiple drafts at various stages of the thinking and writing process. For that especially, I am very grateful. Steve Flusberg, Dharshan Kumaran, and Yula Paluy have provided necessary doses of humor and support. Alexia Toskos Dils and Tania Henetz will be sorely missed office mates, as they were always willing to talk about ideas

and abide my many distractions. Jeremy Glick and Daniel Sternberg have been terrific collaborators and friends.

I had fun and productive conversations with many more: Pavlos Kollias, Orly Fuhrman, Adam November, Nathan Witthoft, Nick Davidenko, Jon Winnawer, Dan Yarlett, Lucas Butler, Daniella Furman, Alan Gordon, Sarah Gripshover, Rebecca Hetey, Ricardo Bion, Justine Kao, Cherish Madu, Nanna Notthoff, Long Ouyang, Grace Tang, and Adriana Weisleder.

I benefitted from the resourcefulness and hard work of several outstanding RAs, including Otto Murphy, Tinu Ngo, and Brennan Klein.

The support staff in the department have been superb: thanks especially to Harry Bahlman, Tamara Danoyan, Roz Grayson, Beth McKeown, and Peter Smith.

Most of all, thanks to my family. My wife, Kristina, has been amazingly supportive. My parents, Tom and Karen, have always been the coolest. And my younger brother, Mark, is someone I will always look up to (despite his diminutive stature). Thanks also to Mike and Julia and Ewan for letting me crash on the couch when I come to town.

Contents

List of Tables

List of illustrations

Chapter 1:
Introduction

Metaphors pervade discussions of social and political issues. Newspaper headlines depict a world where political parties fight wars against each other (and against other metaphorical wars), "Democrats *Battle Back* against Republican '*War on Women*'" (April 18, 2012, *Washington Post*), where hostile nations dance with one another, "*Tango Diplomacy*: Is War with Iran Inevitable?" (May 5, 2006, *Lebanon Wire*), and where economies grow like plants, "From Tiny *Seeds of Ideas*, a Mighty *Economy Grows*" (February, 3, 2012, *Register-Guard*). Do these metaphors affect the way that we conceptualize the issues they are used to describe? Do implicit metaphorical entailments guide how we construe events, reason about difficult issues, and make hard decisions?

Previous work has shown the power of metaphors to shape thought in education (Blackwell, Trzeniewski, & Dweck, 2007; Green, 1993; Mayer, 1993; Petrie & Oshlag, 1979; Sticht, 1993), advertising (Zaltman & Zaltman, 2008), politics (Charteris-Black, 2011; Lakoff, 2002, 2004, 2007, 2008; Schon, 1979; Sternberg, Tourangeau, & Nigro, 1978), and science (Boyd, 1979; Gentner & Jeziorski, 1993; Kuhn, 1979; Pylyshyn, 1979).

Metaphors can be effective pedagogical tools. In an own introductory computer science course I remember being taught how to think like a programmer through metaphor. I learned that there were several phases to software development: the first step is to design a program, the second is to code the program, the third is to test the program, and the fourth is to debug it. The series of roles that I was being taught to embody made some intuitive sense, but I clearly remember the metaphors that were used by the professor to explain exactly what taking on each of these roles would entail. In the design phase, I was to be an "architect" carefully laying out the blueprints of the program. In the coding phase, I was to think of myself as an "engineer" efficiently implementing the design. As a software tester, I was a "vandal," trying to break the program through any means. And, finally, in debugging, I was to be a "detective" looking for clues to find the

problems in the code. These metaphors capture the essence of the various stages of software development and offer a clear and efficient way of communicating how to be a good computer programmer.

Similarly, metaphors in advertisements can influence how people think about products and services. In the mid-90s, for instance, pharmaceutical companies devised a marketing campaign to increase sales of their products in Japan. In the campaign, they described mild or moderate depression as a "cold for the soul" (*kokoro no kaze*), a problem that could be diagnosed and treated medically. In the five years that followed, depression-related visits to the doctor increased 46% and 177 books about depression were published, compared to 27 between 1990 and 1995 (Currie, 2005).

In politics, metaphors often accompany policy initiatives. Legislation is frequently labeled metaphorically, as in the No Child Left Behind Act of 2001 and the Race to the Top Initiative of 2009, and speeches about controversial policies often invoke metaphoric lines of reasoning. President Obama, for instance, has continuously used a "tide" metaphor for the American economy to rally support for financial stimulus and jobs programs: "a rising tide lifts all boats, from the company CEO to the guy on the assembly line" (Obama, 2010). The President uses this metaphor, which was also a favorite of JFK and Ronald Reagan, to make the case that adding capital to the American economy puts money in everyone's pocket, regardless of who specifically gets to cash the stimulus checks.

In science, there are a number of famous examples of metaphors sparking an empirical discovery or theoretical advancement. Ernest Rutherford, for instance, used the structure of the solar system as a model for thinking about the structure of an atom and in the process provoked a paradigm shift in physics (Nersessian, 1992). In cognitive science, there is a long history of using the most advanced technology of the day as a metaphor for the mind: is the mind a clay tablet? a hydraulic system? a telephone switchboard? a computer?

Indeed, members of this dissertation committee routinely use metaphors in their scientific communications. Herbert H. Clark, for instance, invites readers of his work to

think of language as a joint action by comparing it to "two people waltzing, paddling a canoe, playing a piano duet, or making love" (Clark, 1996, p. 3). James L. McClelland argues that one of the aims of the PDP enterprise is "to replace the 'computer metaphor' as a model of the mind with the 'brain metaphor'" (Rumelhart, Hinton, & McClelland, 1986, p. 75). Samuel McClure, in a recent article on the neural underpinnings of addiction, describes the effect of substance of abuse as a "hijacking [of] common reward-related neural circuitry" (Radu & McClure, In press, p. 1). Eve V. Clark emphasizes the complexity of linguistic systems and the challenge faced by a language-learner by describing the "web of skills" (Clark, 2003, p .1) that one must acquire in order to master any language. And Lera Boroditsky describes language as a type of cultural handbook, where syntactic structures, the presence of particular lexical items, and other linguistic patterns force speakers of a language to attend to features of the world that are culturally significant.

Often, metaphors in education, advertising, politics, and science are invoked deliberately to reason about the topics they describe. Speakers who employ them identify important entailments and explain why the metaphors are useful tools for thinking. For instance, Herbert H. Clark (1996) elaborates on his metaphor for language:

> When Fred Astaire and Ginger Rogers waltz, they each move around the ballroom in a special way. But waltzing is different from the sum of their individual actions -- imagine Astaire and Rogers doing the same steps but in separate rooms or at separate times. Waltzing is the joint action that emerges as Astaire and Rogers do their individual steps in coordination, as a couple. Doing things in language is likewise different from the sum of a speaker speaking and a listener listening. It is the joint action that emerges when speakers and listeners -- or writers and readers -- perform their individual actions in coordination, as ensembles.
>
> (p. 3).

In this case, the structural entailments of the metaphor license a particular inference, but Clark does not leave it to readers to make the inference on their own. He explicitly says that language is like a waltz because it is an emergent property of an ensemble, a speaker speaking and a listener listening.

Previous work on the role of metaphor in learning and thinking has focused on the effects of deliberately invoked, instructive metaphors like Clark's extended comparison of language and waltzing (e.g., Mayer, 1993). For instance, Mayer (1983) describes a set of experiments in which students were taught scientific concepts with or without instructive metaphors. He found that students who learned the concept of a radar metaphorically, as a bouncing-pulse, recalled 50 percent more of the principles they were taught and generated twice as many solutions to problem-solving transfer questions than students who learned about radar without the metaphor. In light of these and other similar findings, Mayer (1993) argues that instructive metaphors actively affect how students select, organize, and integrate new knowledge.

Close inspection of mundane linguistic discourse reveals another class of metaphor that suffuses our talk of complex and abstract things. These metaphors are often highly conventional and systematic (Lakoff & Johnson, 1980). For instance, we commonly use terms that are associated with PLANTS[1] to talk about the ECONOMY:

> The *seeds* of economic *growth* were *planted* years ago. Today, they are just starting to *bear fruit,* and soon we will *reap* the rewards.

Interestingly, there are often multiple sets, or families, of metaphors for a given topic. For instance, we can also talk about the ECONOMY as a VEHICLE (e.g., the economy is *off track* or *broken down*) or a BODY (e.g., the economy used to be *healthy* but is now *suffering*) or a MACHINE (e.g., the economy is *broken* and needs to be *fixed*).

Unlike instructive metaphors, instances of these conventional metaphor families often go undetected. To illustrate that people often overlook conventional metaphors, I asked a group of 50 people on Mechanical Turk[2] (Buhrmester, Kwang, & Gosling, 2011) to identify all of the metaphors they could find in an excerpt of a recent speech by President Obama on the economy. In the passage (see Appendix A) the President uses at

---

[1] Following convention, concepts will be referenced by words in all capital letters.

[2] Mechanical Turk (www.mturk.com) is a crowd-sourcing platform operated by Amazon.com where people are paid to complete small tasks online.

least two metaphors. One of the metaphors is the "tide" metaphor for the economy: "An America where a rising tide really did lift all boats." The other includes the PLANT metaphor for the ECONOMY: "we can have a strong and growing economy." Of the 50 people who were asked to identify "as many metaphors as you can" from the report, 100% identified the "tide" phrase, whereas 2%, a single person, recognized "growing" and "strong" as metaphorical ways of talking about the economy.[3]

Nevertheless, these conventional metaphors, like the instructive metaphors above, appear to have structural entailments that license particular inferences. For instance, if the ECONOMY is a VEHICLE and it is broken down, then getting it moving again might require a financial jumpstart. On the other hand, if the ECONOMY is a stunted PLANT, giving a momentary jolt of nutrients, sunlight, and water is unlikely to be an effective longterm solution for the health of the plant. Instead, the economy might be better served by consistent sunlight, water, and nutrients and a supportive environment. It is unclear whether people go beyond what they read to spontaneously make these kinds inferences after hearing or reading conventional metaphors.

This would be an important finding with wide-ranging implications, since much of the language we use, particularly in discussions of social and political issues, is metaphorical (Den Boer, 1998; Graesser, Long, & Mio, 1989; Lakoff & Johnson, 1980; Pollio, Barlos, Fine, & Pollio, 1977; Smith, Pollio, & Pitts, 1981). Den Boer (1998), for instance, found that 19% of language from a broad sample of text was metaphorical, and Pollio et al. (1977) estimate that speakers produce 3,000 novel metaphors every week.

According to Lakoff and Johnson (1980, 1999), people do make metaphorical inferences when processing conventional metaphors. The presence of metaphor families in natural language reflects the metaphorical nature of the human conceptual system. That is, our tendency to talk about the ECONOMY using PLANT terms, reflects the way that we think about the ECONOMY -- namely, as a PLANT that can grow or shrink and be strong or weak, and maybe even as an entity that grows best in a supportive environment with consistent stimuli. On this view, our conception of the ECONOMY is not solely

[3] Indeed, many would argue that these phrases are not metaphoric at all -- that economies literally grow and can be strong or weak.

contained in our representation of a PLANT. It is also subserved by our conceptual representations of VEHICLE, PERSON, and MACHINE. As a result, talking about the ECONOMY as a PLANT encourages inferences about the ECONOMY that are consistent with our knowledge of PLANTS, but talking about the ECONOMY as a VEHICLE or PERSON or MACHINE would encourage different inferences.

Psychological evidence for Lakoff and Johnson's (1980, 1999) theory remains sparse. Most of the evidence comes from patterns in linguistic data (i.e., the presence of extended conventional metaphors in natural discourse). Murphy (1996, 1997) and others (e.g., Keysar & Bly, 1995, 1999; Keysar, Glucksberg, Shen, & Horton, 2000) have argued that there are a number of potential problems with using linguistic data to make psychological claims, since patterns in language do not necessarily reflect patterns of thought.

One possibility is that these conventional metaphors are idioms that do not instantiate any conceptual structure (Keysar, Glucksberg, Shen, & Horton, 2000). Phrases like *kick the bucket* and *red herring,* for instance, have lexicalized meanings ("to die" and "to mislead"), and as a result, neither phrase seems to instantiate a conceptual mapping or structural entailments. Indeed, many lexical items in English have roots in metaphor (Jeffers & Lehiste, 1979) with complex etymologies that are unrecognizable today. It is possible that conventional metaphoric phrases like "a strong and growing economy" or "the economy is stalled" similarly have fossilized meanings. If so, then one would not expect conventional metaphors to support metaphorical reasoning or inference-making.

Some patterns of linguistic data support this alternative view. Despite the observation that metaphor families are generally systematic in English, individual speakers often do not use metaphor families consistently when they speak. Sometimes they do, as in I've been "up and down like a roller coaster" (Tupac) and sometimes they do not, as in "I've been up and down so many times that I feel as if I'm in a revolving door" (Cher). For these reasons, further empirical investigation is needed to understand the psychological implications of the linguistic finding that metaphor is widespread and systematic generally in language.

Previous work in psychology has shown that instances of conventional metaphor families are generative (Thibodeau & Durgin, 2008). Hearing one metaphor from a given metaphor family facilitates the comprehension of another metaphor from the same family. For instance, hearing "*David has a hard time swallowing new ideas. He has to stew over them for days.*" makes it easier to process the sentence "*Otherwise they give him indigestion.*" This is not true of idioms like "*kick the bucket,*" which suggests that conventional metaphors may indeed have some meaningful psychological consequences. Hearing language that compares the ECONOMY to PLANTS, for instance, may actually guide people to think of the ECONOMY as a PLANT in certain respects.

Further, work by Boroditsky (2000) has shown that at least one abstract domain is conceptualized metaphorically. In a set of studies, participants answered ambiguous questions like "If Wednesday's meeting has been moved forward two days, when is the meeting?" Their answers varied systematically depending on whether they were primed to think about themselves moving through time or time moving over them, revealing that people use the relational structure of SPACE to think about the abstract domain of TIME.

Critics, however, argue that the domain of time may be unique in its reliance on metaphor. First, it is only one demonstration of the psychological reality of metaphorical reasoning and some argue that "people's ability to link a single dimension of space to a single dimension of experience is a rather modest example of metaphorical thinking compared with the grandiose ambitions that we have for it" (Pinker, 2007, p. 267). Second, in contrast to other abstract domains like IDEAS, there may be relatively few metaphor families for TIME. When we talk about the sequencing of TIME, we rely on spatial terms extensively (Clark, 1973; Traugott, 1978). As a result, our conceptual representations of SPACE and TIME may be more tightly connected than our conceptual representations of, for instance, the ECONOMY and PLANTS.

It remains an open question whether conventional metaphors for other complex and abstract domains influence the way that we think. Do people spontaneously make metaphorical inferences when they read or hear metaphors in everyday language? Do the implicit structural entailments of these conventional, and often overlooked, metaphors

shape the way people think about the important issues that they describe? What sorts of representations might support this process? Is the influence of the metaphor the result of an online mapping mechanism or are representations inherently metaphorical?

I will use the domain of crime as a test bed for investigating these questions and exploring the role of conventional metaphors in reasoning. I chose crime because it is socially important and because we use a variety of metaphors in talking about it. Both crime and the criminal justice system designed to deal with it, impose tremendous costs on society. Over 11 million serious crimes are reported in the United States each year (FBI Uniform Crime Report, 2008), and the US has the highest per capita imprisonment rate of any country (Walmsley, 2009). Despite being home to only 5% of the world's population, the United States holds 25% of the world's prisoners, with nearly 1% of the US population living behind bars (Liptak, 2008). Addressing the crime problem is important for social policy.

When people talk about crime, they routinely use metaphor. Increases in the prevalence of crime are described as crime *waves*, *surges* or *sprees*. A spreading crime problem is a crime *epidemic*, *plaguing a city* or *infecting a community*. Crimes themselves are *attacks* in which *criminals prey on unsuspecting victims*. And criminal investigations are *hunts* where criminals are *tracked* and *caught*. Are such metaphors just fancy ways of talking, or do they have consequences for how people reason about crime?

Some observations of crime policy in the real world suggest that people may indeed take metaphors as more than just talk. Shifts in metaphors are often accompanied by shifts in policy. In the 1980s Ronald Reagan declared a war on drugs, with smugglers, dealers, and users defined as the enemy. Policies in line with the war on drugs mandated longer, harsher sentences for drug-related crime. Since then, the incarceration rate in the US has more than quadrupled (US Bureau of Justice Statistics, 2011).

Others have taken the CRIME is a VIRUS metaphor seriously and have implemented programs to treat crime as a contagious disease. One crime-prevention program run by an epidemiologist in Chicago treats crime according to the same regimen used for diseases like AIDS and tuberculoses, focusing on preventing spread from person

to person (Kotlotwitz, 2008).

Some criminal justice scholars have even implicated bad metaphor as the root of failure in crime prevention (Kelling, 1991). In one case, a serial rapist attacked 11 girls over a 15-month period before being captured by the police. During those 15 months, the police had information that (had they shared it with the community) could have prevented some of the attacks. Instead, they opted to keep that information secret to set traps for their suspect. The police, on Kelling's analysis, were entrenched in their metaphorical role of hunting down and catching the criminal, and neglected their responsibility to inoculate the community against further harm. The girls, Kelling writes, "were victims… not only of a rapist, but of a metaphor" (p. 1, 1991).

In this dissertation, I explore the role of metaphor in high-level reasoning and conceptual representation in a collection of empirical studies and computational simulations. In Chapter 2, I describe five experiments in which I explore (a) whether conventional metaphoric frames really do influence how people think about important issues; (b) what factors (e.g., context, pervasiveness of the metaphor) contribute to metaphoric reasoning; and (c) whether there any individual differences in people's propensity for thinking metaphorically. I compare two metaphors for crime, CRIME-as-VIRUS and CRIME-as-BEAST, to test whether people suggest different ways of solving a crime problem on the two metaphoric framings of the problem.

In Chapter 3, I investigate the processes that underlie metaphorical thinking. I describe six experiments that explore (a) whether the metaphor framing is an effect driven by spreading activation from the metaphorical words (i.e., lexical priming), rather than their structural entailments; (b) whether the time-course of processing and supporting conceptual context is important for the framing effect; (c) if the framing effect is driven by a deliberative reasoning strategy; and, finally (d) whether it is possible to identify specific features of the target domain (e.g., specific relational structures) that the metaphorical frames serve to make salient.

These experiments suggest, first, that metaphors are powerful in shaping the way we think about important issues. They guide our patterns of reasoning and behavior, in

many cases without our knowledge. Second, the dynamics of this process are driven by a metaphor's ability to constrain how we build a representation of the topic. In the experiments I describe, this ability hinges on the metaphor being presented early in a stream of processing, as a metaphor (and not as a lexical prime), and in a context that allows for at least some ambiguity in reasoning.

To better understand how this process unfolds, in Chapter 4 I present a computational theory of metaphoric inference. I first describe how current, state-of-the-art computational models of analogical and metaphorical reasoning work (e.g., Falkenhainer, Forbus, & Gentner, 1989; Hummel & Holyoak, 1997, 2003). A core feature of these models is a shared commitment to representing concepts as non-overlapping systems of propositional knowledge (e.g., in these models there is an ECONOMY concept and a PLANT concept and these concepts do not share representational space with each other). Reasoning by analogy or metaphor involves, first, bringing two concepts, a source and a target, into alignment, and second, projecting specific aspects of the relational structure from the source onto the target. The primary aim of current models is to describe a mapping mechanism that explains how such an alignment and projection process takes shape.

In contrast, building on work by Hinton (1986) and Rogers and McClelland (2008), I argue that a specific mapping mechanism may not be required for all forms of metaphoric or analogical inference. A core feature of this connectionist approach is a commitment to representing conceptual knowledge in shared weights rather than disjoint propositions. On this view, homologous relational structure can be consolidated in overlapping, distributed representations even across semantically disparate domains (e.g., conceptual knowledge for ECONOMY and PLANT shares representational resources). Reasoning by analogy or metaphor is a natural byproduct of this representational scheme and does not require an on-line mapping mechanism: simultaneously activating a source and target domain will highlight shared relational structure. While this approach is not intended to capture explicit reasoning by analogy or metaphor, it may help to explain why people do not take longer to process conventional metaphors in natural discourse (Blank,

1988; Ortony, 1975; Ortony, Schallert, Reynolds, & Antos, 1978) and why, in the experiments I present below, people do not seem to notice the influence of the metaphor on their decisions.

Finally, in Chapter 5, I will discuss the implications of these findings. I will argue that the metaphors we use to talk about social issues are vital important to the way we reason about, attempt to solve, and behave in relation to issues facing the country and world today. In short, good metaphors lead to effective problem solving; bad metaphors make problems worse.

Chapter 2:
The Metaphor Framing Effect

Using different metaphors can lead people to reason differently about concepts like time, emotion, or electricity (Boroditsky, 2000, 2001; Boroditsky & Ramscar, 2002; Gentner & Gentner, 1983; Gentner, Imai, & Boroditsky, 2002; Gibbs, 1994). In research by Gentner and Gentner (1983) people's reasoning about electricity flow differed systematically depending on the metaphoric frame used to describe electricity (flowing water vs. teeming crowds). Such findings on metaphorical framing are grounded in a larger body of work that has established the importance of linguistic framing in reasoning (Loftus, 1975; Tversky & Kahneman, 1981), and the importance of narrative structure in instantiating meaning (Bransford & Johnson, 1972). But there are still questions about the pervasiveness of the role of metaphor in thinking. Very little empirical work has demonstrated that metaphors in language influence how people think about and solve real-world problems (Pinker, 2007).

In this chapter I investigate whether using different metaphors to talk about crime leads people to reason about crime differently and, in turn, to propose different solutions to a crime problem. I will focus on and extend two contrasting metaphors for crime: crime as a virus and crime as a beast. Do these metaphors lead people to reason about crime (and other problems) consistent with the entailments of the metaphors? For example, might talking about crime as a virus lead people to propose treating the crime problem the same way as one would treat a literal virus epidemic? Might talking about crime as a beast lead people to propose dealing with a crime problem the same way as one would deal with a literal wild animal attack?

To help generate a clear set of predictions, I conducted a norming survey asking 28 participants on Mechanical Turk to describe what should be done to solve a literal virus or beast problem. Participants were asked to imagine a "virus infecting a city" or a "wild beast preying on a city" and then to describe the best way to solve the problem that they had imagined. Those who imagined a "virus infecting the city" universally suggested

investigating the source of the virus and implementing social reforms and prevention measures to decrease the spread of the virus. That is, they wanted to know where the virus was coming from, whether the city could develop a vaccine, and how the virus was spreading. They also wanted to institute educational campaigns to inform residents about how to avoid or deal with the virus and encourage residents to follow better hygiene practices. Those who imagined a "wild beast preying on a city" universally suggested capturing the beast and then killing or caging it. They wanted to organize a hunting party or hire animal control specialists to track down the beast and stop it from ravaging the city.

Might these schematic representations for solving literal virus or beast problems transfer to people's reasoning about crime if crime is metaphorically framed as a virus or a beast? That is, if crime is talked about as a virus, will people suggest diagnosing the root cause of the problem and enacting social reform to treat and inoculate the community?[4] If crime is a beast, will people suggest catching and jailing criminals in order to fight off the crime attack?

In Experiment 1, people were asked to read a report about increasing crime rates in the city of Addison and propose a solution to the problem. For half of the participants, crime was metaphorically described as a beast preying on Addison, and for the other half as a virus infecting Addison. The rest of the report contained crime statistics that were identical for the two metaphor conditions.

Experiment 1

Method

[4] It should be noted that there are two somewhat different metaphorical frameworks that treat crime as an illness. In one, the community or population is seen as an organism, and crime is a disease that is developing inside that organism (e.g., ''Violent crime is a cancer that eats away at the very heart of society.''). In another, the community is seen as individual agents and crime is a contagious disease that can be passed on from one person to another forming an epidemic. In this paper the stimuli did not strongly distinguish between these different varieties of crime as illness metaphors, but doing so would be an interesting extension of this work, as these metaphors suggest somewhat different implications for treating crime.

Participants. The participants were 485 students – 126 from Stanford University and 359 from the University of California, Merced – who participated as part of a course requirement. Gathering data from the two populations allowed us to get a broader cross-section of the general population. This seemed important since people's conceptions of social issues like crime are likely to differ as a function of factors like socioeconomic status and personal experience.

Materials. Participants filled out a survey that included a short paragraph about crime in the fictional city of Addison and some follow-up questions. There were two versions of the crime paragraph that differed only in the embedded metaphor: in one, crime was a beast; in the other, crime was a virus. Most of the paragraph consisted of crime statistics, which were the same in both versions. Half of the participants were given the crime-as-beast version and half the crime-as-virus version. The paragraph read:

> Crime is a {wild beast preying on / virus infecting} the city of Addison. The crime rate in the once peaceful city has steadily increased over the past three years. In fact, these days it seems that crime is {lurking in / plaguing} every neighborhood. In 2004, 46,177 crimes were reported compared to more than 55,000 reported in 2007. The rise in violent crime is particularly alarming. In 2004, there were 330 murders in the city, in 2007, there were over 500.

There were two follow-up questions: 1) In your opinion what does Addison need to do to reduce crime? 2) Please underline the part of the report that was most influential in your decision. This question was aimed at discovering if participants explicitly noticed or made use of the metaphor.

Design. The survey was included in a larger packet of paper questionnaires that were unrelated to this study.

Coding. Proposed solutions to the crime problem in Addison were divided into two categories in line with the norming study: 1) diagnose/treat/inoculate, and 2) capture/enforce/punish. Responses were categorized as "diagnose/treat/inoculate" if they suggested investigating the underlying cause of the problem (e.g., "look for the root cause") or suggested a particular social reform for treating or inoculating the community (e.g., fix the economy, improve education, provide healthcare). Responses were categorized as "capture/enforce/punish" if they focused on the police force or other methods of law enforcement (e.g., calling in the national guard) or modifying the criminal justice system (e.g., instituting harsher penalties, building more jails). For brevity, I will refer to the "diagnose/treat/inoculate" category as "reform" and the "capture/enforce/punish" category as "enforce."

Each participant's response was weighted equally – as a single point towards the analysis. For solutions that solely emphasized either reform or enforcement, the respective category was incremented by a point. Responses that exclusively emphasized one approach were the majority. Occasionally, participants listed both types of suggestions. If the response listed a disproportionate number of suggestions that were consistent with one approach (e.g., if the response listed three suggestions in line with reform and only one in line with enforcement, as in "investigate the root cause, institute new educational programs, create jobs, and hire more police") then it was coded as a full point for the category that was emphasized. But if the response considered both approaches equally, then each category such that each was incremented by 0.5.

Thirty of the 485 responses (6%) did not fit into either category because they lacked a suggestion (e.g., "I don't know", "I need more information", "It should be addressed"). These data were omitted from analysis.

Participants' crime reducing suggestions were coded blindly by two coders. Cohen's K – a measure of inter-rater reliability – was 0.75 indicating high agreement between the coders ($p < .001$). All disagreements between the coders were resolved in discussion between them before analyzing the data.

Results

Overall, participants were more likely to suggest enforcement-oriented strategies (65%) than reform-oriented ones (35%), χ^2[1,N=455] = 41.85, $p < .001$. But, as predicted, the solutions proposed differed systematically with the metaphorical frame in the crime report (see Figure 1). Participants given the crime-as-beast metaphor were more likely to suggest enforcement (74%) than participants given the crime-as-virus framing (56%), *D* [N=455] = .18,[5] bootstrapped $p < .001$.[6]

When asked to identify the most influential aspect of the report, 97% of the participants ignored the metaphor. Only 15 participants (3%) identified metaphor as most influential. Eight of these participants had read that crime was a virus and seven had read that crime was a beast.

[5] *D* will refer to the proportion of responses that emphasize enforcement given the beast frame minus the proportion of responses that emphasize enforcement given the virus frame (i.e., *D* = p(enforcement | beast) - p(enforcement | virus)).

[6] I will report bootstrapped p-values for comparisons in which the dependent variable is dichotomous and there is only one dichotomous independent variable. In other cases, I will use a logistic regression or chi-square model to test particular comparisons. See appendix B for a description of how the bootstrapped p-value was computed. In general, the results from the bootstrapping procedure are nearly identical to those of a logistic regression model. However, using the bootstrap is advantageous for theoretical reasons: bootstrapping is non-parametric and does not require that one make assumptions about the distribution of the data (as is required for regression models).

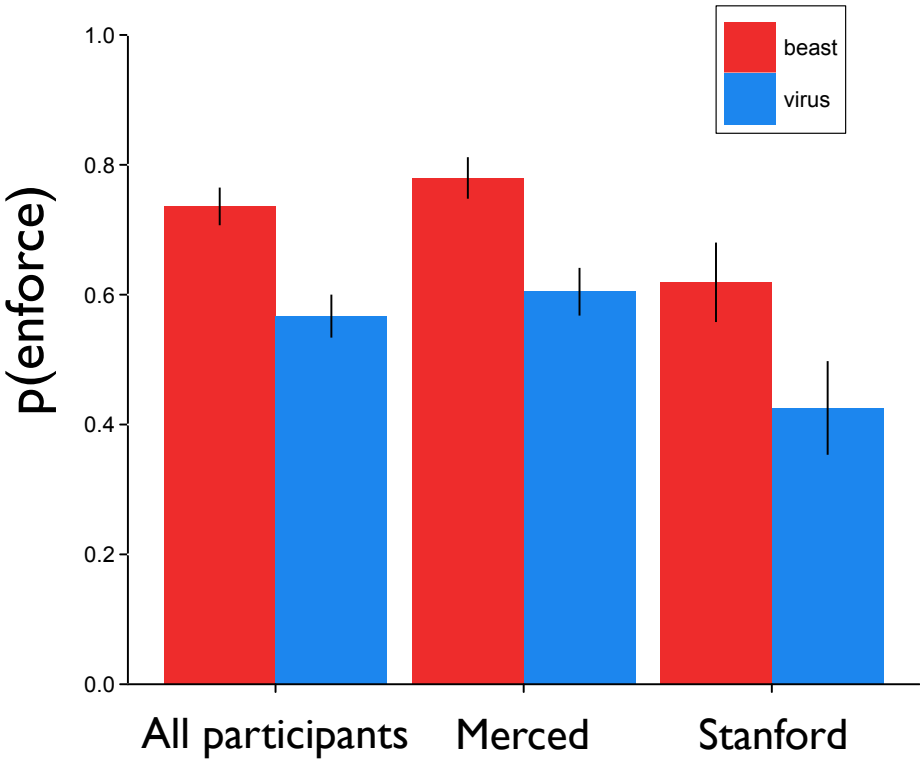

Figure 1: Results of Experiment 1. Framing crime as a "beast" encouraged people to suggest reducing crime through increased law enforcement and punishment relative to framing crime as a "virus." The first set of bars collapses across participants from UC-Merced and Stanford. Data from these populations are shown separately in the second and third sets of bars.

Discussion

In Experiment 1, peoples' crime-reducing suggestions differed with the metaphor used to frame the crime problem. Participants who read that crime was a virus were more likely to propose investigating the root causes of the issue and instituting social reforms than participants who read that crime was a beast. Participants who read that crime was a beast were more likely to propose fighting back against the crime problem by hiring police officers and building jails – to catch and cage the criminals – than participants who read that crime was a virus.

Despite the clear influence of the metaphor, participants generally identified the crime statistics, which were the same for both groups, and not the metaphor, as the most influential aspect of the report. These findings suggest that metaphors can influence how people conceptualize and, in turn, approach solving an important social issue, even if people don't explicitly perceive the metaphor as being especially influential.

In Experiment 2, I made two substantive changes to the task to further test the role of metaphor in reasoning. First, I changed how the metaphoric frame was presented. In Experiment 1, the metaphoric frame was established several times and included vivid relational language. For example, crime was said to be either "preying" and "lurking," or "infecting" and "plaguing" the community. These metaphorical verbs explicitly specified relations between crime and the community. Is specifying relations explicitly in this way necessary for people to make appropriate inferences, or might people be able to spontaneously extract the relevant relational inferences given a minimal metaphorical suggestion? Might a single carefully chosen and appropriately placed word be enough to instantiate a metaphorical frame and induce different reasoning strategies?

So in Experiment 2 I removed the relational verbs from the report, replacing them with a single word metaphor that described crime as a "virus" or "beast" in the introductory sentence. The two conditions differed only in this one word, and otherwise included all the same information.

The second change I made was to add an additional follow-up question: What is the role of a police officer in Addison? This question aimed to disambiguate the modal crime-reducing suggestion from Experiment 1, which was "increase the police force." In that context, I interpreted the response (and close variants of it) as a suggestion for increased law enforcement and punishment. However, police officers do not just catch and punish criminals. They also serve as crime deterrents, educators, and role models and it is possible that some participants intended for the increased police presence to serve in this way. Including this question allowed these participants an opportunity to specify how they saw the increased police force impacting the community.

Experiment 2

Method

Participants. Experiment 2 was conducted online with 347 participants recruited on Amazon's Mechanical Turk. They were paid $1.60 – consistent with a $10/hour pay rate since the study took 5 to 6 minutes for most participants to complete. Running the study online enabled us to sample from a more diverse population than that available on college campuses generally (Buhrmester, Kwang, & Gosling, 2011).

I used Mechanical Turk's exclusion capabilities and tracked IP addresses to ensure that participants were not sampled repeatedly (within or across versions of experiments).[7] I also restricted the study to Turkers with a 95% or better performance record to ensure that I was sampling high quality participants.[8]

At the end of the study participants were asked about their language, location, and certain background information (age, gender, political affiliation). I then restricted my analysis to residents of the United States who were native English speakers, and that excluded 94 participants. This left data from 253 participants for analysis, 157 female and 96 male. Their ages ranged from 18 to 66, with a mean age of 32 (median 29). Eighty-two reported an affiliation to the Democratic Party, 57 reported an affiliation to the Republican Party, and 114 were Independent.

Materials. The crime report in this study was similar, but not identical to the one used in Experiment 1. It instantiated the beast or virus metaphor for crime with a single word. It also included a number of words that intuitively seem to instantiate different meanings on the two different framings (e.g., "good shape," "vulnerabilities," "defense systems," "weakened," "succumbed"). It read as follows:

---

[7] This applies not only to this experiment, but across all experiments that were run on Mechanical Turk: no Mechanical Turk User ID was allowed to participate in more than one of the metaphorical framing experiments having to do with crime or in experiments where the virus and beast frames were used.

[8] On Mechanical Turk, "Requesters" have the opportunity to give positive or negative feedback to their participants, which can then be used as a criterion for future "Requesters."

> Crime is a {beast / virus} ravaging the city of Addison. Five years ago Addison was in good shape, with no obvious vulnerabilities. Unfortunately, in the past five years the city's defense systems have weakened, and the city has succumbed to crime. Today, there are more than 55,000 criminal incidents a year - up by more than 10,000 per year. There is a worry that if the city does not regain its strength soon, even more serious problems may start to develop.

After participants read the paragraph, they answered three follow-up questions in the following order: 1) In your opinion what does Addison need to do to reduce crime? 2) What is the role of a police officer in Addison? 3) Please copy the part of the report that was most influential and paste it in the text area below. Questions one and two were free-response. Question three was copy and paste: participants were shown the report adjacent to an open text field and were asked to copy the portion of the report that was most influential in their reasoning and paste it into the open text field.

Design. Each step of the experiment was presented on a separate screen. The initial crime report was presented on a screen by itself. After participants read the report and clicked a button indicating they had finished reading it, the report disappeared and the first follow-up question appeared on a screen by itself. Similarly, each subsequent question was shown on a separate screen. On the final screen, participants provided their age, gender, political affiliation, first language, and country of residence.

Participants were instructed not to use the “back” button on their browser. If they did, the session was terminated. This ensured that participants did not reread the crime report when they were later asked questions about it. This seemed important because participants’ interpretation of the report may change on a second reading, after they know the content of the follow-up questions.

The experiment was implemented in Adobe Air and hosted on the author’s academic website. In order to participate, Turkers followed a link to the survey website. Upon finishing the survey, they were given a random completion code, which they submitted to the Mechanical Turk website as verification that they had participated.

Coding. Crime-reducing suggestions were coded into two categories (reform and enforcement) as in Experiment 1. But in Experiment 2 I also coded whether the participant exclusively suggested increasing the police force. For these responses, I planned to use the follow-up question about the role of a police officer in Addison to specify whether the participant thought a police officer's primary role was as an instrument of social reform and prevention or an instrument of law enforcement and punishment.

Interpretations of the role of a police officer were coded into two groups that were analogous to the categories created for the first question: 1) crime deterrent, and 2) law enforcer and punisher. Interpretations that emphasized the police officer's role in preventing crime, educating youth, or serving as a role model in the community were coded as "crime deterrent." Interpretations that emphasized the police officer's role in catching criminals, responding to crime reports, or punishing offenders were coded as "law enforcer and punisher." As in Experiment 1, each response contributed one point to the analysis. This point either went to one of the two categories or was split evenly between them.

Seven (3%) crime-reducing suggestions and 18 (7%) police officer interpretations were eliminated from the analysis because they lacked a suggestion or interpretation. It is possible that relatively more police officer interpretations fell into this category because the question was not prefaced with "In your opinion" (several responses to this question were a variant of "the report didn't say what the role of a police officer in Addison was").

Answers to both of the free response questions were coded blindly by two coders. Inter-rater reliability was high for both: Cohen's K for crime-reducing suggestions was .86 ($p < .001$); Cohen's kappa for interpretations of the role of a police officer was .72 ($p < .001$). All disagreements between the coders were resolved in discussion between them before analyzing the data.

Results

The results of Experiment 2 replicate the findings from Experiment 1. Participants were again more likely to suggest reducing crime through enforcement (62%) than reform (38%) overall, χ^2[1,N=242] = 13.67, $p < .01$. But the tendency towards enforcement was more pronounced for participants who read that crime was a beast (71%) than for those who read that crime was a virus (54%), D[N=242] = .17, $p < .05$ (see Figure 2).

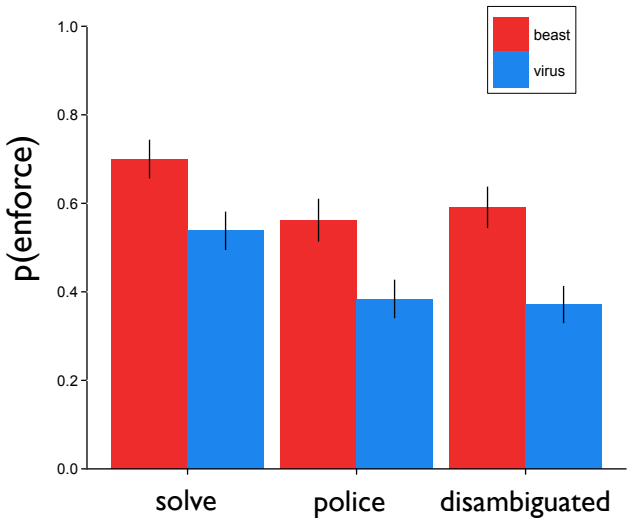

Figure 2: Results of Experiment 2. Framing crime as a "beast," relative to "virus," caused people to emphasize enforcement in their suggestions for reducing crime (first set of bars) and in their interpretations of the role of a police officer in Addison (second set of bars). The third set of bars shows disambiguated crime-reducing suggestions: participants who exclusively proposed "increasing the police force" as the way to reduce crime in Addison were "disambiguated" with responses to a follow-up question about the role of a police officer in Addison.

Of the responses that emphasized enforcement, 81 (31%) exclusively suggested increasing the police force. Asking people to specify the function of a police officer in Addison further clarified the effect of the metaphor. Because "police" responses were previously coded as enforcement, recoding them created an overall shift to the reform category in both conditions, with a larger shift in the virus condition as predicted. With

the "police" responses disambiguated, 37% of the responses advocated enforcement in the virus condition, and 59% advocated enforcement in the beast condition, D[N=242] = .22, $p < .01$.

Further, as in Experiment 1, most participants did not report the metaphor as being influential in their reasoning. Only 18 of the 242 participants (7%) did so. Of these, 12 had read that crime was a virus and six had read that crime was a beast.

Effects of the individual difference measures were tested by comparing logistic regression models. One model was fit to predict crime-reducing suggestions with a regressor for the metaphoric frame and a second model was fit with an additional regressor for a given individual difference measure. These models were compared using an ANOVA with Maximum Likelihood Deviance (ML Deviance) representing the additional variance in predicting crime-fighting suggestions accounted for by adding regressor(s) to the model. ML Deviance has a chi-square distribution with the number of added parameters as its degrees of freedom (Myers, 1990).

Including predictors for main effects of gender, age, or political affiliation or interactions between these measures and the frame did not significantly improve the fit of the model (main effects for gender: χ^2[1,239] = 1.97, p = .16; age: χ^2[1,239] = .26, p = .61; political affiliation: χ^2[1,239] = 0.03, p = .86; interactions between the frame and gender: χ^2[2,238] = 1.55, p = .28; age: χ^2[2,238] = .33, p = .85; political affiliation: χ^2[2,238] = 0.05, p = .97).

Discussion

Results of Experiment 2 replicate and extend the findings of Experiment 1. Manipulating the metaphor used to frame the issue of crime influenced how people approached solving the crime problem. When crime was framed as a virus, participants were more likely to suggest social reform, but when it was framed as a beast, they were more likely to suggest law enforcement and punishment.

Remarkably, presenting an otherwise identical report with only one word different in the introductory frame ("Crime is a {virus / beast} ravaging the city of Addison") yielded different problem solving suggestions just as in Experiment 1. While in Experiment 1, the metaphoric frame was established using vivid verbs with rich relational meaning (e.g., crime was said to be either "preying" and "lurking", or "infecting" and "plaguing"), the findings of Experiment 2 showed that these elements need not be specified explicitly. In Experiment 2, people spontaneously inferred these relations inferences even given a single metaphorical noun.

In Experiment 2 participants were asked about the role a police officer should play in Addison. This allowed for a clearer interpretation of their crime-reducing suggestions and boosted the statistical power to detect the influence of the metaphor.

Again, despite the clear influence of the metaphor, participants generally identified the crime statistics, which were the same for both groups, and not the metaphor, as the most influential aspect of the report.

Despite the consistency of the findings across Experiments 1 and 2 there are some limitations of these results. One is that both experiments use a free-response method for eliciting participants' crime-reducing suggestions. The drawback of this approach is that it leaves open the possibility that the effect is merely the result of a weak prime that makes frame-consistent responses more available in the short term. That is, one possible explanation for the results of these two experiments is that the metaphorical frame affects what kind of solution comes to mind easiest. When faced with a complete set of options, people may realize they had neglected to attend to other alternatives and no longer show the influence of the metaphor. For example, a participant who read the "beast" frame may not have spontaneously thought to address underlying problems in the economy or educational system. However, if these factors are made explicitly available as response options, the participant may recognize them as good ideas and may re-bound from the metaphorical framing (i.e., show no metaphorical framing effect).

To test for this, in Experiment 3, participants were presented with a list of four (or five) possible approaches to the crime problem and asked to rank order them according to

how effective they would be at reducing crime in Addison. These included two options that were more consistent with a social reform approach (education, economy) and two (or, in some cases, three) options that were more consistent with an enforcement-oriented approach (police, jails, neighborhood watches). If the metaphor offers more than a weak prime, the results of Experiment 3 should be similar to those of Experiments 1 and 2. But if the metaphor merely serves to make some approaches to crime more available than others, then there should be no systematic effect of the frame in Experiment 3.

Another advantage of this method over the free response method is that it eliminates the need for human coders.

Experiment 3

Method

Experiment 3 was run in three nearly identical versions, which I will describe as Experiments 3A, 3B, and 3C. There were two differences between the three versions of Experiment 3. First, Experiments 3A and 3B included five response options while Experiment 3C included only 4. One response option was eliminated in Experiment 3C because it did not emphasize enforcement and punishment as strongly as the other two options in the enforcement category. Second, the method that was used to gather peoples' crime-fighting suggestion differed between Experiment 3A and Experiments 3B and 3C. In Experiment 3A participants were asked to rearrange the response options in place; in Experiments 3B and 3C participants were asked to drag and drop the response options from one text box into another text box. This change was made to ensure that participants carefully considered each response option.

In addition, data from Experiment 3C comes from two separate experimental sessions: one was run with Experiment 7 of this dissertation (N = 164) and one was run with Experiment 10 of this dissertation (N = 251). In presenting Experiments 7 and 10, I will include only the data that was run specifically in that experimental session; however,

I have collapsed these data for presentation in Experiment 3 to help describe the metaphoric framing effect and to highlight factors that contribute to the effect. There were no differences between these data with respect to the analyses I report for Experiment 3C.

Participants. Each version of Experiment 3 was conducted online with participants recruited and paid through Amazon's Mechanical Turk. There were 939 participants (353 in version A, 171 in version B, and 415 in version C). Mechanical Turk had been updated between Experiments 2 and 3 and, in Experiment 3, it became possible to restrict the sample to Turkers living in the United States.

Of the 353 Turkers in Experiment 3A, there were 212 females and 141 males, whose age ranged from 18 to 65 (mean = 34, median = 25). Of these, 134 identified as Democrats, 144 as Independents, and 75 as Republicans.

Of the 171 Turkers in Experiment 3B, there were 92 females and 79 males, whose age ranged from 18 to 75 (mean = 31, median = 25). And of these, 77 identified as Democrats, 64 as Independents, and 30 as Republicans.

Of the 415 Turkers in Experiment 3C, there were 212 females and 203 males, whose age ranged from 18 to 75 (mean = 31, median = 25); 155 participants identified as Democrats, 210 as Independents, and 50 as Republicans. In Experiment 3C, 6 participants had less than a high school degree; most (357, 86%) had completed at least some college; but 15 participants had a doctorate degree.

Materials. The crime report used in this study was identical to that of Experiment 2 in which the beast or virus metaphor for crime was instantiated with a single word.

Immediately after reading the crime report, participants were presented with a set of possible approaches that Addison could adopt in response to the crime problem. In versions A and B, there were five options:

1. Increase street patrols that look for criminals.
2. Increase prison sentences for convicted offenders.

3. Develop neighborhood watch programs and do more community outreach.
4. Reform education practices and create after school programs.
5. Expand economic welfare programs and create jobs.

In Experiment 3C, the third option -- neighborhood watch -- was excluded because it was relatively ambiguous with respect to its emphasis on enforcement or reform.

As in Experiment 2, I investigated whether people noticed the effect of the metaphor. I used two different methods for exploring this question in Experiment 3. In Experiments 3A and 3B, participants were asked: "The report you read started 'Crime is a _____ ravaging the city of Addison.' Please fill in the blank." In Experiment 3C, they were asked to "copy the part of the report that was most influential and paste it in the text area below."

After answering this question, participants answered a number of personality and background questions. The personality questions included a shortened version of the Big Five Inventory that had ten questions, the BFI-10, (Benet-Martinez & John, 1998; John, Donahue, & Kentle, 1991; John, Naumann, & Soto, 2008), the need-for-cognition scale (Cacioppo & Petty, 1982; Cacioppo, Petty, & Kao, 1984), and the fascism scale (f-scale) (Adorno, Frenkel-Brunswik, Levinson, & Sanford, 1950) (see Appendix C for a complete list of the personality questions).

The BFI-10 has 10 questions designed to measure the "big five" personality dimensions: openness, agreeableness, conscientiousness, neuroticism, and extroversion. The need-for-cognition scale is an instrument that measures "the tendency for an individual to engage in and enjoy thinking" (Cacioppo & Petty, 1982, p. 116) by asking people to indicate to what degree statements like "I prefer complex to simple problems" are appropriate descriptions of themselves. The f-scale is an instrument that measures authoritarian tendencies by asking people to indicate to what degree statements like "Obedience and respect for authority are the most important virtues children should learn" are characteristic of what they believe. Only participants in Experiment 3C completed the fascism and need-for-cognition questionnaires.

After answering the personality questions, participants were asked for age, gender, first language, political affiliation, and educational history (the educational history question was only asked in Experiment 3C).

Design. The design of Experiment 3 was identical to that of Experiment 2 except that participants were asked to rank order the crime reducing suggestions from (1) most likely to reduce crime to (5 or 4) least likely to reduce crime. In version A, participants saw the five options listed in a random order and were asked to reorder them in place (see top panel of Figure 3). The survey software would not allow participants to advance until they had moved at least one option. In versions B and C (see bottom panel of Figure 3), participants saw the four or five options listed in a random order on the left side of the screen and an empty box on the right side of the screen; their task was to drag and drop the options from the left side of the screen over to the right. The survey software would not allow participants to move on in the experiment until they moved all of the options from the text box on the left side of the screen over to the text box on the right side of the screen. The latter method was adopted because it seemed to require that participants engage more deeply in the experiment. Qualtrics survey software was used to run each version of the study.

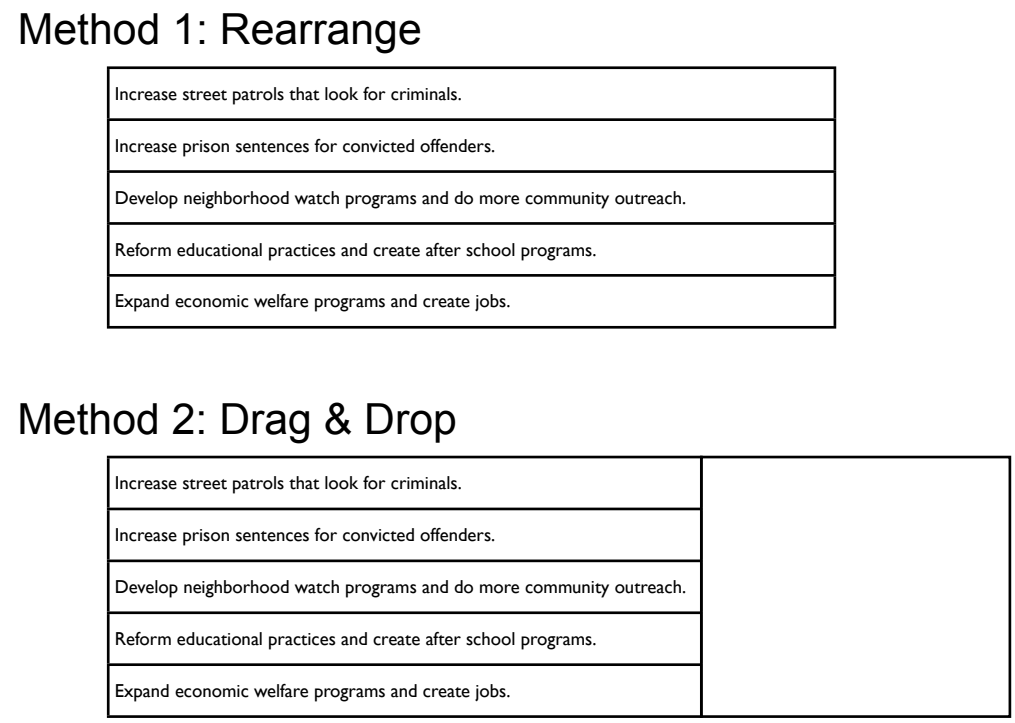

Figure 3: Two methods for gathering crime-reducing suggestions in Experiment 3. The "rearrange" method (top) was used in version A; the "drag & drop" method (bottom) was used in versions B and C.

Coding. Numerical codes were computed for participants' crime-reducing suggestions and individual difference measures.

Crime-reducing suggestions. Responses were coded in two ways. On the first, only the top ranked response was coded -- as "reform" or "enforce." Three of the crime-fighting options (street patrols, prison sentences, and neighborhood watches) were counted as enforcement-oriented and the remaining two (educational reform and economic welfare) were counted as reform-oriented. On the second method for coding responses, all five (or four) options were used to generate a continuous numeric value for each participant. The degree to which each response emphasized enforcement was established by asking a separate group of 35 Turkers to rank the five options on a single continuous dimension ranging from exclusively emphasizing social reform (0) to exclusively emphasizing enforcement (100).

The results of this norming study confirm the categorical distinction between the three enforcement-oriented and two reform-oriented responses: "street patrols" (M = 87.21, sd = 13.5), "prison sentences" (M = 85.11, sd = 22.37), and "neighborhood watches" (M = 58.69, sd = 25.77) were rated as more enforcement-oriented than reform-oriented while the reverse was true for "educational reform" (M = 17.14, sd = 27.13), and "economic welfare" (M = 20.82, sd = 30.93).

These ratings were also used to compute a continuous numeric value for each participant. This was done by multiplying each person's ordered response set by each option's rated value (weight). That is, for each participant, I multiplied the weight associated with their the top-ranked response by 2, the weight associated with their second-ranked response by 1, etc, down to the weight associated with their lowest-ranked response, which was multiplied by -2. Specifically, in Experiments 3A and 3B, the equation used to compute this value was:

2*weight(rank1) + 1*weight(rank2) + 0*weight(rank3) + -1*weight(rank4) + -2*weight(rank5)

In Experiment 3C, the equation was:

2*weight(rank1) + 1*weight(rank2) + -1*weight(rank3) + -2*weight(rank4)

Multiplying the third option by 0 in Experiments 3A and B ensured that this dimension had the same range (from -204.43 to 204.43) for all three versions of the experiment. Participants' values on this measure were roughly normally distributed (see Figure 4) with a mean of -44.93 and standard deviation of 116.46. High values along this dimension reflect an emphasis on enforcement over reform across the entire ordering of crime-reducing options (e.g., 204.43, the highest value along this dimension would be given to a participant who ranked "street patrols" as the method most likely to reduce crime, "prison sentences" as second most likely, "economic welfare" as second to least-likely, and "educational reform" as least likely).

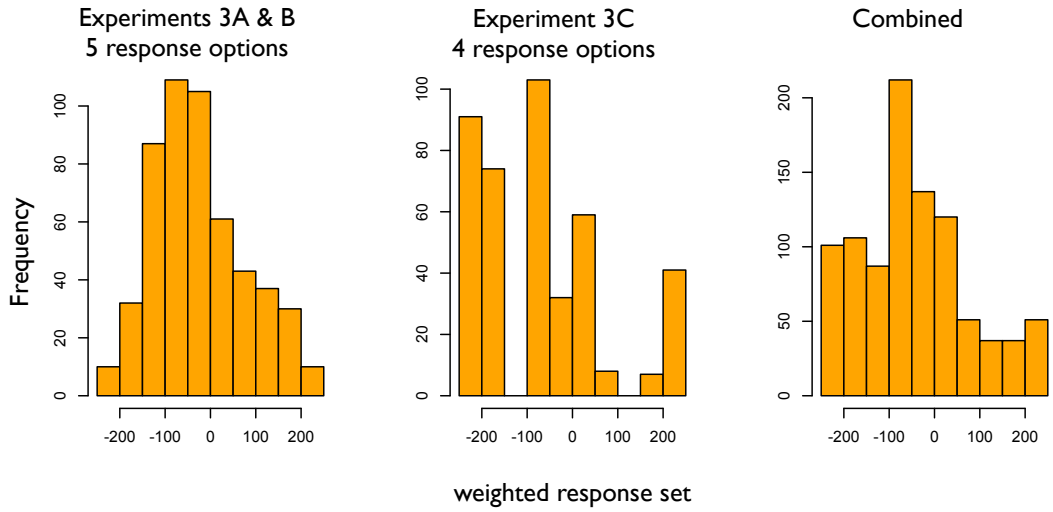

Figure 4: Histograms of the continuous dependent measure that utilizes participants' full ordering of the response set. High values along this dimension reflect a response set that emphasized enforcement-oriented approaches to the crime problem over reform-oriented approaches.

Individual differences. Participants' political affiliation was coded numerically as -1 = Democrat, -.5 = Liberal-leaning Independent, 0 = Center independent, .5 = Conservative-leaning Independent, 1 = Republican (participants who indicated that they were Independents were asked a follow-up question to determine if they thought of themselves as relatively more conservative, relatively more liberal, or right in the middle of the political spectrum).

Scores for extraversion, agreeableness, conscientiousness, neuroticism, openness, need-for-cognition, and fascist tendencies were computed according to the instructions in Appendix C. However, covariance between the personality measures made it difficult to find a reliable and consistent influence of these dimensions across experiments. As a result, I ran a principal components analysis on the seven dimensions across 1,893 participants (from Experiments 3C, 5, 6, 7, 10, and 11; i.e., all of the experiments that included measures of the seven personality dimensions) to extract variably that was consistent across the dispositional measures.

	extra							
agreeable	0.159	agree						
conscientious	0.177	0.202	consc					
neuroticism	-0.27	-0.25	-0.24	neuro				
openness	0.051	0.021	0.05	0.092	open			
nc	0.086	0.007	0.312	-0.22	0.273	nc		
f	-0.01	0.062	0.14	0.001	-0.18	-0.26	f	
pc1 - professionalism	0.536	0.473	0.658	-0.66	0.237	0.607	-0.1	pc1
pc2 - dogmatism	0.179	0.359	0.158	-0.27	-0.63	-0.54	0.695	0

Table 1: Correlations between the big five personality dimensions, need-for-cognition scale, and fascism scale, and the computed principal components (professionalism and dogmatism)across all experiments in which this information was gathered (i.e., Experiments 3C, 5, 6, 7, 10, and 11).

The first two principal components captured 26% and 20% of the variance in the seven personality variables respectively and 46% cumulatively. There was a large drop in the proportion of variance captured by the subsequent components (14%, 13%, 12%, 9%, 7%) indicating that restricting analyses to the first two principal components is sufficient to uncover any potential dispositional component of the metaphor framing effect that could be attributable to these seven dimensions.

The first principal component loaded most highly on conscientiousness (.658), need-for-cognition (.607), extraversion (.536), and agreeableness (.473) in the positive direction and neuroticism (-.659) in the negative direction. It is was also positively correlated with educational history (.159). With respect to specific items on the personality questionnaires, this component was highly correlated with statements that described a proclivity for hard work and taking responsibility (e.g., the dimension was highly correlated with items like "...tends to be lazy", r = -.600. from the BFI-10 and "I like to have the responsibility of handling a situation that requires a lot of thinking," r = .482, on the need-for-cognition scale). As a result, I will call this dimension *professionalism.* Participants

who were conscientious, extroverted, agreeable, enjoyed thinking, and less neurotic have high *professionalism* scores (see Figure 5).

The second principal component loaded most highly on the f-scale (.695) in the positive direction and openness (-.632) and need-for-cognition (-.535) in the negative direction. It was also positively correlated with political affiliation (.241). With respect to specific items on the personality questionnaires, this component was highly correlated with statements that described a proclivity for creative thinking (e.g., the dimension was highly correlated with items like "...has few artistic interests," r = -.557, from the BFI-10, "I would rather do something that requires little thought than something that is sure to challenge my thinking abilities," r = .418, from the need-for-cognition scale, and "Young people sometimes get rebellious ideas, but as they grow up they ought to get over them and settle down," r = -0.546, from the f-scale). As a result, I will call this dimension *dogmatism* in the analyses below. Participants who reported a low need-for-cognition and who agree with items related to having an unquestioning faith in a higher power (or fascist dictator) on the f-scale have high *dogmatism* scores (see Figure 5).

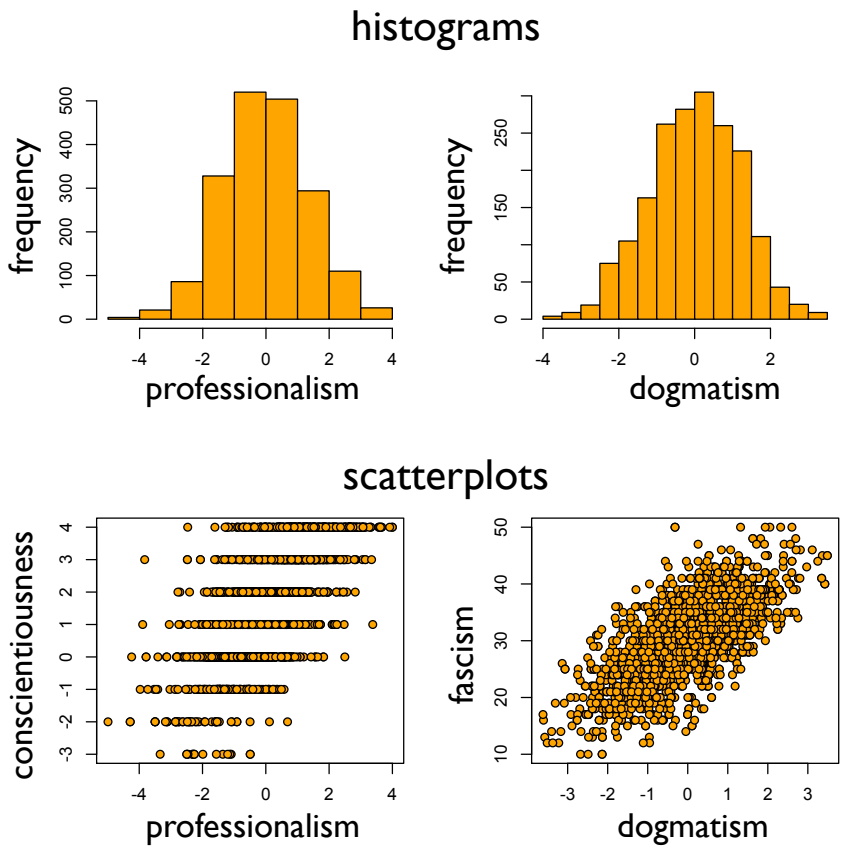

Figure 5: Visualizations of the two principal components computed over the seven personality dimensions measured in Experiments 3C, 5, 6, 7, 10, and 11. On the top row, histograms of the distribution of the first and second principal components (professionalism and dogmatism) reveal that both are normally distributed among the participants sampled. On the bottom row, the leftmost plot illustrates the lack of a correlation between the two principal components while the rightmost plot illustrates the positive relationship between the second principal component (dogmatism) and scores of the f-scale, a measure of authoritarian tendencies.

Results

Collapsing across the three versions of Experiment 3 reveals that participants were more likely to select a response that emphasized reform as their top choice (64%) than one that emphasized enforcement (36%), $\chi^2[1,N=939] = 70.33$, $p < .001$. Experiments 3A and C follow this pattern (with 61% and 81% of participants respectively choosing a response that emphasized reform in the two versions of the experiment): for 3A, $\chi^2[1,N=353] = 18.59$, $p < .001$; and for 3C, $\chi^2[1,N=415] = 159.88$, $p < .001$.

Experiment 3B, however, reveals a different pattern. In that version, people were more likely to select a response that emphasized enforcement (74%), $\chi^2[1,N=171] = 38.37$, $p < .001$. The difference in this main effect across the three versions of Experiment 3 was significant: including parameters for "version" (A, B, C) in a logistic regression model predicting participants' top choice significantly improved the fit of the model, $\chi^2[2,936] = 158.77$, $p < .001$.

On the continuous dependent measure, collapsing across the three versions of the experiment reveals a similar pattern. There was a tendency for participants to emphasize reform ($M = -44.93$, $sd = 116.46$), $t[938] = 11.82$, $p < .001$. However, only the results of Experiments 3A and C followed this pattern (3A: $M = -48.2$, $sd = 86.05$, $t[352] = 10.53$, $p < .001$; 3C: $M = -70.71$, $sd = 129.85$, $t[414] = 11.09$, $p < .001$), while in Experiment 3B there was a main effect of enforcement ($M = 24.44$, $sd = 109.33$), $t[170] = 2.92$, $p < .01$. The difference in this main effect across the three versions was significant, $F[2,936] = 44.41$, $p < .001$.

Variability in this main effect across the three versions of Experiment 3 is likely attributable, to some degree, to the change that was made to the method by which participants indicated their preferred ordering of the response options. However, it should be noted that participants were equally likely to choose the response option that was, by chance, at the top of the list, suggesting that they thoughtfully considered their suggestions on both of the rank order measures.

In version A, 19% of the top ranked responses were the responses that randomly showed up on top of the list; in version B, 22% of the top ranked responses were the responses that randomly showed up on top of the list. Neither of these percentages significantly differ from what one would expect by chance (i.e., 20% when there are five response options): for 3A, $\chi^2[1,N=353] = .23$, $p = .63$, and for 3B, $\chi^2[1,N=171] = .53$, $p = .47$. In version C, 27% of participants' first-ranked responses were the responses that randomly showed up on top of the list. This percentage was not significantly different from what one would expect by chance (i.e., 25% since there were four response options), $\chi^2[1,N=415] = .502$, $p = .48$.

Influence of the frame. The critical result of Experiments 1 and 2, that the frame systematically influenced people's crime-reducing suggestion, was replicated in all three versions of Experiment 3. Overall, participants who read the beast frame suggested an enforcement-oriented response as their top choice more often (42%) than participants who read the virus frame (31%), D[N=939] = .11, $p < .001$ (see Figure 6).

In version A, 44% of responses in the beast condition emphasized enforcement compared to 33% of responses in the virus condition, D[N=353] = .11, $p < .05$; in version B, 81% of responses in the beast condition emphasized enforcement compared to 67% in the virus condition, D[N=171] = .14,$p < .05$; in version C, 25% (27% and 23% in the first and second session, respectively) of responses in the beast condition emphasized enforcement compared to 14% (13% and 16% in the first and second session, respectively) in the virus condition, D[N=415] = .11, $p < .01$. Despite the variability in the main effect of enforcement across the three versions of this experiment, the swing in responses attributed to the frame was highly consistent: in the range of 11-14% for all three versions.

The influence of the metaphoric frame could also be seen on the continuous dependent measure. Overall, the beast frame yielded responses that were more enforcement-oriented (M = -33.29, sd = 122.46) than the virus frame (M = -56.15, sd = 109.32). A linear regression model that included a regressor for the frame, along with regressors for the version of the experiment (A, B, or C), was a better fit to the data than one that only included parameters for version, F[1,935] = 9.87, $p < .01$. The effectiveness of the frame on the continuous dependent measure did not differ significantly across versions of the experiment, although it did approach significance: a regression model that included interaction terms between the frame and version was a marginally better fit to the data than one that did not include interaction terms, F[2,933] = 2.49, $p = .083$. Numerically, the difference in the continuous measure was greatest in version B, although the difference was in the predicted direction across all three versions (D = 2.70, 43.36, and 31.35 for versions A, B, and C, respectively).

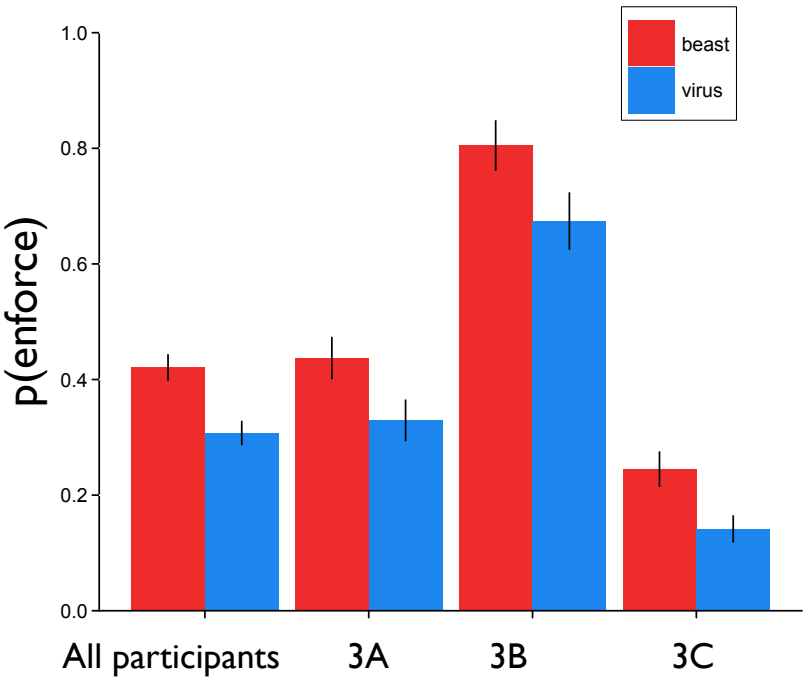

Figure 6: Results of Experiment 3. People who read that crime was a beast were more likely to select en enforcement-oriented response to crime as their first choice than people who read that crime was a virus. The first set of bars collapses over three versions of the experiment. The data for these three versions of the experiments are illustrated separately as 3A, 3B, and 3C.

Noticing the metaphor. In versions A and B, participants were asked to recall the missing word from the sentence "Crime is a _____ ravaging the city of Addison." In sum, 236 participants (45%) were able to recall the metaphoric frame (40% in version A and 55% in version B).

The influence of the metaphor was slightly, but not significantly, weaker among participants who were able to recall the metaphoric frame. The swing in the proportion of first responses that emphasized enforcement given the beast frame was smaller for participants who recalled the metaphor (8%) than for participants who could not recall the metaphor (13%). This was true not just of the aggregate data set, but of the individual versions as well: the frame was less effective among participants who could recall the

metaphor (*D*[N=524] = 6% and 10%, in versions A and B respectively) than among participants who could not (*D* = 13% and 15%). However, a logistic regression model that included a parameter for recalling the metaphor as well as an interaction term between the frame and recalling the metaphor was not significantly better than a model that included parameters for the frame and version, χ^2[2,519] = 1.11, p = .57.

In version C, participants were asked to identify the "most influential" part of the report. Eleven people (3%) included the metaphoric frame in the portion of the report that they copied and pasted into the text box; the small number of the participants that identified the metaphor prevents a comparison across groups to see whether the frame was more or less influential among them (however, see summary of Chapter 2 for further discussion of this issue).

Interestingly, in all three versions, people were more likely to remember or identify the virus frame (51% in 3A & 3B; 5% in 3C) than the beast frame (39% in 3A & B; 0% in 3C), {P(remember | virus) - p(remember | beast)}[N=524] = .13, $p < .01$ for Experiments 3A and 3B.

Individual differences. To test the effect of each individual difference measure, I used the model comparison method described in the previous experiment. The baseline model had predictors for version and frame.

Personal background. I found no main effect of age, gender or education on participants' top choice or weighted response set (age: χ^2[1,934] = .017, p = .90 and F [1,934] = .061, p = .81; gender: χ^2[1,934] = .020, p = .89 and F[1,934] = .022, p = .88; education (for Experiment 3C): χ^2[1,411] = .0010, p = .97 and F[1,411] = .34, p = .56).

There were also no interactions between these background measures and the frame (age: χ^2[2,933] = 1.06, p = .59 and F[2,933] = .80, p = .80; gender: χ^2[2,933] = .29, p = .87 and F[2,933] = .017, p = .98; education: χ^2[2,410] = 1.25, p = .54 and F[2,410] = 2.32, p = .10).

Political Affiliation. There were, however, systematic differences in response patterns by political affiliation. Republicans were more likely to suggest enforcement-

oriented responses (55%) than Independents (33%) or Democrats (32%) as their first choice approach for reducing crime in Addison, $\chi^2[1,934] = 24.90$, $p < .001$. A similar difference in emphasis across political affiliation was found in predicting participants' weighted response set, $F[1,934] = 25.11$, $p < .001$.

Further, Democrats and Independents were numerically more likely to be systematically influenced by the frame than Republicans (D[N=939] = .12, .13, and -.03 for Democrats, Independents, and Republicans, respectively) (see Figure 7). However, this difference was not significant in predicting participants' top choice, $\chi^2[1,933] = 1.72$, $p = .19$ or weighted response set, $F[1,933] = .32$, $p = .57$. Grouping Independents and Democrats and comparing them to Republicans made this comparison marginally significant in predicting participants' first choice, $\chi^2[1,933] = 2.84$, $p = .09$, but not their entire response set, $F[1,933] = 1.52$, $p = .22$.

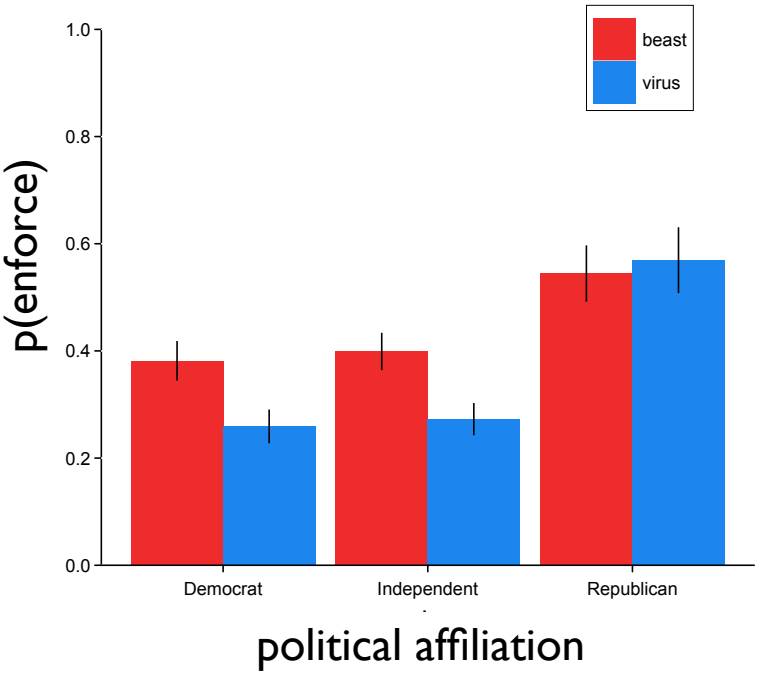

Figure 7: Results of Experiment 3. Democrats', Independents', and Republicans' top crime-reducing suggestion across Experiments 3A, 3B, and 3C.

Personality. With regard to the dispositional measures, I will only report analyses on data from Experiment 3C, which is the only version of the experiment that included

the fascism and need-for-cognition questionnaires. For this data set, the first principal component, professionalism, was not significant in predicting participants first choice, χ^2 [1,412] = .91, p = .34. It was marginally predictive of their weighted response set, F [1,412] = 3.64, p = .057. Participants who were high in professionalism were less likely to emphasize enforcement. There was no interaction between this dimension and the frame in predicting participants' first choice, χ^2[2,411] = .94, p = .63, or in predicting participants' entire response set, F[2,411] = 1.88, p = .15.

Participants with high scores on the second principal component, dogmatism, were more likely to emphasize enforcement-oriented approaches to the crime problem both in their first choice, χ^2[1,412] = 22.30, $p < .001$, and in their entire response set, F [1,412] = 36.43, $p < .001$. There was no interaction between social conservatism and the metaphoric frame, χ^2[1,411] = .07, p = .79, and F[1,411] = 1.36, p = .24.

The results of the analysis of this experiment, as well as those that follow, reveal that most of the variance associated with the dispositional measures is captured by the f-scale. As a result, for ease of interpretability, from now on I will run and report analyses with participants' f-scores rather than principal components. Participants who had a high score on the f-scale were more likely to emphasize enforcement in their first choice, χ^2 [1,412] = 17.79, $p < .001$, and in their entire response set, F[1,412] = 35.28, $p < .001$. There was no interaction between f-score and the metaphoric frame, χ^2[1,411] = .14, p = . 70, and F[1,411] = .33, p = .56.

Best Model. The best fitting model for Experiment 3C for predicting both dependent measures included main effects for frame, dogmatism, and social conservatism and no interaction terms. In addition, a regressor for political affiliation significantly predicted participants weighted response set. All regressors except the one for the virus frame (and the intercept) predicted enforcement (see Table 2).

regressor	logistic			linear		
	β	*se*	*p*	*β*	*se*	*p*
intercept	-3.45	0.61	< .001	-175	27.67	< .001
frame(virus)	-0.65	0.26	< .05	-24.34	12.12	< .05
political affiliation				28.67	9.20	< .01
f-score	0.074	0.018	< .001	4.23	0.84	< .001

Table 2: The best fitting regression models for Experiment 3C for predicting participants' top choice crime-reducing suggestion (logistic) and the weighted response set (linear). Participants who read the virus framing were less likely to emphasize enforcement; participants high in conservatism, and/or authoritarian tendencies (f-score) were more likely to emphasize enforcement.

Discussion

In Experiment 3 the critical finding from Experiments 1 and 2 was replicated when participants were given a set of response options to consider and rank. Both their first choice response option and the weighted ordering of their response preference showed a systematic influence of the frame. This replication is important because it rules out the possibility that the metaphor serves only as a weak prime in making some response options more available than others. It is also more representative of the types of policy decisions people make in the real world -- for instance, when choosing among multiple options in a voting context. However, it should be noted that the influence of the metaphor may be smaller in Experiment 3 (on the order of 11-14%) compared to Experiments 1 and 2 (on the order of 18-22%). This suggests that seeing a space of possible response options may wash away the prime for some small proportion of participants.

In addition, Experiment 3 uncovered some of the individual differences that are relevant to this effect. I found, predictably, that Republicans were more likely to

emphasize enforcement in their suggestions than Democrats and Independents and there was some indication that they may be less likely to be swayed by the metaphor.

I also found, not surprisingly, that people high in authoritarian tendencies (i.e., people who scored high on the fascism scale) were more likely to emphasize enforcement in their crime-reducing suggestions. This dimension did not interact with the frame, suggesting that even people who are high in the f-scale are susceptible to this metaphor framing effect. No other dispositional measures reliably predicted participants' first choice or weighted response set.

One interesting puzzle in this data concerns the variability in the main effect of enforcement across Experiments 1-3 and within the different versions of Experiment 3. I will explore some possible explanations for the variability across experiments in the discussion at the end of the chapter. Within Experiment 3, there are three identifiable causes of this variability. The first relates to the specific set of response options that participants are asked to consider. In versions A and B, participants considered five response options, three of which were coded as enforcement whereas in version C, participants considered only four response options, two of which were coded as enforcement. Removing the option that described a neighborhood watch program in version C increased peoples' likelihood of emphasizing reform, both with respect to their first-choice suggestion (which increased by 31%), $\chi^2[1,938] = 100.07$, $p < .001$ and weighted response set (D[N=939] = 95.22), $F[1,938]=37.89$, $p < .001$.

The second factor contributing to this variability relates to the individual characteristics of the participants in the three versions of the experiment. Participants in the first two versions, A and B, were, on average, more conservative than participants in third version of the study, $t[937] = 1.96$, $p < .05$, which may help to explain why there is a drop in the tendency to emphasize enforcement in version C.

Finally, the third factor relates to the timeline for running these versions of the experiment. In general, as this dissertation has evolved, the political and social climates have changed, most noticeably with respect to the worsening economy. As a result, the response option that emphasizes economic welfare has become more popular in later

versions of the study. This effect may be particularly true of the Mechanical Turk population, as some proportion of participants may have turned to Turk to supplement their income as a consequence of the economic climate[9]. Fully, 55% of participants suggest improving economic welfare as the best way of approaching the crime problem in Addison in Experiment 3C (compared to 33% in versions A and B; although note that there was one extra response option in versions A and B).

While these three factors likely play a large role in modulating participants' baseline preference for enforcement- or reform-oriented solutions, they do not account for all of the variability in the main effect and there may be additional factors at play that cannot be identified from the available data. Most importantly, despite this variability, there is a consistent effect of the metaphoric frame on participants responses in the predicted direction.

One potential limitation of these studies concerns the temporal dynamic of the effect. It is possible that the metaphor frame only affects people in the short term, and that, over time, it may come to have less of an impact on how people think about crime in Addison. However, it is also possible that the metaphoric frame can affect not only how people propose solving the problem of crime, but also how they go about gathering information for future problem solving. If participants seek out information that is likely to confirm the initial bias suggested by the metaphor, this may be a mechanism for metaphors to iteratively amass long-term effects on people's reasoning (Plous, 1993). I test this possibility in Experiment 4.

Experiment 4

Rather than asking participants to make a crime-reducing suggestion as in previous studies, the task in Experiment 4 was to select some aspect of the city of Addison to investigate further (in preparation for making a crime-fighting suggestion in the future).

[9] Informed speculation based on participant comments.

Method

Participants. Experiment 4 was conducted online with participants recruited on Amazon Mechanical Turk. One hundred fifty two Turkers participated for pay, of whom 76 were female and 76 were male. Their ages ranged from 18 to 70, with a mean age of 31.85 (median = 29). Sixty-nine participants identified as Democrats, 58 as Independents, and 25 as Republicans.

Materials. The crime report used in Experiment 4 was the same as the crime report for Experiments 2 and 3. However, instead of submitting a crime-reducing suggestion, participants were asked to choose between four issues that related to crime as a source of additional information about the problem. They were told that this information should be used to help them make a more informed crime-reducing suggestion in the future. The instructions read as follows: "Now imagine that Addison has consulted you about the crime problem. You have the resources to investigate one of the following four issues. Please select one from the list below." The issues included: 1) the education system and availability of youth programs, 2) the economic system including the poverty level and employment rate, 3) the size and charge of the police force, and 4) the correctional facilities including the methods by which convicted criminals are punished. Choosing to gather additional information about the educational system or economic system was coded as a social reform category of response; gathering additional information about the police force or criminal justice system was coded as an enforcement and punishment category of response.

Design. The design of Experiment 4 was similar to that of the previous two experiments. Participants read the report and follow-up questions on separate screens. After answering the target follow-up question, participants were asked what aspect of the report was most influential in their choice. In this experiment, participants were asked

about their political affiliation, age, and gender, but they were not asked to fill out personality questionnaires.

Results

Overall, participants in Experiment 4 were more likely to gather information about reform options (67%) than enforcement options (33%), $\chi^2[1,N=152] = 17.79$, $p < .001$. However, as in Experiments 1-3, this emphasis on reform was more pronounced among participants who read the crime-as-a-virus metaphor (78%) than participants who read the crime-as-a-beast metaphor (60%), $D[N=152] = .18$, $p < .05$ (see Figure 8).

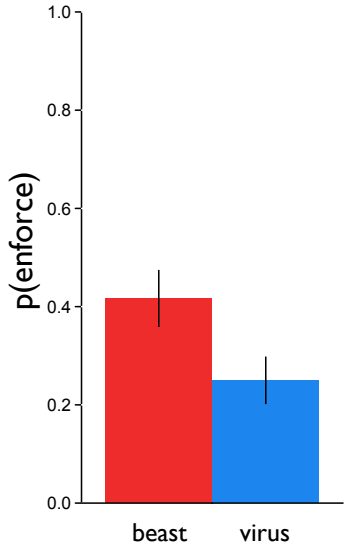

Figure 8: Results of Experiment 4. People who read that crime was a beast (as opposed to a virus) were more likely to gather information about the city having to do with punishment and enforcement, thereby confirming and strengthening the initial influence of the metaphor.

As in Experiments 1-3, when given the opportunity to identify the most influential aspect of the report, the vast majority of participants ignored the metaphor. Only 27 participants (15%) reported that the metaphor influenced their decision; in this experiment, participants were equally likely to identify "virus" (14, 18%) as they were "beast" (11, 15%), $\chi^2[1,151] = .23$, $p = .63$.

There were no differences by age, gender, or political affiliation and none of these measures interacted with the frame (age: $\chi^2[1,150] = .46$, $p = .49$ and $\chi^2[2,149] = 2.71$, $p = .26$; gender, $\chi^2[1,150] = .67$, $p = .41$ and $\chi^2[2,149] = .84$, $p = .66$; political affiliation: $\chi^2[1,150] = .18$, $p = .33$ and $\chi^2[2,149] = .15$, $p = .93$).

Discussion

In Experiment 4 I found that the effect of metaphorical framing persisted even when the context of participants' judgment was changed. In Experiments 1-3 participants were asked to make a crime-reducing suggestion while in Experiment 4 they were asked to forage for additional information. The results reveal that the metaphorical frame influences how people go about gathering information for future problem solving. People tended to seek additional information about the city that confirmed their initial (metaphor-induced) suspicion about how to solve crime.

In sum, the studies presented so far demonstrate that even minimal (one-word) metaphors can significantly shift people's representations and reasoning about important real-world domains. The specific cues that have been used in these experiments are metaphors. In several cases these metaphors have been instantiated by the word "virus" or "beast." I have speculated that people have schematic representations for solving literal virus and beast problems that are transferred into or instantiated within the domain of crime to inform how people think about these issues. If this is indeed the case, then the domain of crime may not be the only domain in which the virus and beast metaphors can systematically affect peoples' thinking. It could be the case that using the virus and beast metaphors could influence how people think about other important, structurally similar, policy issues like homelessness or education.

As a first step towards exploring this important question, in Experiment 5 I present participants with a somewhat different issue. Rather than having people read a metaphorically framed report about a crime problem, they read about a problem relating to high school drug use. The same metaphoric frames -- virus and beast -- were used to

describe this problem to see if these frames have a consistent influence on the way people think in different contexts. Will thinking about the high school problem as a virus make them more likely to suggest reform-oriented approaches than thinking about the problem as a beast?

Experiment 5

Method

Participants. Two-hundred twenty-one Turkers participated in Experiment 5 for pay. One-hundred thirteen participants were female and 108 were male. Their ages ranged from 18-65 (M = 31.74 and median = 25). Seventy-two identified as Democrats, 107 as Independents, and 42 as Republicans. The educational history of participants spanned the range from "some high school" to "doctorate" with most participants (194, 88%) having at least some college.

Materials and Design. In Experiment 5, participants read about a high school with a drug problem instead of reading about a city with a crime problem. The story was structurally very similar to the crime story and it was metaphorically framed as a "virus" or "beast". The story read as follows:

> Drugs are a {virus infecting/beast preying on} Addison High School. Five years ago the school was in good shape, with no obvious vulnerabilities. Unfortunately, in the past five years the school's defense systems have weakened, and it has succumbed to a major drug problem. During the past school year, there were more than 400 incidents involving drugs – up by more than 300 incidents from 2009. Teachers have reported a variety of behavioral problems: skipping or being late to class, disruptive behavior, inability to keep up academically. School officials fear that if they do not get the problem under control soon, even more serious issues may start to develop.

After reading the story, participants were asked to rank a set of four possible responses to the drug problem by dragging and dropping them into a response window. Two of the responses emphasized enforcement:

1) Increase drug searches and security camera surveillance in the school.
2) Increase punishments for students found with drugs, including expulsion.

The other two responses emphasized reform:

3) Expand the extracurricular opportunities at the school so that students have more school-sponsored activities to keep them busy.
4) Begin an educational campaign to teach students the dangers of drug use in high school.

After submitting their suggestion, participants were asked to identify which aspect of the story most influenced their suggestion. They did this by copying and pasting a subset of the report into a response window. Then, they were asked to complete three personality questionnaires and a set of background questions.

Results

Overall, participants were equally likely to suggest enforcement-oriented solutions (46%) as they were to suggest reform-oriented solutions (54%) to the drug problem, $\chi^2[1,N=221] = 1.63$, $p = .20$. However, as in the case of crime, participants were more likely to endorse enforcement-oriented solutions when they read that the problem was a “beast” (56%) than a “virus” (36%), $D[N=221] = .20$, $p < .01$ (see Figure 9).

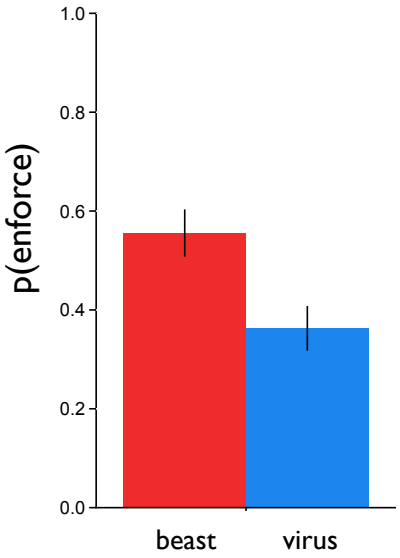

Figure 9: Results of Experiment 5. Metaphoric frames not only influenced how people thought about crime, but also how they thought about a drug problem. People were more likely to endorse enforcement-oriented responses when a the drug problem was framed as a beast than a virus.

Noticing the frame. Seven (3%) of the 221 participants identified the metaphor as influential in their decision. As in the case of crime, participants were numerically more likely to identify the "virus" frame (4.4%) than the "beast" frame (1.9%) as important.

Individual differences. Using logistic regression models to predict participants' first ranked choice revealed no main effect of age or interaction between age and the metaphoric frame, $\chi^2[1,219] = .019$, $p = .89$ and $\chi^2[2,218] = .15$, $p = .93$.

There was, however, an interaction between gender and the frame such that males given the beast frame were particularly likely to suggest enforcement-oriented solutions, $\chi^2[2,218] = 5.98$, $p < .05$ (there was no main effect of gender, $\chi^2[1,219] = .68$, $p = .41$). In fact, females did not show an effect of the frame in this experiment. There was also a main effect of education such that the more educated a person was, the more likely they were to endorse a reform-oriented approach to the drug problem, $\chi^2[1,219] = 4.60$, $p < .05$. There was no interaction between education and the frame, $\chi^2[1,218] = 18$, $p = .67$.

There was also a main effect of political affiliation such that Republican participants were more likely to endorse enforcement-oriented solutions (57%) than Independents (45%) or Democrats (40%), $\chi^2[1,219] = 4.70$, $p < .05$. There was no interaction between political affiliation and the frame, $\chi^2[1,218] = .16$, $p = .69$.

Finally, there was an effect of authoritarian tendencies. Participants who scored high on the f-scale were more likely to emphasize enforcement as a solution to the drug problem, $\chi^2[1,218] = 20.425$, $p < .01$. There was no interaction between f-score and the frame $\chi^2[1,217] = .177$, $p = .67$. Including this predictor came at the expense of the regressors for educational history and political affiliation.

Best Model. The best fitting logistic regression model for predicting participants' first choice drug-reducing suggestion included regressors for main effects of frame, gender, and f-score, as well as an interaction term between the frame and gender. Males and people with authoritarian tendencies were more likely to endorse enforcement-oriented approaches to the problem. However, males who read that the problem was a virus were more likely to endorse reform-oriented approaches to the problem (see Table 3).

regressor	logistic		
	β	*se*	*p*
intercept	-3.46	0.80	< .001
frame(virus)	-0.187	0.40	= .64
male	1.14	0.43	< .01
f-score	0.102	0.02	< .001
male * virus	-1.39	0.59	< .05

Table 3. Best fitting regression model for the results of Experiment 5. Being male or high in authoritarian tendencies (f-score) predicts an increased emphasis on enforcement; being male and receiving the virus frame, however, predicts an increased emphasis on reform.

Discussion

In Experiment 5, I found that peoples' thinking about a high school drug problem was affected by metaphoric frames in much the same way as the frames affected peoples' thinking about a city's crime problem. When the drug problem was described as a beast, people were more likely to suggest enforcement-oriented solutions than when it was described as a virus. This suggests that the metaphoric frames, virus and beast, are not specific to descriptions of crime *per se*, but also apply to descriptions of other structurally similar problems as well.

There are two interesting differences between the results of Experiment 5 and those of previous studies relating to the individual difference measures. First, unlike previous studies, in Experiment 5, I found an interaction between the frame and gender such that the frame was only influential among males. This may suggest that males and females have different thoughts about or personal experience with high school drug problems. Second, Republicans seem much more susceptible to making inferences that were consistent with the entailments of the frame in the drug domain than in the crime domain. Among Republicans, there was no numeric effect of the frame in their crime-reducing suggestions in Experiment 3C (1% swing in likelihood of suggesting enforcement given the beast frame) and a large numeric effect of the frame in their drug-reducing suggestions in Experiment 5 (13% swing in likelihood of suggesting enforcement given the beast frame). A post-hoc test of this difference in the proportion of congruent responses among Republicans (44% in Experiment 3C and 62% in Experiment 5) was marginally significant, $\chi^2[1,90] = 2.95$, $p = .086$. It may be the case that there is less of a "party line" approach to high school drug problems among Republicans than to crime problems more generally.

Summary of Experiments 1-5

Experiments 1-5 reveal that metaphoric frames can influence how people think

about important policy issues (see Figure 10). When a crime or drug problem was framed as a beast, relative to virus, people were biased towards suggesting enforcement-oriented solutions to the problem or to gather information about policies related to law enforcement. This was true using a variety of dependent measures and a variety of contexts.

The size of this effect was consistent across these experiments, in the range of a 10-22% swing in the probability of suggesting an enforcement-oriented approach to the problem attributable to the frames. Quantitatively, this swing is roughly equivalent to (or, in many cases, even greater than) the size of the predicted difference between Democrats and Republicans across these experiments.

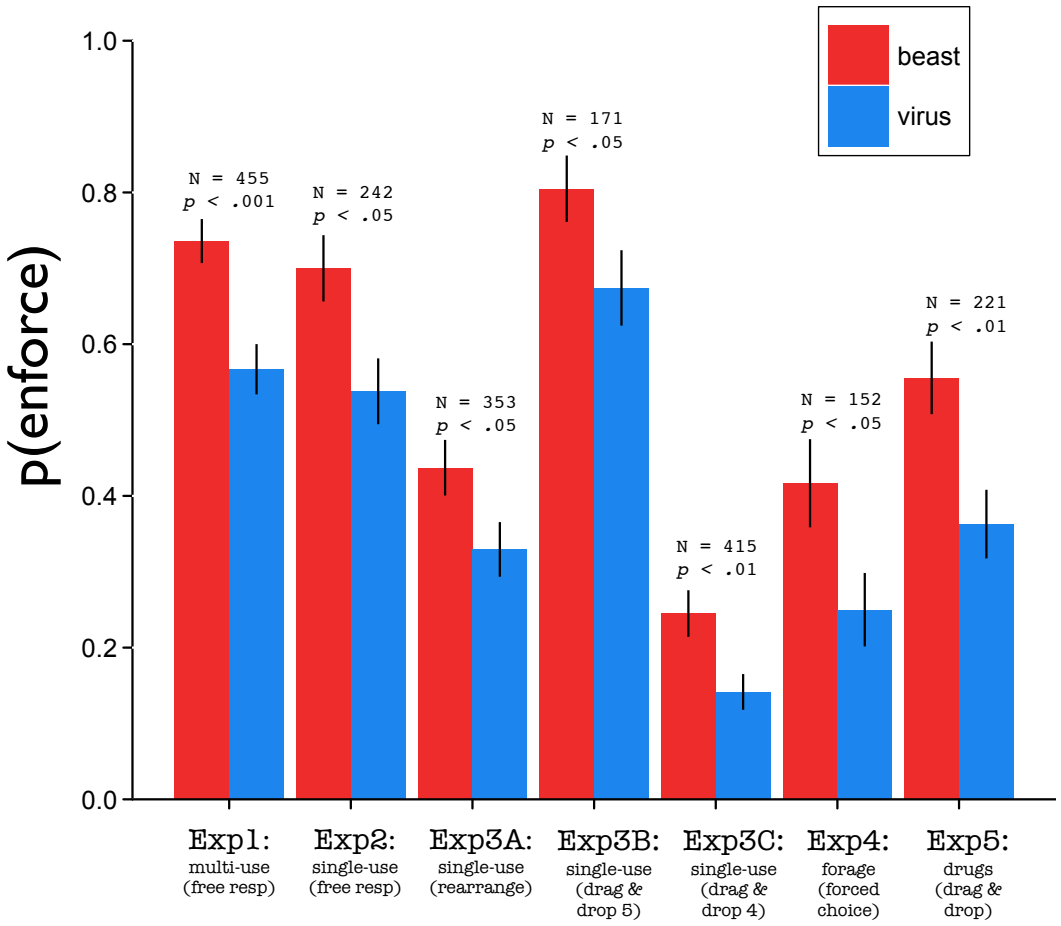

Figure 10: Results from Experiments 1-5 illustrating the basic metaphor framing effect. When crime (or drugs) is a beast, people are more likely to emphasize enforcement.

Interestingly, there was some indication that Republicans were more stable (i.e., less persuadable by the metaphoric frame) in their thinking about crime than Democrats or Independents. Collapsing data across Experiments 1-4 and comparing Republicans to non-Republicans reveals a marginally significant interaction between political affiliation and frame: non-Republicans, but not Republicans, were more likely to suggest reform when given the virus frame, $\chi^2[1,1326] = 2.74$, $p < .10$ (in this case the baseline model was fit with regressors for each experiment/version and a regressor for frame). In addition to the influence of political affiliation, there was a consistent effect of the fascism-scale individual difference measure: people with more authoritarian tendencies were more likely to endorse enforcement.

One striking feature of the results is that participants consistently did not identify the metaphor as influential in their thinking about crime or drugs. When participants were asked to identify the part of the report that most influenced their decision, they tended to point towards the numerical information. In sum, 65 of the 1485 participants in Experiments 1, 2, 3C, 4, and 5 (4%) identified the metaphor as influential. Collapsing across these data reveals that participants who identified the metaphor as influential were not more likely to have been influenced by it, $\chi^2[1,1483] = 2.54$, $p = .11$. Indeed, in Experiments 3A and 3B, in which participants were cued to recall the metaphoric frame from the surrounding first sentence of the report, people who remembered the frame showed a numeric tendency to be less influenced by the frame.

Of note, the virus metaphor seems to have been more noticeable or memorable than the beast metaphor. Among the 76 people who identified the metaphor as influential, 50 (66%) had read that crime was a virus. Similarly, among the 236 people who were able to remember the metaphor given the surrounding sentence as a cue, 133 (56%) had read that crime was a virus. It is unclear why this is. Both frames are highly conventional in the context of crime -- a Google search of “Crime is a virus” and “Crime is a beast” returns about 250,000 hits for each. It is possible that the virus metaphor is less promiscuous than the beast metaphor (i.e., less frequently used as a metaphor vehicle in

domains other than crime) making it somewhat less conventional overall, however, further work is needed to explore this possibility. Understanding this difference may be an interesting project for future inquiry.

The one major difference in the results across experiments relates to the main effect of enforcement. In Experiments 1, 2, and 3B people were more likely to endorse enforcement-oriented suggestions for reducing crime, while in Experiments 3A, 3C, and 4 people were, on the whole, more likely to suggest reform-oriented solutions. Some of the reasons for this difference were explored in the discussion of Experiment 3. There, I describe the influence of the response set (there were 3 enforcement-oriented responses and only 2 reform-oriented responses in Experiments 3A and 3B), differences in the political affiliation of participants across versions of Experiment 3, and the increasing popularity of thinking about the economy as a potential remedy for crime.

Here, I wish to identify one more source of variability across experiments. In general, it seems to be the case that people were more likely to suggest reducing crime through reform when they were asked to consider a specific set of response options. Comparing the baseline tendency to suggest enforcement or reform across experiments (1 & 2 vs. 3 & 4), I found that 64% of the responses emphasize enforcement when using a free response measure compared to 36% when participants are asked to consider a particular set of response options[10], $\chi^2[1,1787] = 135.34$, $p < .001$.

In discussing the results of these five experiments, I explored some of the possible mechanisms that may support the effect. First, results of Experiments 3-5 suggest that the effect is not driven by making some response options more available than others. If this were the case then seeing the space of response options in those experiments should wash away the prime. Instead, people seem to have schematic representations for virus and beast problems and these representations influence how people construct a representation of the target domain of crime (or drugs). Evidence for this view comes from (a) the norming study, in which participants' held highly consistent views about how to approach a virus or beast problem and (b) the fact that the virus and beast frames affect how people

[10] These percentages add up to 100 only by coincidence.

think about multiple domains.

There remain a number of questions about how this process unfolds and what metaphorically instantiating a schematic representation entails. I will investigate some of these questions in Chapter 3.

Chapter 3:
How does it Work?

In Chapter 2, five experiments demonstrated that metaphors can systematically shape the way we think. These experiments also revealed some clues about how this process may unfold. For instance, Experiments 3-5 showed that the effect is not the result of making specific response more available given a particular framing, since people show an effect of the metaphor when they are asked to rank pre-defined options. However, numerous questions about the underlying mechanisms and representations that subserve metaphoric reasoning remain.

One possibility is that the metaphor framing effect may be driven by simple lexical priming. Activation from the word "beast" may spread to words like "catch" and "cage" and activation from the word "virus" may spread to words like "investigate" or "reform," which get carried over into people's thinking about crime. If this is the case, then the bias shown in Experiments 1-5 is not an effect of the metaphor at all, but of the word "beast" or "virus." To test for this, in Experiment 6, I have people do a synonym listing task before reading the non-metaphorically framed crime report. If the framing effect is the result of simple lexical priming, then people should be systematically biased in this experiment as they were in Experiments 1-5. However, if the effect is driven by the activation of a more complex knowledge structure that is elicited by the metaphor, then there should be no effect of the lexical prime.

A second question that I explore in this chapter relates to the time-course of how metaphors influence the construal of complex issues. One possibility is that metaphors influence reasoning by providing people with a knowledge frame that structures how they think about the information given in the crime scenario (Bransford & Johnson, 1972). On this view, after being exposed to the metaphor, participants assimilate all further information they receive into the knowledge structure activated by the frame, instantiating any ambiguous information in a way that would be consistent with the metaphor. For instance, words like "defense systems" and "vulnerable" may be read one

way on the beast framing, possibly activating mental images of police officers and jails, and another way on the virus framing, possibly activating mental images of a city in a more general state of disrepair. If this is the case, if metaphors actively coerce incoming information, then metaphors should have the most impact when they are presented early. This was the structure of the report in Experiments 1-5: the metaphoric frame was presented in the first sentence of the report.

Alternatively, if metaphors simply activate a stored package of ideas and do not encourage the kind of active assimilation process described above, then they should be most effective when they are presented late in the narrative, as close to when people are asked to reason about a solution as possible. This way, the memory of the metaphor should be fresh and any knowledge activated by it should have the best chance to influence reasoning. This was the structure of the report in Experiment 7.

A related question that I explore in this chapter is -- what is the role of explicit, deliberative reasoning in the metaphor framing effect? In Experiments 1-5, there is some suggestion that people do not use the metaphor explicitly to inform their thinking or crime-fighting response. People tend not to identify the metaphor as influential and only about half of the participants in Experiments 3A and 3B could recall the frame when asked. However, asking people to identify the most influential aspect of the report or to recall the frame from memory are weak tests of the hypothesis that people are using the metaphor implicitly. People may in fact actively use the metaphor to guide their suggestion, but choose to point towards the numeric information in the report in order to seem rational and objective. Similarly, not being able to remember the frame only circumstantially suggests that people are not using it explicitly. It remains possible that people actively use a schematic representation of a virus or beast problem, or an abstraction of these problems, to guide their subsequent decision and that some simply cannot remember the precise word used in the metaphorical frame when asked at the end.

Further, if it is the case that people are only implicitly using the metaphor to reason in Experiments 1-5, it may still be possible that the underlying mechanisms that support metaphoric reasoning are qualitatively the same using an explicit reasoning

strategy. For instance, if metaphoric reasoning involves a mapping process similar to the mapping process that is thought to underlie analogical reasoning (e.g., Gentner, 1983; Gick & Holyoak, 1980) then people may show a similar, or even greater, influence of the metaphor when they are asked to explicitly use the metaphoric frame to guide their thinking.

I test these possibilities in Experiments 8 and 9. In Experiment 8 I present people with both metaphors and ask them to match a response option with each. On this task, people can compare the two source representations to find specific relationships between crime and a virus problem and crime and a beast problem. If the response options map on to distinct structural elements of crime, then making these structural elements "pop out" through comparison, should reveal an even bigger difference between the framing conditions.

In Experiment 9 I employ a different method to explore explicit metaphoric reasoning: I present people with a single metaphoric frame, ask them to interpret this frame, and then invite them to take the speaker's perspective in ranking the four crime-fighting options. On this method, people do not have the opportunity to compare the frames, so identifying the structural correspondences between the source and target domains may be more difficult to find. As a result, this experiment is a clearer test of whether people do (or can) use these metaphors to explicitly inform how they think about crime.

Finally, in Experiments 10 and 11 I investigate the claim that the metaphors serve to instantiate a schema or knowledge structure of the source domain. What is a metaphorical knowledge structure or schema? Is it a specific relation, a system of relations, or an elaborate mental model? In Experiment 10, I replace the noun metaphors "virus" and "beast" with specific relations: "spreading" and "overwhelming." If these relations alone can systematically influence peoples' thinking about crime then the function of the metaphoric frame may be to elicit these relations in the context of crime.

In Experiment 11, I test whether eliciting these particular relations is all that the metaphor does. That is, even if there is a correspondence between the results of the noun

and relational frame conditions in Experiment 10, there may still be interesting differences between these conditions. For instance, the metaphor may provide the reader with a rich pallet of relations to choose from. In the context of making a suggestion for reducing crime, "spreading" and "overwhelming" may come to mind. However, in another context, as in making a sentencing decision, other relational structure may be more salient.

Together, these studies help to shed light on the mechanisms through which metaphors influence our reasoning.

Experiment 6

In Experiment 6 I tested whether the influence of the metaphor observed in the previous chapter could have come about through simple spreading activation from lexical associates of the words "beast" and "virus." Perhaps simply hearing a word like beast, even outside of the context of crime, would activate associated words like "hunting" and "caging". These activated lexical associates might then color people's descriptions of how to solve the crime problem. To test for this possibility I dissociated the words "beast" and "virus" from the rest of the crime report. Before reading about crime in Addison, participants were asked to generate a synonym to the word "beast" or the word "virus" – thereby priming representations for a beast or a virus. They then read the same report about crime as in Experiments 2-4, but with the metaphorical word omitted ("Crime is ravaging the city of Addison"). Might a non-metaphorical lexical prime have the same effect as a metaphor?

Method

Participants. Experiment 6 was conducted online with 236 participants. Of these participants, 136 were female and 100 were male. Their ages ranged from 18 to 81, with a mean age of 29 (and median age of 26). Seventy-six reported an affiliation to the

Democratic Party, 48 reported an affiliation to the Republican Party, and 112 were Independent.

Materials. The crime report used in Experiment 6 was the same as the crime report used for Experiments 2-4 (i.e., the crime report that instantiated the metaphor with a single word), except that it did not contain a virus or beast metaphor. The first sentence of the report read: "Crime is ravaging the city of Addison."

Design. Before participants read the crime report, they were shown the word "beast" or the word "virus" and were asked to "list a synonym" for it. After completing this task, they were presented with the paragraph on crime in Addison on a separate screen. They were then asked to suggest a method for reducing crime and to interpret the role of a police officer in Addison. Participants responded to both of these questions in free response form. They were not asked to identify the part of the report that was most influential in their suggestion since there was no metaphor frame in the report.

Experiment 2, in which the frame was a single word and participants crime-reducing suggestions were in free-response form, was run concurrently with Experiment 6. Comparisons will be made between the results of Experiment 2 (metaphor frame condition) and Experiment 6 (synonym listing condition) below.

Coding. Answers to the free response questions were coded as they were in Experiment 2, into two categories: "enforce" and "reform." Fifteen crime-fighting suggestions (6%) and 21 police officer interpretations (9%) did not fit into either category. In every case this was because the response lacked a suggestion or interpretation.

Answers to both of the free response questions were coded blindly by two coders. Inter-rater reliability was high for both: Cohen's *K* for crime-reducing suggestions was .87 ($p < .001$); Cohen's *K* for interpretations of the role of a police officer was .84 ($p <$.

001). All disagreements between the coders were resolved between them before analyzing the data.

Results

The synonyms that participants listed were analyzed to ensure that the lexical prime had the intended effect. Of the 124 participants in the crime-as-beast condition, all except one listed a synonym of "beast". The modal response was "animal", but others included "monster," "mongrel," "invader," etc. The single respondent who did not list a synonym to "beast" instead wrote "I forget what a synonym is." This participant's subsequent responses were omitted from the analyses reported below. Of the 112 participants in the crime-as-virus condition, all listed a synonym of virus. In this case, the modal response was "disease", but others included "bug," "cold," "sickness," "illness," etc.

Overall, participants were significantly more likely to suggest enforcement (64%) than reform (36%), $\chi^2[1,N=221] = 18.0$, $p < .001$. There was no difference in this main effect between the metaphor frame condition (61% suggested enforcement) and the lexical prime condition, $\chi^2[1,461] = .79$, $p = .37$.

There was also no difference between participants who were lexically primed with "beast" (64% suggesting enforcement and punishment) and those who were lexically primed with "virus" (65%), $D[N=221] = -.01$, $p > .8$. This pattern was marginally different from that of the metaphor frame condition, in which people were more likely to be enforcement-oriented when they read the beast framing (70%) than the virus framing (54%), $\chi^2[1,459] = 3.41$, $p = .065$.

Further, disambiguating the responses that called for an increase to the police force did not differentiate the groups in Experiment 6. Sixty-eight of the responses (31%) were disambiguated. Of these, 29 (43%) interpreted the role of a police officer as a crime deterrent, 37 (54%) interpreted the role of a police officer as a law enforcer or punisher, and two responses could not be disambiguated.

Importantly, disambiguating these responses did not reveal a difference between the two synonym groups: participants who were lexically primed with "virus" were no more likely to suggest enforcement (50%) than those who were lexically primed with "beast" (51%), D[N=219] = .04, $p > .9$. However, it did reveal a difference between this pattern of results and those of the metaphor framing condition, in which participants were more likely to suggest enforcement given the beast frame (63%) than the virus frame (41%), χ^2[1,457] = 5.42, $p < .05$ (see Figure 11).

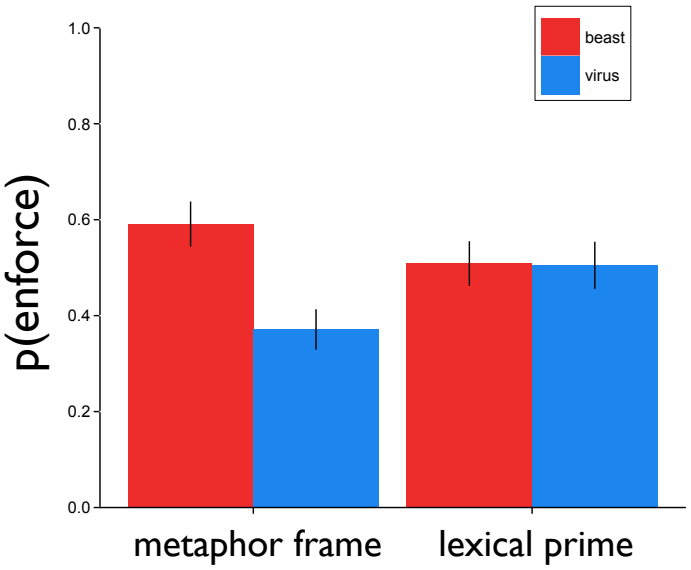

Figure 11: Results of Experiment 6. The metaphor frame, but not the lexical prime influenced peoples' disambiguated crime-reducing suggestions.

Discussion

In Experiment 6 I tested whether the influence of the metaphor observed in the studies from the first chapter could have come about through simple spreading activation from lexical associates of the words "beast" and "virus." To test this, I dissociated the words "beast" and "virus" from the story, so that they could act as non-metaphorical lexical primes. These disconnected lexical primes did not yield differences in people's

crime-fighting suggestions. These results suggest that metaphors act as more than just isolated words – their power appears to come from participating in elaborated knowledge structures.

Additionally, the results of Experiment 6 shed some light on this population's baseline preference for reducing crime. That is, in Experiment 2 it might have been the case that participants had a general preference for reducing crime through enforcement and that it was the crime-as-virus frame alone that shifted peoples' responses. The results of this lexical prime experiment, however, suggest that the population does not seem to favor either of the two crime-reducing suggestions absent a metaphoric frame and that both frames are influential.

Experiment 7

In Experiment 7 I investigated the time-course of how metaphors influence people's construal of and reasoning about problems. One possibility is that metaphors influence reasoning by instantiating a knowledge frame that structures subsequent information. After being exposed to the metaphor, participants may assimilate all further information they receive into this knowledge structure, instantiating any ambiguous information in a way that would be consistent with the metaphor. For example, words like "vulnerabilities," "defense," "weakened" may take on different meanings depending on whether they are understood in the context of viruses or beasts (Bransford & Johnson, 1972; Ortony, 1978). If this is the case, if metaphors actively coerce incoming information, then metaphors should have the most impact when they are presented early, such that their impact can accumulate in the course of assimilating further information.

Alternatively, if metaphors simply activate a fossilized package of ideas and do not encourage the kind of assimilation process described above, then they should be most effective when they are presented late in the narrative, as close to when people are asked to reason about a solution as possible. This way, the memory of the metaphor should be fresh and any knowledge activated by it should have the best chance to influence

reasoning. In Experiment 7, I repeated the design of Experiment 4 (in which participants forage for additional information about a city's crime problem), but moved the metaphorical frame so that instead of being the first sentence in the crime report it was the last.

Method

Participants. One-hundred seventy-two Turkers participated for pay. Among these Turkers, 157 were female and 96 were male. Their ages ranged from 18 to 81, with a mean age of 31 (and median age of 28). Fifty-five reported an affiliation to the Democratic Party, 40 reported an affiliation to the Republican Party, and 77 were Independent. This background information was solicited from participants after they completed the experiment.

Materials and Design. The materials and task in Experiment 7 were identical to those of Experiment 4 except, instead of presenting the metaphor frame at the beginning of the report, I presented the metaphor frame at the end of the report, as shown below. All other aspects of the design were identical to Experiment 4. The paragraphs used were:

> Five years ago Addison was in good shape, with no obvious vulnerabilities. Unfortunately, in the past five years the city's defense systems have weakened, and the city has succumbed to crime. Today, there are more than 55,000 criminal incidents a year - up by more than 10,000 per year. There is a worry that if the city does not regain its strength soon, even more serious problems may start to develop. Crime is a {beast / virus} ravaging the city of Addison.

Experiment 4, in which the frame was a single word and participants foraged for additional information about Addison, was run concurrently with Experiment 7. Comparisons will be made between the results of Experiment 4 (metaphor-first condition) and Experiment 7 (metaphor-last condition) below.

Results

Overall, participants in Experiment 7 were more likely to gather information relating to the city's social situation (67%) than the criminal justice system (33%), χ^2 [1,N=172] = 19.55, $p < .001$. This main effect is consistent with what was found in the metaphor-first condition, χ^2[1,322] = .002, $p = .96$.

However, unlike the metaphor-first condition, there was no effect of the metaphorical frame when the metaphor was presented at the end. In the metaphor-last condition, participants who were read that crime was a beast were about equally likely to gather additional information about the city's social situation (69%) as participants who read that crime was a virus (64%), D[N=172] = -.05, $p > .4$. This pattern was significantly different from the metaphor-first condition, χ^2[1,320] = 4.31, $p < .05$. That is, significantly more participants were influenced by the metaphor when it was presented at the beginning of the report than at the end of the report (see Figure 12).

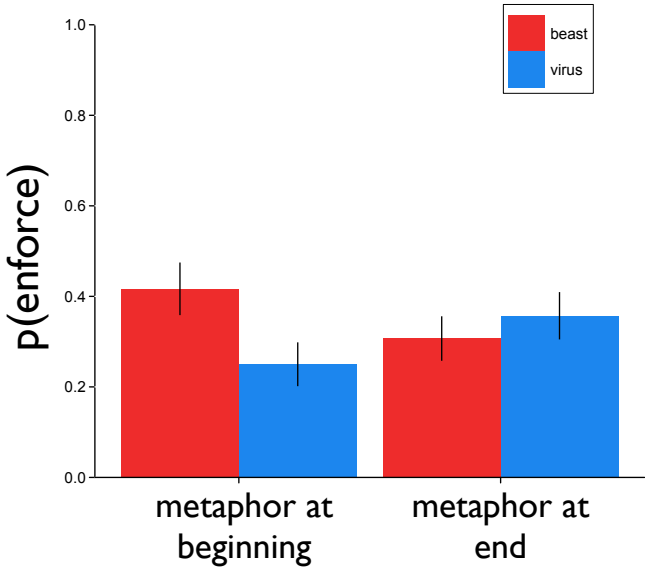

Figure 12: Results of Experiment 7. Including the frame at the beginning of the report, but not the end, influenced peoples' crime-reducing suggestions.

Consistent with previous experiments, when participants were given the opportunity to identify the most influential aspect of the report, the vast majority ignored the metaphor. Only 18 participants (10%) reported that the metaphor influenced their decision. Participants were about equally likely to identify the beast frame (12%) as they were to identify the virus frame (9%).

Discussion

In Experiment 7, consistent with previous work on meaning instantiation, I found that the metaphors were most effective when they were presented early in the narrative and were then able to help organize and coerce further incoming information. For example, Bransford and Johnson (1972) demonstrated that a procedural description of washing clothes was understood and remembered best when participants knew the topic of the passage before they heard the description. When the topic was given at the end of the passage or not at all, participants reported being unable to make sense of what they had heard and were able to recall few details of the description on a memory test.

While the crime passage I used was clearly not as ambiguous as the procedural description of washing clothes used by Bransford and Johnson, it did contain many words and phrases that would likely be interpreted differently in the different contexts represented by the metaphoric frames. For instance, in the context of an attacking beast the meaning of the words ‘‘vulnerable’’ and ‘‘defense system’’ may be different from what the same words would be taken to mean in the context of a spreading virus. Previous work has demonstrated that contextual cues can strongly influence how people interpret seemingly unambiguous text (Elman, 2004 ; Ortony, 1978; Rumelhart, 1979).

These results are particularly striking since in Experiment 7, the metaphorical frame appears in much closer proximity to the measure of interest. It would have been reasonable to predict that a metaphorical frame that is more fresh in mind should have the

largest effect. Instead, the way a metaphorical frame is integrated into the narrative appears to be more important.

This finding also helps allay a possible worry about the findings in the previous experiment. In Experiment 6, I moved the words “virus” and ”beast” out of the crime story, and asked participants to generate synonyms to these words before they read about crime. When the words appeared in this way as disconnected lexical primes, they had no influence over people’s crime-fighting suggestions. Of course, one possibility is simply that taking the words out of the narrative also made them more distant in time from the measure of interest. Results of Experiment 7 suggest that it is integration at the right point in the narrative rather than simple temporal distance that modulates the effect of the metaphor. In Experiment 7, the words “virus” and “beast” occurred immediately prior to the measure of interest, and yet had no effect.

Finally, the fact that the metaphor did not have an effect when it was read immediately before participants made their crime-fighting judgment suggests that the influence of the metaphor may be implicit. If people were explicitly using the metaphor to reason about the problem, one might expect the frame to have the same, if not a bigger, affect on people’s reasoning when encountered at the end of the story, immediately preceding their judgment.

In concert with data from Experiments 1-5, in which, for instance, only a small minority of participants identify the metaphor as important, these findings bolster the claim that the metaphor implicitly affects how people reason about crime. Nevertheless, these results do not rule out the possibility that some people deliberately used the metaphor to reason about crime. For instance, it may be the case that people only use the metaphor explicitly when it is encountered first, at the beginning of the report. As a result, a stronger test of this question is needed to assert that this metaphoric reasoning effect is the result of implicit processing.

In Experiments 8 and 9 I manipulate the degree to which people are forced to attend to and reason with the metaphor in order to explore whether the metaphor framing effect is the result of people deliberately reasoning about the relationship between the

source and target domains.

In Experiment 8 I present participants with both metaphors and ask them to match one response option to each. In this task, people have the opportunity to compare and contrast the frames, which should make the structural correspondences and incongruities between the target and source domains more salient (Christie & Gentner, 2010; Clement & Gentner, 1991; Gentner, 1983; Markman & Gentner, 1993). If this process -- identifying structural correspondences and incongruities between the target and source domains -- is important for the metaphor framing effect, then comparison should increase the influence of the frame. This finding would also demonstrate that people, in principle, can use the metaphors explicitly to reason about crime.

Experiment 8

Method

Participants. Two-hundred twenty-six Turkers participated in exchange for pay. One-hundred four of the participants in Experiment 8 were female and 122 were male. Their ages ranged from 18 to 65 with a mean of 30.88 (median = 25). Eighty-two of the participants identified themselves as Democrats, 104 as Independents, and 40 as Republicans. Participants' educational history spanned the range from "some high school" to "doctorate" with most (200, 88%) having "some college."

Materials and Design. Participants read the standard crime report minus the metaphor frame ("Crime is ravaging the city of Addison"). After reading the report, they were told that two city officials were engaged in a debate over how to solve the crime problem in the city, and that these two officials were using contrasting metaphors to support their message. Both metaphors were presented on this screen as well as the standard four response options. Participants were instructed to match the response options

with the candidates (they did not know anything about the candidates except for their metaphor preference). The instructions for this section read as follows:

> The city's officials know that they have to change certain policies in response to crime, but they aren't sure which policies to change or how much to change them. Two of the city's officials are leading this debate and they tend to talk about the crime problem in different ways, using different metaphors.
>
> One metaphorically compares crime to a virus, as in "Crime is a virus ravaging the city of Addison."
>
> The other metaphorically compares crime to a beast, as in "Crime is a beast ravaging the city of Addison."
>
> If you had to guess, which of the four crime-reducing options listed under the "Items" menu is supported by each of the officials. Pick the “Item” that you think is most consistent with the virus metaphor, then drag and drop that item into the "Virus" box. Do the same with the proposal that is most consistent with the beast metaphor and drag and drop it into the "Beast" box (1 in each).

After submitting these responses, participants were asked to fill out the three personality questionnaires and answer a series of background questions.

Results

In this experiment, unlike previous experiments, participants provided two data points each: they associated one crime-reducing suggestion with the virus frame and one with the beast frame. Because these data were not sampled independently -- since (a) participants were aware of both of their choices and (b) a given response option could not be associated with both frames -- the analyses for Experiment 8 will proceed somewhat differently from those of previous experiments.

Overall, suggestions that emphasized enforcement (63%) were numerically more popular than suggestions that emphasized reform (37%). This was true of responses that

were matched to the beast frame (87%), but not true of responses that were matched to the virus frame (40%) (see left panel of Figure 13).

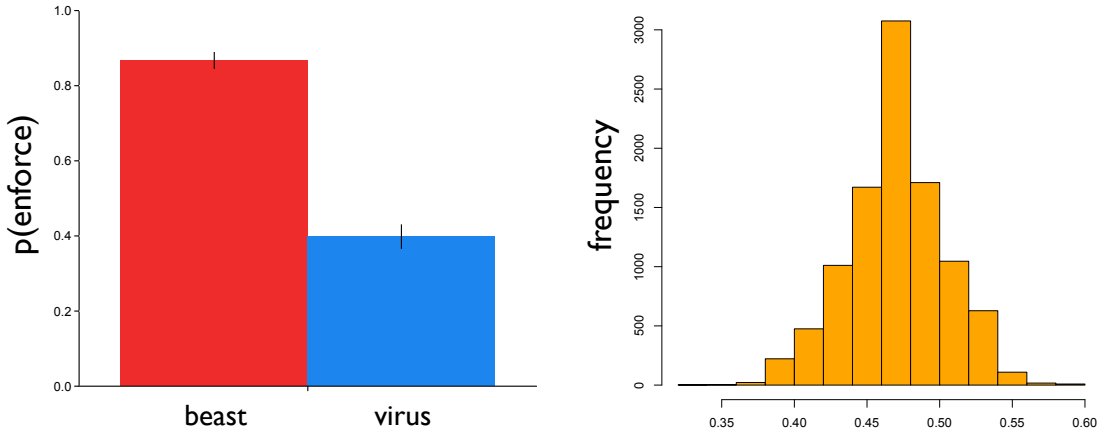

Figure 13: Results of Experiment 8. The plot on the left illustrates the proportion of responses that emphasize enforcement by frame. The plot on the right illustrates a bootstrapped distribution of the increased likelihood of emphasizing enforcement given the beast frame.

Given that participants provided two data points each, one way to investigate a possible systematic effect of the frames is to quantify the number of "congruent" responses that each participant provided. In this analysis "enforcement" responses were coded as congruent with the beast frame and "reform" responses were coded as congruent with the virus frame. If a participant associated a "reform" suggestion with the beast frame or an "enforcement" suggestion with the virus frame, that association was coded as incongruent.

By chance, one would expect an equal proportion of participants to submit two congruent responses (1/2 * 2/3 = 1/3), one congruent and one incongruent response (2* [1/2 * 1/3] = 1/3), and two incongruent responses (1/2 * 2/3 = 1/3), which, for a sample of 226, would predict that approximately 75 participants would be in each category. If, however, significantly more than half of the responses are matched 'congruently' with the

frame (i.e., more than what we would expect by chance), then there must be some systematic relationship between the frames and responses. In fact, I found that 129 (57%) participants submitted two congruent responses, 74 (33%) submitted only one congruent response, and 23 (10%) submitted no congruent responses. This observed distribution was significantly different from what one would expect by chance, $\chi^2[2,N=226] = 74.61$, $p < .001$.

Of the 74 participants who gave only one congruent response, significantly more submitted two enforcement-oriented responses (91%) than two reform-oriented responses (9%), $\chi^2[1,N=74] = 48.65$, $p < .001$. This is evidence of a response bias towards enforcement.

In addition to computing the distribution of congruent responses by participant, I bootstrapped a distribution of the difference between the proportion of enforcement-oriented responses in the beast and virus conditions. I did this by randomly assigning half of the participants to contribute their "virus" response and the other half to contribute their "beast" response. Then I computed the proportion of enforcement-oriented responses in the two groups and subtracted the proportion in the "virus" group from that of the "beast" group (i.e., p(enforcement | beast) - p(enforcement | virus), where 113 participants were randomly assigned to be in the virus group and the rest were assigned to be in the beast group). I repeated this procedure 10,000 times to find the distribution of this difference and to test whether it was significantly different from 0 (see right panel of Figure 13).

The mean of this distribution was .47 and the standard deviation was .034. At the low end of the range, the distribution did not include 0 (it went from .35 to .59) suggesting that that people were much more likely to associate enforcement-oriented solutions with the beast framing than the virus framing, $t[9999] = 1389.69$, $p < .001$.

Individual Differences. Surprisingly, Republicans were numerically more likely to have selected crime-reducing suggestions that were congruent with both frames (70%) than Democrats (55%) or Independents (54%) and numerically less likely to submit only

one congruent response (18%) than Democrats (38%) or Independents (35%). When Independents and Democrats were grouped in contrast to Republicans, this difference was marginally significant, $\chi^2[2,N=226] = 5.13$, $p = .077$.

This is interesting because in previous experiments Republicans tended to be the least likely to be swayed by the metaphor. This suggests that Republicans may be more effective at comparing the frames. However, it also raises the possibility that Republicans are able to differentiate their own options from those suggested by the frame. In this experiment, unlike previous ones, participants were asked to take someone else's perspective in making their crime-fighting suggestion. Therefore, it may be the case that Republicans actively reject the virus framing of crime in Experiments 1-5 but are capable of providing congruent responses when reasoning from someone's else's shoes in the current experiment.

In addition to finding differences by political affiliation, I found that people who were high in authoritarian tendencies were less likely to submit congruent responses, $F[1,224] = 4.838$, $p < .05$. When predictors for both political affiliation and f-score were included in the model, they were both significantly predictive (see Table 4). This was the best model fit to these data.

Predictor	*β*	*se*	*p*
intercept	1.99	0.199	< .001
Republican	0.25	0.124	< .05
f-score	-0.02	0.007	< .001

Table 4: Results of Experiment 8. Summary of a linear regression model in which regressors for political affiliation and f-score predict the number of congruent crime-reducing suggestions that a given participant submitted (0, 1, or 2). Republicans were more likely to suggest two congruent crime-reducing suggestions; people high in authoritarian tendencies were less likely to suggest two congruent crime-reducing suggestions.

This pattern of results is interesting because Republicans tend to be much higher in authoritarian tendencies ($M = 34.95$, $sd = 6.15$) than Democrats ($M = 27.57$, $sd = 6.50$) or Independents ($M = 29.06$, $sd = 6.60$) and there is a high correlation between the numerical coding of political affiliation and f-scores among participants in this experiment: $r[224] = .430$, $p < .001$. The fact that the regressors for Republican and f-score make different predictions about participants' likelihood of associating both frames with congruent crime-reducing approaches therefore suggests that this task may dissociate participants who are politically or economically conservative from those who are socially conservative.

In addition, females (72%) were marginally more likely than males (50%) to submit two enforcement-oriented responses, $\chi^2[1,225] = 3.2447$, $p = .072$. No other individual differences predicted participants' likelihood of submitting two congruent responses (age: $\chi^2[1,225] = 1.29$, $p = .26$; education: $\chi^2[1,225] = .18$, $p = .67$; gender: $\chi^2[1,225] = .14$, $p = .71$).

Discussion

Participants in this experiment where given both metaphoric frames for crime and were asked to associate one response option with each of these frames. In doing so, they were systematically more likely to associate reform-oriented responses with the virus metaphor and enforcement-oriented responses with the beast metaphor. The context of this task and the format of the data prevent direct comparisons between these results and the results of previous experiments; however, numerically, the magnitude of the difference by frame is far larger in this experiment than in previous versions ($D = .47$ in Experiment 8 and $D < .25$ in Experiments 1-5).

There are three interesting differences between this task and the tasks used in previous experiments. The most obvious is that participants were presented with both frames in this experiment and only one frame in previous experiments. The second difference between this experiment and previous ones is that participants' attention was

directed towards the frames. The word “metaphor” was used to describe the frames in this experiment (i.e., “Two of the city's officials are leading this debate and they tend to talk about the crime problem in different ways, using different metaphors.”) and, unlike previous versions, the metaphors were presented outside the context of the report on the same screen as the response options. Finally, unlike previous versions of the experiment participants were asked to take the speakers perspective in making their crime-fighting judgment (i.e., “If you had to guess, which of the four crime-reducing options listed under the "Items" menu is supported by each of the officials.”).

It is possible that any of these changes was responsible for the increased magnitude of the effect. For instance, seeing both frames may have encouraged participants to make comparisons between them. This may have highlighted the various ways in which crime is similar to and different from the virus and beast problems and facilitated the matching process. Alternatively, it might be the case that attending to the metaphor is sufficient to facilitate metaphoric reasoning and that taking the speaker’s perspective ensures that people make a frame-consistent suggestion. The degree to which this experiment sheds light on the mechanisms that underlie the metaphor framing effect in Experiments 1-5 will depend on which of these changes are responsible for increasing the magnitude of the effect.

In Experiment 9, I test this second possibility: whether drawing people’s attention to a single frame and asking them to reason from the speaker’s perspective are sufficient for yielding an enhanced metaphor framing effect. If so, then it might be the case that explicit and implicit metaphoric reasoning rely on the same underlying processes. For instance, if the effectiveness of the metaphor is the result of an active mapping process between the source and target domains, then one might predict that making people aware of the metaphor would facilitate the process of finding structural correspondences between the source and target domains (Blanchette & Dunbar, 2000; Chen, 1996; Gentner, Bowdle, Wolff, & Boronat, 2001; Gick & Holyoak, 1980; Hall, 1989; Holyoak & Koh, 1987).

Similarly, if there is a pragmatic component to processing these metaphors (e.g., Searle, 1979), then one might predict that people would be more likely to find meaning in the metaphor when it is given special attention.

Alternatively, however, if the standard metaphor framing effect relies on a more implicit process that involves a dynamic interaction between the frame and the content of the crime report, then making people aware of the metaphor may not have a facilitative effect. Indeed, making people aware of the metaphor may interfere with the effectiveness of the frame.

Experiment 9

Method

Participants. Three-hundred thirty-five Turkers participated in Experiment 9 in exchange for pay. One-hundred sixty-four of these participants were in a replication condition (their data was presented in Experiment 3C) and 171 were included in the explicit condition. Ages of the participants ranged from 18 to 75 with a mean of 31.36 (median = 25); 179 were female and 156 male. One-hundred twenty-six of the participants identified as Democrats, 156 as Independents, and 53 as Republicans. The educational history of the participants spanned the range from "some high school" to "doctorate" with most participants having at least some college (283, 84%).

Materials. There were two conditions in this experiment: one was a replication condition and one included a novel manipulation. The critical materials in both conditions were the same. They included the standard crime report and solution set (with 4 drag & drop options).

Design. In the explicit condition, participants were shown the first sentence of the crime report (i.e., the sentence that included the metaphoric frame) out of context, before

they had read the crime report. They were told that it was a metaphor and were asked to describe what it meant in a text box that was provided. The instructions for this portion of the experiment read as follows:

> Imagine that someone used the following phrase to describe a city's crime problem:
>
> "Crime is a {beast/virus} ravaging the city"
>
> Please describe (briefly) what you think the speaker is seeking to convey with the metaphor.

On the following screen they were given the complete crime report and told to "Imagine that the speaker said the following." Then, they were shown the four crime-reducing suggestions and asked to take the speaker's perspective and rank order the methods for reducing crime in the city. The instructions for this section read: "Do you think the person who used the crime metaphor would prefer some of the following crime-reduction programs over others? Please read each of the proposals listed below carefully and consider which ones you think the speaker would prefer."

After submitting their crime-fighting suggestion, participants were asked to complete three personality questionnaires and a set of background questions.

Results

As described in Experiment 3C, the results of the replication condition reveal an overall bias towards the reform-oriented suggestions (79%; $\chi^2[1,N=164] = 58.88$, $p < .001$) and a systematic influence of the frame: both with respect to participants' first choice ($D[N=164] = .137$, $p < .05$) and their weighted set of responses ($D[N=164] = 50.46$, $t[162] = 2.60$, $p < .05$).

This pattern of results is not similar to those of the explicit condition, in which there was a bias in the opposite direction, towards enforcement-oriented solutions (73%), $\chi^2[1,N=171] = 34.67$, $p < .001$, and no systematic influence of the frame with respect to

participants' first choice (D[N=171] = .002, $p > .9$) or weighted response set (D[N=171] = 10.09, t[169] = .47, $p = .64$).

The difference in the baseline tendency to endorse enforcement-oriented solutions was statistically significant between the implicit and explicit frame conditions on both measures: in predicting participants first-choice, χ^2[1,332] = 92.303, $p < .001$, and in predicting their weighted response sets, F[1,332] = 216.13, $p < .001$. However, there was no interaction in the effect of the frame across conditions on either measure, χ^2[1,331] = 1.63, $p = .20$ and F[1,331] = 1.17, $p = .28$ (see Figure 14).

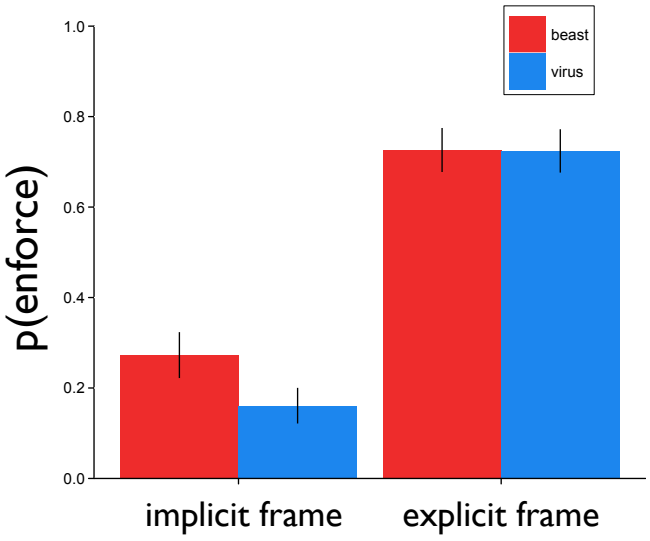

Figure 14: Results of Experiment 9. In the standard metaphor framing condition (implicit) people who read the beast metaphor for crime relative to the virus metaphor for crime were more likely to endorse enforcement-oriented solutions. When participants' attention was directed at the metaphor and they were asked to take the speaker's perspective in ranking the crime-fighting suggestions (explicit), this effect went away.

Metaphor Interpretations. Qualitatively, interpretations of the crime-as-a-beast metaphor emphasized the urgency and scale of the crime problem whereas interpretations of the crime-as-a-virus metaphor emphasized more systemic aspects of the problem -- for instance, the fact that crime was spreading. Word-frequency counts in the two conditions help to illustrate this difference (see Table 5). Many of the words people used to describe

the crime problem across the two conditions were the same (e.g., “crime”, “city”, “control”), but there were some interesting differences across the two frequency-rank lists. For instance, words like “rampant” and “destroy,” that emphasize the overwhelming nature of the crime problem, appear in interpretations of the beast frame. In contrast, the word “spread” selectively shows up in interpretations of the virus frame.

frequency rank	beast	virus
1	city	crime
2	crime	city
3	speaker	**spread**
4	problem	virus
5	trying	rate
6	high	speaker
7	beast	people
8	**rampant**	problem
9	**destroy**	trying
10	control	control

Table 5: Word frequency counts by frame condition in Experiment 9. People interpreting the crime-as-a-beast frame tend to emphasize the urgency and scale of the crime problem whereas people interpreting the crime-as-a-virus frame tend to emphasize more systemic aspects of the problem (e.g., that it is spreading).

Individual Differences. In the implicit frame condition, there were significant differences by political affiliation and authoritarian tendencies. Participants who were Republican were more likely to endorse enforcement-oriented suggestions (62%) than non-Republicans (18%), $\chi^2[1,162] = 10.91$, $p < .001$ and $F[1,162] = 18.12$, $p < .001$. Similarly, participants high in the f-scale tended to be more enforcement-oriented than participants low in the f-scale, $\chi^2[1,412] = 17.79$, $p < .001$, and $F[1,412] = 35.28$, $p < .001$. There were no effects of age, gender, or educational history.

Not surprisingly, these individual difference measures did not have the same effect in the explicit condition, likely because participants were taking the speaker's perspective in ranking the crime-fighting suggestions. Participants with more education were marginally more likely to suggest reform-based solutions to the crime problem, χ^2 [1,168] = 3.74, p = .05 and F[1,168] = 3.34, p = .07. Otherwise, none of the individual difference measures predicted participants responses (political affiliation: χ^2[1,168] = .16, p = .69 and F[1,168] = .17, p = .68; f-score: χ^2[1,168] = .01, p = .99 and F[1,168] = .92, p = .34; Age: χ^2[1,168] = .08, p = .78 and F[1,168] = .12, p = .73; Gender: χ^2[1,168] = .09, p = .76 and F[1,168] < .01, p = .98).

There were significant interactions between condition (implicit and explicit) and political affiliation and between condition and authoritarian tendencies. People who were politically conservative or who scored high on the f-scale were more likely to endorse enforcement-oriented solutions in the implicit condition but not in the explicit condition (Political affiliation: χ^2[1,330] = 9.39, p < .01 and F[1,330] = 2.68, p < .01; f-scale: χ^2 [1,330] = 10.11, p < .01 and F[2,330] = 16.10, p < .001).

Discussion

Two aspects of the explicit framing condition in Experiment 9 distinguish it from the standard implicit metaphor framing experiment. First, in the explicit case participants' attention was directed at the frame when they were asked specifically to interpret the metaphor, whereas in the implicit condition, participants' attention was not directed explicitly towards the frame. Second, participants were asked make their crime-fighting suggestions from the metaphor-user's perspective not their own in the explicit condition but not in the implicit condition. In a third way, the design of the explicit condition was consistent with the implicit condition, but unlike that of the previous experiment: participants were only exposed to one metaphor frame.

With these changes in place, I found two important differences between the results of the implicit and explicit conditions. First, and most importantly, the metaphor

did not systematically affect participants' crime-fighting suggestions in the explicit condition. People who read that crime was a beast were equally likely to emphasize enforcement as people who read that crime was a virus. Second, the results of the implicit condition differ from those of the explicit condition in peoples' baseline tendency to emphasize enforcement. In the implicit condition, people were much more likely to emphasize reform whereas in the explicit condition peopler were much more likely to emphasize enforcement.

Together, these results suggest that the effectiveness of the metaphor, at least in the case where people encounter a single metaphor frame, hinges on it being processed and used implicitly. It may be the case that when participants process the metaphor explicitly they do not integrate the subsequent "information" in the report in a way that is consistent with the frame. Instead, they may find some feature of the frame to focus on -- perhaps an affective (Mial, 1987) or imagistic (Gibbs & O'Brien, 1990) component of the metaphor.

This may be because the virus and beast metaphors for crime are highly conventional (Bowdle & Gentner, 2005). According Bowdle and Gentner's career of metaphor hypothesis, novel metaphors are understood through a comparison process (Gentner, 1983) whereas conventional metaphors are understood through a categorization process (Glucksberg & Keysar, 1990). Inviting people to use a comparison process to comprehend a conventional metaphor may change what aspects of the metaphor are salient or meaningful for the reader.

Alternatively, it may be the case that the results of this experiment differ from the implicit metaphor condition because people are asked to take the speaker's perspective in ranking the crime-fighting responses. It may be the case that people try to read into the language to get a sense of the speaker's personality in some way that distracts them from processing the report in a way that would be consistent with Experiments 1-5.

One outstanding question about these results relates to the form of the representations that subserve the metaphor framing effect. Is it the case that peoples' schematic representations of the virus and beast problems are transferred into the domain

of crime to structure how people think about crime and criminals and social institutions? Or, might people use a more abstract schema of spreading versus attacking?

In Experiments 10 and 11, I investigate whether the metaphoric frame highlights specific relational components of the crime domain. Does framing crime as a beast make direct aspects of the problem -- the overwhelming nature and urgency of it -- more salient? Does framing crime as a virus make more systemic aspects of the problem -- that it has the capacity to spread -- more salient?

In Experiment 10 I test this possibility by substituting particular relations for the metaphoric frame. Using the results of Experiment 9, in which people wrote out specific interpretations of the frames, I substituted the word "overwhelming" for "beast ravaging" and "spreading in" for "virus ravaging." If the metaphor in Experiments 1-5 works by weighting specific relational structure of the crime domain, then the results of the relational frame should mirror the results of the metaphoric framing condition.

Experiment 10

Method

Participants. Four-hundred fifty-three Turkers participated for pay, of whom 251 were in a replication "metaphor frame" condition and 202 were in a "relation frame" condition. Two-hundred forty-two participants were female and 211 were male. The ages of the participants ranged from 18-75 with a mean of 32 (median = 25). The educational background of the participants spanned the full range of options from "some high school" (2%) to "doctorate" (4%) with most participants having at least "some college" (385, 85%).

Materials and Design. In Experiment 10, participants read a crime report that was framed with a non-metaphoric analog of "virus" or "beast": namely, "spreading in" or "overwhelming." These words were commonly used by participants in Experiment 8 to

describe what a speaker might be seeking to convey through the use of the virus or beast metaphors for crime. The report read as follows:

> Crime is {spreading in/overwhelming} the city of Addison. Five years ago Addison was in good shape, with no obvious vulnerabilities. Unfortunately, in the past five years the city's defense systems have weakened, and the city has succumbed to crime. Today, there are more than 55,000 criminal incidents a year - up by more than 10,000 per year. There is a worry that if the city does not regain its strength soon, even more serious problems may start to develop.

After reading the report, participants were presented with the standard set of four response options, which they were asked to rank order by dragging and dropping them into a response window. Once they had submitted their response, they were asked to identify which aspect of the report was most influential in their decision, complete three personality questionnaires, and answer a set of background questions.

Results

Collapsing across the two conditions, there was an overall bias in favor of the reform-oriented crime-fighting options (82%; $\chi^2[1,N = 453] = 186.93$, $p < .001$) and a systematic influence of the frame. This was true of participants' first choice (D[N=453] = .11, $p < .01$) and weighted set of responses (D[N=453] = 25.24, $t[451] = 2.07$, $p < .05$). There were no differences between conditions on any of these three measures (baseline preference for reform: $\chi^2[1,451] = .047$, $p = .82$; first choice: $\chi^2[2,449] = .125$, $p = .94$; weighted response set: $F[2,449] = .004$, $p = .996$) (see Figure 15).

Fifteen participants (3%) identified the frame as influential in their response. Marginally more participants in the relation condition identified the frame (10, 5%) than in the metaphor condition (5, 2%), $\chi^2[1,451] = 3.06$, $p = .08$. Further, numerically more participants who read the "virus" or "spreading" frame identified the metaphor as important (4%) than participants who read the "beast" or "overwhelming" frame (2%), although this difference was not significant, $\chi^2[2,451] = 2.42$, $p = .12$.

Individual Differences. As in previous experiments, Republicans were more likely to endorse enforcement-oriented responses (31%) than Independents (19%) and Democrats (10%), $\chi^2[1,450] = 16.17$, $p < .001$. There was no difference in the influence of the frame across political affiliations, $\chi^2[1,449] = 1.25$, $p = .26$.

People who scored high on the f-scale were more likely to choose an enforcement-oriented suggestion as their first choice, $\chi^2[1,450] = 20$, $p < .001$, and to emphasize enforcement in their weighted response set, $F[1,450] = 31.51$, $p < .001$. There was no interaction between dogmatism and the frame in predicting participants' first-choice, $\chi^2[1,449] = .0007$, $p = .98$ or weighted response set, $F[1,449] = 1.25$, $p = .26$.

No other individual difference measures significantly predicted participants' first choice or weighted response set or interacted with the frame to predict these measures (Age: $\chi^2[1,450] = 1.96$, $p = .16$ and $\chi^2[1,449] = 2.08$, $p = .15$; $F[1,450] = 10.40$, $p < .01$ and $F[1,449] = .36$, $p = .5$; Education: $\chi^2[1,449] = 5.14$, $p < .05$ and $\chi^2[1,448] = 2.23$, $p = .14$; $F[1,450] = 2.33$, $p = .13$ and $F[1,449] = 4.21$, $p < .05$; Gender: $\chi^2[1,450] = .87$, $p = .35$ and $\chi^2[1,449] = .43$, $p = .51$; $F[1,450] = 1.00$, $p = .33$ and $F[1,449] = 1.72$, $p = .19$).

Best Model. The best model included regressors for frame, political affiliation, f-score, and age (in predicting participants' weighted response set). People who read that crime was a virus were less likely to emphasize enforcement. People who were politically conservative, had authoritarian tendencies, or were older were more likely to emphasize enforcement (see Table 6).

regressor	logistic			linear		
	β	*se*	*p*	*β*	*se*	*p*
intercept	-2.99	0.67	< .001	-204.4	32.69	< .001
frame(virus)	-0.66	0.26	< .05	-18.04	10.35	= .06
political affiliation	0.674	0.18	< .01	39.18	8.19	< .001
f-score	0.059	0.02	< .001	3.95	0.83	< .001
age				1.67	0.55	< .01

Table 6: Results of Experiment 10. Best models for predicting participants' first choice method for responding to crime (logistic) and weighted response set (linear). Both models include regressors for frame, political affiliation, and f-score. The model predicting participants' weighted response set additionally includes a predictor for age.

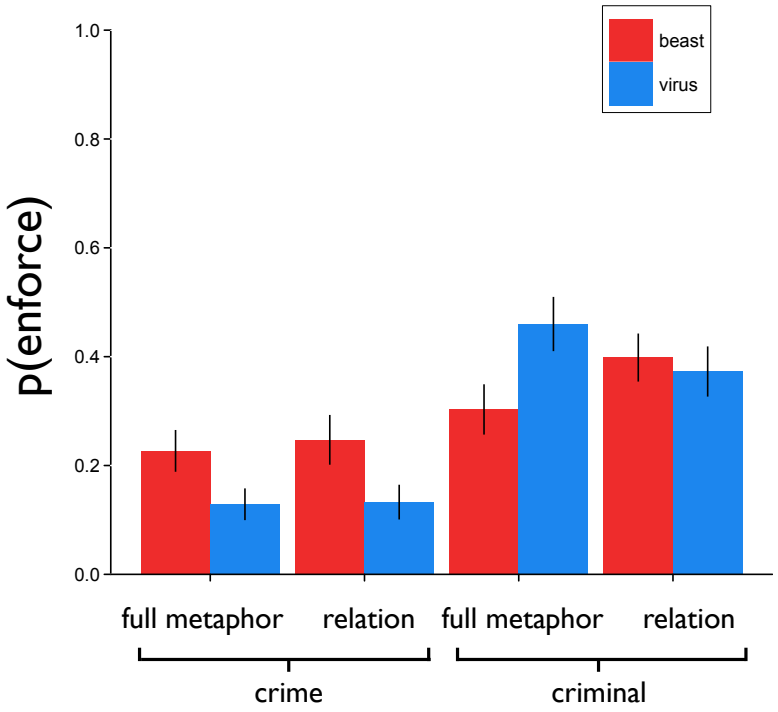

Figure 15: Results of Experiment 10 (crime) and 11 (criminal). Both the metaphorical and the relational frames influenced peoples' crime-reducing suggestions; only the metaphorical frames influenced peoples' sentencing judgment.

Discussion

In Experiment 10 I replaced the noun frames "beast" and "virus" with specific relations "overwhelming" and "spreading" to investigate whether the metaphor's influence relies on the highlighting of specific relational structure in the target domain. The results of the relational condition closely match the results of the noun condition, suggesting that the noun metaphor frames in Experiments 2-5 may work by eliciting particular relations in the target domain.

Alternatively, it is possible that both the noun frames and the relational frames are two routes to instantiating a more abstract schematic structure. For instance, the "virus" and "spreading" frames may both instantiate a more general image-schema of containment or the idea of a systemic problem. In contrast, the "beast" and "overwhelming" frames may both instantiate a more general image-schema of attack or the idea of a direct problem.

One final experimental question I will explore in this dissertation is whether the virus and beast frames only instantiate a single relation in the context of crime. That is, one possibility is that the metaphor "Crime is a virus" always instantiates some notion of "spreading" and that the metaphor "Crime is a beast" always instantiates some notion of "overwhelming." However, an alternative possibility is that the noun frames have the potential to highlight a variety of relational structure. In the context of solving crime, ideas of "spreading" and "overwhelming" are most informative. However, in another context, for instance when making a sentencing decision for a criminal, it might be the case that a different relation or set of relations is called to mind.

I test this possibility in Experiment 11 by asking people to read the same crime report as in Experiment 10 (noun: "virus" or "beast"; or relational: "spreading" or "overwhelming") before making a sentencing decision. Specifically, participants are asked whether "increasing prison sentences" or "mandating job training programs while incarcerated" would be a more effective sentence for convicted criminals.

In this context, it may be the case that the noun metaphors offer qualitatively different solutions from what they suggest in the crime-fighting context. Intuitively, a salient way of dealing with a literal virus that has been captured would be to quarantine it, while a salient way of dealing with a literal beast that has been captured would be to train or domesticate it. If the virus and beast metaphors appeal to these alternative schemas in a sentencing context, than the results of Experiment 11 should be different from the results of Experiment 10 in two ways. First, the results of the relational condition should not mirror the results of the noun condition, since the relations "spreading" and "overwhelming" are not as relevant in the sentencing context. And, second, participants who read the "beast" frame should be less likely to endorse harsher, more enforcement-oriented responses than participants who read the "virus" frame.

Experiment 11

Method

Participants. Four-hundred thirty-two Turkers participated in Experiment 11 for pay. One-hundred ninety-nine were in the metaphorical frame condition and 233 were in the relation frame condition. Of these participants, 230 were female and 202 were male; their ages ranged from 18-75 with a mean of 33 (median = 25); the educational history of the participants spanned the range from "some high school" (6, 1%) to "doctorate" (16, 4%) with most having at least "some college" (368, 85%); 133 identified as Democrats, 222 as Independents and 77 as Republicans.

Materials and Design. In the noun and relational conditions of the experiment, participants read a framed version of the crime report. The frame was either noun (virus vs. beast) or relational (spreading vs. overwhelming). After reading the report, they were told that the city had been able to arrest a number of criminals in recent weeks and that

the city officials were debating how to handle the criminals. These instructions read as follows:

> In recent weeks Addison has stepped up its policing efforts and has arrested numerous criminal offenders. As a result, the city's policy makers and judges have gathered to discuss how these criminals should be handled. The judges have found that they are frequently hearing similar cases and hope to come to some consensus in their sentencing judgments.
>
> Imagine that they have asked you to participate in this discussion. Addison's city officials would like to know how you think they should focus their limited attention and resources.
>
> One of the most frequent cases that the judges have been hearing involves someone accused of armed robbery. If convicted of this offense, the current sentencing guideline suggests a prison term of 2 to 12 years, with the most frequent sentence lasting about 7 years. The judges have decided that this sentence is not doing enough to reduce crime though, and they have proposed adopting one of the following two modifications.

Participants were presented with the two options listed below. The first was categorized as more direct/enforcement-oriented and the second was categorized as more systemic/reform-oriented.

1) Increase standard prison terms for criminals convicted of armed robbery to the maximum of 12 years.
2) Mandate that criminals convicted of armed robbery complete job training programs while incarcerated.

Participants were then asked to identify the part of the report that was most influential in their crime-reducing suggestion.

In all three versions of the experiment, participants were asked to complete three personality questionnaires and a set of background questions.

Results

Overall, participants were more likely to suggest mandating job training, the reform option, (62%) than to suggest extending prison sentences, $\chi^2[1,N=432] = 23.15$, $p < .001$. This was true in the metaphor condition (62% reform), $\chi^2[1,N=199] = 11.10$, $p < .001$, as well as the relation condition, (61% reform); $\chi^2[1,N=233] = 12.06$, $p < .001$.

Numerically, participants were more likely to suggest enforcement-oriented solutions after reading the virus/spreading frame (41%) than after reading the beast/ overwhelming frame (36%), but this difference was not statistically significant, $\chi^2[1,N=432] = 1.32$, $p = .25$.

Breaking the data down by condition, however, reveals a statistically significant effect of the frame in the metaphor condition: participants were harsher given the virus frame (46% enforcement) than the beast frame (30% enforcement), $\chi^2[1,N=199] = 4.55$, $p < .05$. There was no effect of the frame in the relation condition: participants were equally likely to suggest increasing prison sentences given the virus frame (37%) as the beast frame (40%), $\chi^2[1,N= 233] = .071$, $p = .79$. There was a marginally significant interaction between the frame and the framing condition (noun vs. relational), $\chi^2[1,428] = 3.82$, $p = .05$.

Noticing the Metaphor. Overall, 23 (5%) participants identified the frame as important in making their decision. Fourteen of these participants were in the noun frame condition (constituting 7% of the participants in this condition) while 9 were in the relation frame condition (constituting 4% of participants in this condition). In the noun framing condition, 9 of the 14 (64%) participants who identified the metaphor as important had read the virus frame; in the relation framing condition 3 of the 9 (33%) participants who identified the metaphor as important had read the “spreading” frame.

Individual Differences. As expected, Republicans were more likely to suggest increasing prison sentences (49%) than Independents (43%) or Democrats (25%). This

pattern could be seen in both the noun condition (50%, 43%, and 23%, respectively) and the relational condition (49%, 43%, 26%, respectively), $\chi^2[1,427]=14.51$, $p < .01$. There was no interaction between political affiliation and framing condition, $\chi^2[1,426] < .001$, $p = .98$.

Participants with more education were marginally more likely to suggest mandating a job training program for prisoners, $\chi^2[1,427] = 3.33$, $p = .068$. This effect did not differ across framing conditions, $\chi^2[1,426] = 2.53$, $p = .11$.

As in the crime-fighting judgment, participants who scored high on the f-scale were more likely to suggest that prison sentences should be extended, $\chi^2[[1,427] = 28.58$, $p < .001$. This measure did not interact with the frame ($\chi^2[1,426] = .09$, $p = .76$) or the framing condition ($\chi^2[1,426] = .18$, $p = .67$).

No other individual difference measures were predictive (age: $\chi^2[1,427]=.311$; $\chi^2[2,426]=.312$; gender: $\chi^2[1,427]=1.17$, $p = .84$; $\chi^2[2,426]=1.86$, p .17).

Best Model. The best fitting model included regressors for condition (metaphor vs. relation), frame (virus/spreading vs. beast/overwhelming), political affiliation, f-score, and an interaction between frame and condition (see Table 7). Participants who were politically conservative, had authoritarian tendencies, or received the virus framing were more likely to suggest extending prison sentences.

Predictor	Logistic		
	β	***se***	***p***
intercept	-2.96	0.59	< .001
condition(relation)	0.345	0.30	= .25
frame(virus/spread)	0.538	0.31	= .08
political affiliation	0.359	0.15	< .05
f-score	0.073	0.02	< .001
frame(virus/spread)*condition(relation)	-0.656	0.42	= .12

Table 7: Results of Experiment 11. Best model for predicting participants' first choice method for sentencing a criminal.

Discussion

In Experiment 11, participants were asked to make a sentencing decision rather than a crime-fighting decision. The context of the decision was changed to investigate whether the "virus" and "beast" noun metaphors offer a single relation each for thinking about crime (e.g., "spreading" for "virus" and "overwhelming" for "beast").

As predicted, the results of Experiment 11 differed from those of Experiment 10 in two substantial ways. First, the results of the relational condition did not mirror the results of the noun condition. Second, the direction of the effect in the noun condition was different in the sentencing context. People who read that crime was a virus were more likely to suggest extending prison sentences for criminals than people who read that crime was a beast.

People who read the virus metaphor may have been more enforcement oriented than people who read the beast metaphor on this measure because this context tapped into different virus and beast schemas: once caught, viruses are to be quarantined and beasts can be trained.

These results also suggest that the findings from previous experiments was not driven by a difference in, for instance, the affective components of the metaphors (Mial, 1987). If, for instance, beasts made people fearful and being fearful made people more enforcement-oriented, then we would expect people who read that crime was a beast to be more enforcement-oriented when making a sentencing decision.

More generally, these results suggest that metaphors can be productive in a number of different ways. Hearing that crime is a "virus" or "beast" does not seem to elicit only one specific relation. Instead, the metaphor seems to offer a rich pallet of relations, and the specific contribution of the metaphor may be determined by contextual factors.

Summary of Experiments 6-11

In Chapter 3, six experiments explored the underlying mechanisms that support metaphoric reasoning. In Experiments 6 and 7 I found that the the framing effect depended on the word "virus" or "beast" being presented as a metaphor early in peoples' stream of processing. The synonym listing task in Experiment 6 did not yield results that mirrored the metaphor framing condition; similarly, reading the frame at the end of the report, as in Experiment 7, had no effect on peoples' reasoning. The virus and beast metaphors may influence peoples' crime-fighting suggestions by shading how they interpreted specific words in the crime report and, therefore, how they built up representations of the crime problem: words like "defense systems" and "vulnerable" may take on one interpretation after reading that crime is a virus and a different meaning after reading that crime is a beast, leading people to conceptualize the crime problem differently on the two framings.

In Experiments 8 and 9 I found that the effect of the metaphor was not the result of deliberate reasoning. In Experiment 8 I found that people were capable of using the metaphors to deliberately reason about crime: when people were presented with both frames and asked to reason from the speaker's perspective, the metaphor framing effect increased dramatically. However, this boost in the power of the metaphor did not translate into Experiment 9, in which people were given a single metaphor and asked to explicitly use it to make their crime-fighting suggestion. These particular metaphors for crime, and conventional metaphors in general, are most powerful when they are processed implicitly.

Finally, in Experiments 10 and 11, I explored whether the noun metaphoric frames influence peoples' reasoning by eliciting particular relational structure in the target domain. In Experiment 10, I found that replacing the noun frames, "virus" and "beast," with specific relations, "spreading" and "overwhelming," had a similar effect as the corresponding noun frames on peoples' crime-fighting suggestions. People who read that crime was "overwhelming" were more likely than people who read that crime was "spreading" to endorse enforcement-oriented responses to crime.

In Experiment 11, participants were asked to make a sentencing judgment rather than a crime-fighting suggestion. In that context, the virus and beast metaphors had a different kind of influence on peoples' judgment: people who read that crime was a virus were more enforcement-oriented than people who read that crime was a beast. And, in the sentencing context, the relational frames were ineffective. The results of these experiments suggest that the meaning of a given metaphor is not fossilized. Metaphors can instantiate different meanings in different contexts.

General Discussion

Metaphoric Reasoning and Mental Representation

Together, these experiments suggest that metaphors dynamically affect reasoning. The meaning that a metaphor takes on is shaped by its surrounding context. This is why priming people with "virus" and "beast" as lexical items rather than metaphors had no effect on their reasoning about crime and why changing the judgment people are asked to make changes how metaphor influences their judgment. However, a metaphor's surrounding context -- particularly, information that is processed after the metaphor -- is also shaped by the meaning that a metaphor takes on. As a result, metaphors that are processed early in a stream of language are more influential than metaphors that are processed late. In the case of the virus and beast metaphors for crime, this process appears to unfold in the absence of any deliberate or strategic control of the language processor. Indeed, people were largely unaware that they had been influenced by a metaphoric frame.

One further question relates to the nature of the mental representations that subserve this effect of metaphoric reasoning? To what extent is the effect of metaphor on reasoning driven by peoples' long-term, stable representations of crime and to what extent is the effect of metaphor on reasoning driven by peoples' short-term constructed representations of crime?

One possibility is that the effect is driven by peoples' long-term representations of crime. Peoples' conceptual representation of crime is contained metaphorically, in conceptual representations of other, more experiential domains (Lakoff & Johnson, 1980, 1999). On this view, the metaphor frame serves to activate the particular component of the crime representation that is housed in our representation for a "virus" or "beast" problem.

An alternative possibility is that the effect is driven by peoples' short-term constructed representation of crime. People have disjoint representations of each of these domains (crime, virus, beast) and that, upon encountering the metaphor, people map the source domain onto the target domain and project inferences from the source domain to the target domain (Gentner, 1983). In other words, one possibility is that metaphor is like analogy (Gentner, Bowdle, Wolff, & Boronat, 2001).

There is also at least one other intermediate possibility. On this third approach, conceptual representations are overlapping and distributed and simultaneously activating multiple conceptual representations serves to highlight shared relational structure (Rogers & McClelland, 2004, 2008).

Each of these accounts is capable of explaining the data from Chapters 2 and 3. However, theoretical arguments against the first possibility, for strong metaphoric representation, and a lack of psychological evidence in support of the view suggest that the first possibility may not be the best metaphor for describing the metaphor reasoning effect (Murphy, 1996, 1997). First, most of the evidence that is taken as supporting this view is linguistic, and therefore circular. Extended metaphors in language suggests the psychological reality of conceptual metaphors and extended metaphors in language are evidence of conceptual metaphors. Without converging psychological evidence, the presence of extended conceptual metaphors in language does not strongly support the notion of conceptual metaphor. Second, the presence of multiple metaphor families for a given target domain (e.g., ECONOMY-as-PLANT and -PERSON and -MACHINE and -VEHICLE) appears to contradict the spirit of conceptual metaphor theory -- in which conceptual representations for more experiential domains are said to contain, and

therefore simplify, representations for more abstract domains. For these reasons (and others: see Murphy 1996, 1997), the strong view of metaphoric representation is generally seen to be untenable.

This leaves two other possibilities: that metaphoric reasoning is the result of an on-line analogical mapping mechanism and that metaphoric reasoning is the result of a more dynamic constraint satisfaction process. Given the neural evidence for overlapping, distributed representation (Dilkina, McClelland, & Plaut, 2008; Rogers & McClelland, 2004; Rogers et al., 2004) as well as behavioral evidence, suggesting that conventional metaphors are processed as quickly and easily as non-metaphoric language (Gibbs, 1980; Harris, 1976; Ortony, Schallert, Reynolds, & Antos, 1978; Verbrugge & McCarrell, 1977), the latter possibility may be more likely. However, some have argued that such a system is, in principle, unable to account for effects of analogical and metaphorical reasoning. In Chapter 4, I will present a computational model that seeks to demonstrate how such a process could unfold -- namely, how overlapping distributed representations, and no on-line mapping mechanism -- can support metaphorical and analogical reasoning.

Limitations

Before moving on to presenting a computational model of metaphoric reasoning, I wish to acknowledge some important limitations of the empirical results I described in Chapters 2 and 3. First, all of the studies that I present contrast two particular metaphors in primarily one domain. This approach is advantageous in that it allows for some comparison across experiments, and facilitates the careful and systematic study of the factors that influence metaphoric reasoning. However, the drawback of this approach is that the results may or may not generalize beyond these particular metaphors for crime.

Indeed, over the course of my graduate career, I have found conflicting results in other domains and with other metaphor frames. For instance, I have used the “virus” and “beast” frames in contexts other than crime and drugs and have not seen the same

systematic influence of the frames on peoples' thinking. This was true of the framing of a homeless problem, a cheating problem, and a terrorism problem (see Appendix D for these stories). I predicted that the virus frame would make people approach each of these problems with an emphasis on social reform and that framing them as a beast would make people approach them with an emphasis on enforcement. This was not the case.

Further, I have explored other metaphoric frames in other domains that has yielded results both in support of and in opposition to the notion that conventional metaphors influence how people reason about complex domains. In some cases I found predicted results -- for instance, I found that when the educational system was framed as a garden rather than a factory, people were less likely to think that standardized tests were an effective way to measure student achievement. However, in other cases, I found no effect of metaphor on how people reasoned about the economy (see Appendix E for these stories). Future work will aim to identify some of the conditions of metaphors and their surrounding context that might predict which metaphors in which contexts influence reasoning.

Chapter 4:
A Connectionist Approach to Metaphorical Reasoning

In the previous two chapters, I presented several experiments showing that metaphors can shape the way we think. When crime was framed as a beast, rather than a virus, people were more likely to suggest solving crime by increasing enforcement measures than through social reform. Further, I identified several psychological factors that contribute to the effectiveness of a metaphoric frame: notably, that (a) the source domain (virus or beast) was instantiated as a metaphor (b) the metaphor was presented early in a stream of processing and (c) the metaphor was processed and used implicitly.

In light of these findings, I argued that metaphors dynamically influence reasoning. In the experiments presented above, the virus and beast metaphors constrained the constructed representation of the target domain, crime in Addison -- in part, by shading the interpretation of the language that followed the metaphor. However, the meaning that the virus and beast metaphors took on (and, hence the way that the frame influenced how people reasoned about crime), was also shaped by their surrounding context. A given frame influenced peoples' judgment about crime in one way (e.g., the beast metaphor made people more enforcement-oriented), but it influenced peoples' judgment about criminals in another way (e.g., the beast metaphor made people more reform-oriented).

In this chapter, I present a computational perspective on the processes that underlie this type of metaphoric reasoning. I will use the Rumelhart model (Rumelhart, 1990; Rumelhart & Todd, 1993), a connectionist model that has been used to simulate a variety of phenomena related to conceptual development and semantic cognition (Rogers & McClelland, 2004), to argue that this type of metaphoric reasoning is a natural byproduct of our conceptual system. That is, I will argue that fundamental, biologically grounded principles of learning (i.e., progressive differentiation) combined with a sub-symbolic scheme for conceptual representation (i.e., one in which knowledge lives in shared weights rather than disjoint propositions) can naturally give rise to this type of

metaphoric reasoning. Indeed, on this view, metaphoric reasoning is a core feature of the conceptual system: a form of generalization over homologous relational structure.

In making this argument, I will contrast my computational approach with other, state-of-the-art models of analogical reasoning -- namely, the Structure Mapping Engine (SME: Falkenhainer, Forbus, & Gentner, 1989) and Learning and Inference with Schemas and Analogies (LISA: Hummel & Holyoak, 1997, 2003, 2005). I will make this comparison for two reasons: first, the processes that underlie metaphorical reasoning are thought to be the same as those that to underlie analogical reasoning (e.g., Gentner, Bowdle, Wolff, & Boronat, 2001). Second, proponents of these modeling approaches offer a very different theory of how metaphorical or analogical reasoning comes about. Indeed, some go so far as to say that a sub-symbolic system, like the Rumelhart model, is, in principle, not capable of supporting any type of analogical or metaphorical reasoning. On this view, symbolic representations and a specific mapping mechanism are *necessary* for any system to carry out analogical or metaphorical reasoning (Doumas, Hummel, & Sandhofer, 2008; Hummel, 2010; Gentner & Markman, 1993). Hummel (2010), for instance, argues that "symbol systems permit qualitatively different kinds of processing (such as learning and inference) than do non-symbolic systems (a difference so profound that our symbolic species dominates the planet, whereas our nonsymbolic cousins do not)" (p. 961). For these reasons, I will refer to this approach as the *Structured* approach to analogical or metaphorical reasoning and my alternative as the *Emergent* approach.

In what follows, I will first describe the *Structured* view of analogical and metaphorical reasoning, noting both advantages and disadvantages of this perspective. Second, I will describe an *Emergent* alternative that attempts to build upon the strengths of the *Structured* models and improve upon their weaknesses. Third, I will describe a series of theoretical simulations designed to show that a sub-symbolic system without a specific mapping mechanism can, in principle, capture important features of analogical or metaphorical reasoning. Fourth, I will describe two additional simulations that model an empirical demonstration of metaphoric reasoning. Finally, I will discuss some of the

implications of this approach and how they inform our thinking about the process of metaphorical reasoning.

The Structured Approach to Analogical and Metaphorical Reasoning

Structured models of analogical and metaphorical reasoning endorse the notion that these processes can be decomposed into several sub-components. In a recent review, Gentner and Forbus (2011) describe four major sub-processes of analogy:[11]

1. *Retrieval*: Given a situation, find an analog that is similar to it.
2. *Mapping*: Given two situations, align them structurally to produce a set of correspondences that indicate 'what goes with what,' candidate inferences that follow from the analogy, and a structural evaluation score which provides a numerical measure of how well the base and target align.
3. *Abstraction*: The results of comparison may be stored as an abstraction, producing a schema or other rule-like structure.
4. *Rerepresentation*: Given a partial match, people may alter one or both analogs to improve the match. (Gentner and Forbus, 2011, p. 267).

In SME and LISA these component processes are implemented as distinct systems that come together to give rise to analogical reasoning. Of particular importance, these models of analogy share a basic commitment to treating questions of conceptual representation and mapping as separate issues. In both SME and LISA this is achieved by hand coding structured symbolic (or hybrid) conceptual representations (e.g. of objects and relations) and implementing a distinct mechanism dedicated to mapping (or binding)

[11] It should be noted that this is not the *only* way to sub-divide analogy. For example, Hall (1989) suggests that analogy consists of Recognition, Elaboration, Evaluation & Transfer, and Consolidation. However, there is a great deal of overlap between the ways in which different researchers conceive of these sub-processes (e.g. Elaboration is very similar to Mapping).

that operates over these representations (Gentner & Forbus, 2011; but see Leech, Mareschal, & Cooper, 2008).

The success that these models have had at simulating a wide range of behavioral findings suggests that they may capture certain important features of human analogical processing (Hummel & Holyoak, 2005; Gentner & Forbus, 2011). Indeed, this success has led some researchers to argue that the *algorithmic* processes which allow these models to exhibit analogical behavior, necessitate symbolic representations and a specific mapping mechanism (Doumas, Hummel, & Sandhofer, 2008; Hummel, 2010; Gentner & Markman, 1993). Indeed, some proponents of the *Structured* approach argue that the Rumelhart model, specifically, is incapable of performing any kind of analogical or metaphorical processing because, as Opfer and Doumas (2008) articulate, the model "represents neither relations (e.g., goal-directed) nor relation-filler bindings explicitly. Consequently, [the Rumelhart] model cannot use relational information to drive inference." They go on to explain that "successful models of semantic cognition must be able to learn explicit representations of properties and relations from examples and must bind these representations to novel arguments" (p. 723).

This view constitutes an *a priori* constraint on any theory or model of analogy or metaphor, and implies that there are classes of models (and cognitive systems) that cannot *in principle* capture analogical reasoning (namely, any model that does not obviously include structured representations and a mapping mechanism). In particular, models that rely on low-level associative learning mechanisms and fully distributed representations to simulate cognitive processing, like many connectionist models (e.g., Rogers & McClelland, 2008), are thought to be incapable of supporting fully analogical capabilities (Holyoak and Hummel, 2008; Hummel, 2011; Kemp & Tenenbaum, 2008; Marcus, 2001; Marcus & Keil, 2008; Opfer & Doumas, 2008). A common view states that while non-symbolic models "excel at learning complex correlations between features, they fail to represent abstract operations over variables, structured representations, and contrasts between individuals and kinds; and it is not clear how well they can do any of these things in principle" (Marcus & Keil, 2008, p. 722). Some

theorists have gone so far as to suggest that the attempt to capture higher-level cognitive abilities such as analogy in these sorts of models involves "suck[ing] the essence out, then force-fit[ting] what's left into an associationist straitjacket" (Holyoak and Hummel, 2008, p. 389).

To summarize, then, the approach taken by SME and LISA embodies two related hypotheses: (1) that the sub-components of analogy outlined above, especially a specific mapping mechanism that operates over structured representations, are necessary for analogical processing, and (2) that models of a certain class (i.e. in which these components are not explicitly built in, or in which distributed representations and low-level learning mechanisms carry the burden of processing) cannot in principle come to instantiate these processes, and therefore cannot implement analogical processing.

The Emergent Approach to Analogical Reasoning

In contrast to the structured approach outlined above, I will argue that questions of mapping are inseparable from questions of conceptual representation. On this view, the concept of *mapping* may be thought of as a computational level description (Marr, 1982) of a fundamentally integrated phenomenon, but without direct implications for the algorithmic or implementation levels. That is, the way that concepts are stored and represented in the system (i.e. that they are learned and distributed) may naturally give rise to analogical mapping over the course of development. Previously, proponents of this emergent approach have had success simulating certain aspects of the development of analogical reasoning (see especially Leech et al., 2008). However, a major criticism of this work has been its failure to capture our ability to use analogy to drive inferences and facilitate knowledge acquisition (e.g. Markman & Laux, 2008; Opfer & Doumas, 2008).

While this approach builds on previous work by Hinton (1986) and Rogers and McClelland (2008), it represents a novel theory of metaphorical reasoning and suggests that some forms of metaphorical reasoning may be better thought of as generalization over relational structure. This connectionist approach may help to explain why, for

instance, people do not take longer to process metaphors in natural discourse and why, in the experiments I presented above, people do not seem to notice the influence of the metaphor on their decisions.

In the section that follows, I will use a toy domain -- of family structures -- to demonstrate that the Rumelhart model can exhibit two key features of analogical reasoning. First, I demonstrate that the model can make an inference based on shared relational structure between a source and target domain in the face of conflicting feature-based similarity, a hallmark of mature analogical reasoning (Gentner, 1983). Second, I show that the model can leverage shared relational structure between two domains in order to facilitate learning (Carey, 2009; Gentner, 2010a). Later, I will use the same model to simulate empirical findings on a task that involves metaphorical reasoning.

Who Walks the Dog?: A Connectionist Exploration of Analogical Reasoning

General Modeling Framework

The learning task used in the first set of simulations is inspired by Hinton's (1986) family tree model, one of the first attempts to address relational learning in a connectionist network. Previous empirical work has shown that family trees are closely related to analogy, to the point that practice with family trees facilitates analogical inference in young children (e.g. Mutafchieva & Kokinov, 2007). The goal of the model is to learn "statements" that are true about the various members of a family, including identity information, perceptual features, and relations between family members. The input to the model consists of activating a Subject unit, corresponding to a particular family member, and a Relation unit. The Relation units correspond to the different kinds of relationships that can hold between subjects and objects (e.g., "is_named", "is_a", "has", "parent_of", "daughter_of").

The network is wired up in a strictly feed-forward fashion, as shown in Figure 16, such that the input propagates forward through the internal layers, resulting in a set of

predictions over the Object layer. Over the course of training, the network's weights change (via backpropagation of the cross-entropy error on the output units) in order to better predict which objects hold for each particular Relation to each Subject. As the model also contains intervening layers of units between the input and output layers, it is forced to re-represent the inputs as a distributed pattern of activation over these internal layers[12].

These distributed patterns are built up slowly over the course of development by an error-driven process of progressive differentiation. It is forced to learn to re-represent inputs as distributed patterns of activation over many processing units in order to improve its predictions of output patterns. The modeling approach is consistent with an account of learning in which children (and adults) are constantly making predictions about what they will experience in the world, and using their observations (i.e. what they actually do experience) to make better predictions in the future.

[12] For convenience, I often use labeled or even localist input and output units; however, these are not meant to be interpreted as explicit symbols. Rather, I interpret these patterns as observed states of the world, including linguistic and nonlinguistic perceptual-motor experience, which are used both to predict future states and to provide corrective feedback to these predictions (Flusberg et al., 2010; Rogers & McClelland, 2008). What is important is the statistical structure encoded by the entire set of input and output vectors.

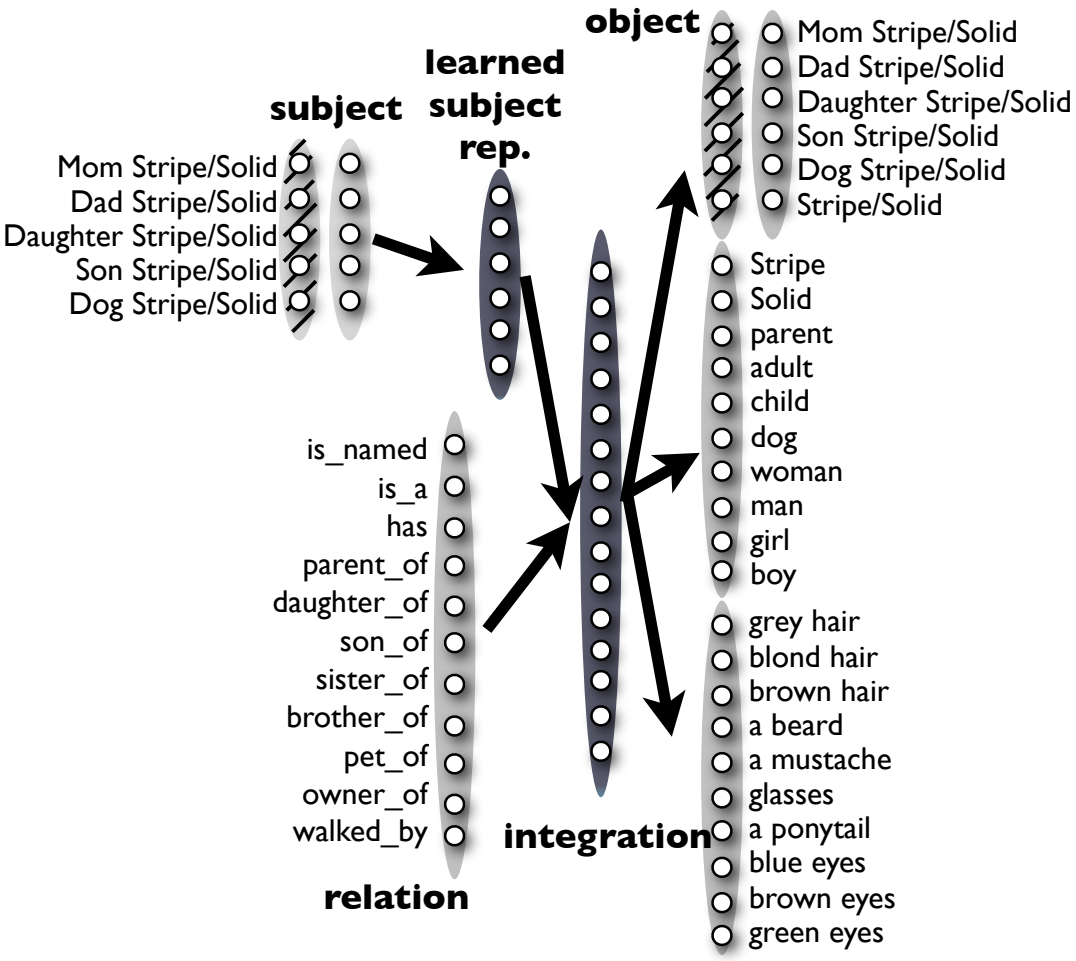

Figure 16. The network architecture. Note that in some cases, more or fewer units were used in the Subject, Relation, or Object layer to accommodate more or fewer families or family members.

In this simple world, the network learns about two families: the Stripes and the Solids (pictured in Figure 17 and described in Table 8). Each family has a mother, a father, a daughter, a son, and a pet dog. Each individual has an identity (e.g. “Daughter” and “Stripe”) given by the “is_named” Relation. In these training patterns there is no output overlap between the families. Specifically, there is one unit that represents membership in the Stripe family and one that represents membership in the Solid family. In addition, there is one output unit associated with each individual that represents their name.

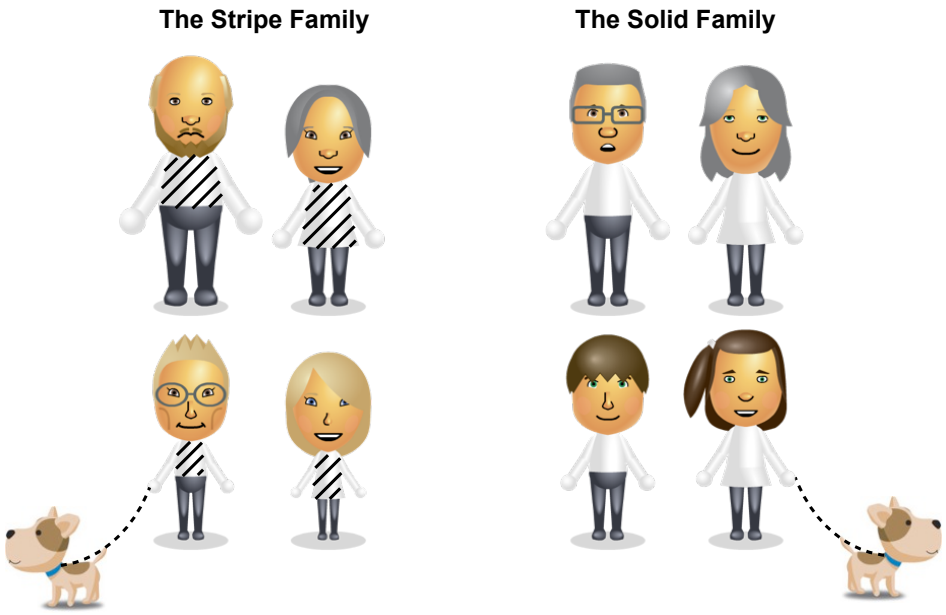

Figure 17. An illustration of the Stripe and Solid families, which served as the source and target domain for most simulations.

Each individual can be described as belonging to a variety of categories - e.g., human, dog, parent, child - given by the "is_a" relation. In this context, corresponding family members in the two families share numerous features. For instance, the mother in the Stripe family and the mother in the Solid family are both "human", "parent", "woman", etc.

Each individual has a set of perceptual features, such as grey hair, a mustache, or glasses, given by the "has" relation. In this context there is also feature overlap across families as well as feature overlap within families. For instance, the mother in the Stripe family and the mother in the Solid family both have grey hair, while the daughter in the Stripe family and the son in the Stripe family both have blonde hair.

Mom_{Stripe}	*is_named*	Stripe, Mom_{Stripe}
Dad_{Stripe}	*is_named*	Stripe, Dad_{Stripe}
$Daughter_{Stripe}$	*is_named*	Stripe, $Daughter_{Stripe}$
Son_{Stripe}	*is_named*	Stripe, Son_{Stripe}
Dog_{Stripe}	*is_named*	Stripe, Sog_{Stripe}
Mom_{Stripe}	*is_a*	Stripe, human, adult, parent, mom
Dad_{Stripe}	*is_a*	Stripe, human, adult, parent, dad
$Daughter_{Stripe}$	*is_a*	Stripe, human, child, daughter
Son_{Stripe}	*is_a*	Stripe, human, child, daughter
Dog_{Stripe}	*is_a*	Stripe, dog
Mom_{Stripe}	*has*	grey hair, ponytail, brown eyes
Dad_{Stripe}	*has*	a bald head, beard, brown eyes
$Daughter_{Stripe}$	*has*	blond hair, blue eyes
Son_{Stripe}	*has*	blond hair, glasses, brown eyes
Dog_{Stripe}	*has*	blond hair, fur, brown eyes
Mom_{Stripe}	*parent_of*	Son_{Stripe}, $Daughter_{Stripe}$
Dad_{Stripe}	*parent_of*	Son_{Stripe}, $Daughter_{Stripe}$
$Daughter_{Stripe}$	*daughter_of*	Mom_{Stripe}, Dad_{Stripe}
Son_{Stripe}	*son_of*	Mom_{Stripe}, Dad_{Stripe}
$Daughter_{Stripe}$	*sister_of*	Son_{Stripe}
Son_{Stripe}	*brother_of*	$Daughter_{Stripe}$
Dog_{Stripe}	*pet_of*	Son_{Stripe}
Son_{Stripe}	*owner_of*	Dog_{Stripe}
Dog_{Stripe}	*walked_by*	Son_{Stripe}
Mom_{Solid}	*is_named*	Sollid, Mom_{Solid}
Dad_{Solid}	*is_named*	Solid, Dad_{Solid}
$Daughter_{Solid}$	*is_named*	Solid, $Daughter_{Solid}$
Son_{Solid}	*is_named*	Solid, Son_{Solid}
Dog_{Solid}	*is_named*	Solid, Dog_{Solid}
Mom_{Solid}	*is_a*	Solid, human, adult, parent, mom
Dad_{Solid}	*is_a*	Solid, human, adult, parent, dad
$Daughter_{Solid}$	*is_a*	Solid, human, child, daughter
Son_{Solid}	*is_a*	Solid, human, child, daughter
Dog_{Solid}	*is_a*	Solid, dog
Mom_{Solid}	*has*	grey hair, green eyes
Dad_{Solid}	*has*	grey hair, mustache, glasses, brown eyes
$Daughter_{Solid}$	*has*	brown hair, ponytail, green eyes
Son_{Solid}	*has*	brown hair, green eyes
Dog_{Solid}	*has*	brown hair, fur, brown eyes
Mom_{Solid}	*parent_of*	Son_{Solid}, $Daughter_{Solid}$
Dad_{Solid}	*parent_of*	Son_{Solid}, $Daughter_{Solid}$
$Daughter_{Solid}$	*daughter_of*	Mom_{Solid}, Dad_{Solid}
Son_{Solid}	*son_of*	Mom_{Solid}, Dad_{Solid}
$Daughter_{Solid}$	*sister_of*	Son_{Solid}
Son_{Solid}	*brother_of*	$Daughter_{Solid}$
Dog_{Solid}	*pet_of*	$Daughter_{Solid}$
$Daughter_{Solid}$	*owner_of*	Dog_{Solid}
Dog_{Solid}	*walked_by*	??? ($Daughter_{Solid}$)

Table 8. These tree-part propositions represent knowledge about the Stripes' and Solids' families. They are training patterns for the network in Simulation 3.

Finally, the members of each family sit in various relations to one another. For example, the mother in the Stripe family is the "mother_of" the son and daughter in the Stripe family. While there is between-family overlap, across all training patterns in all contexts the similarity of each person to all their family members is greater than to any non-family member (see Figure 18).

The underlying model parameters were identical in all of the simulations that I present. In all cases, the learning rate was .005 and the network was trained for 10,000 epochs. Results were averaged over 10 runs of each network in order to provide statistical tests. The hidden layers were identical in each case: there were 6 Subject Representation units and 16 Integration units. In all presented simulations, error on the training patterns was very low by the end of training (average cross-entropy error < .35).

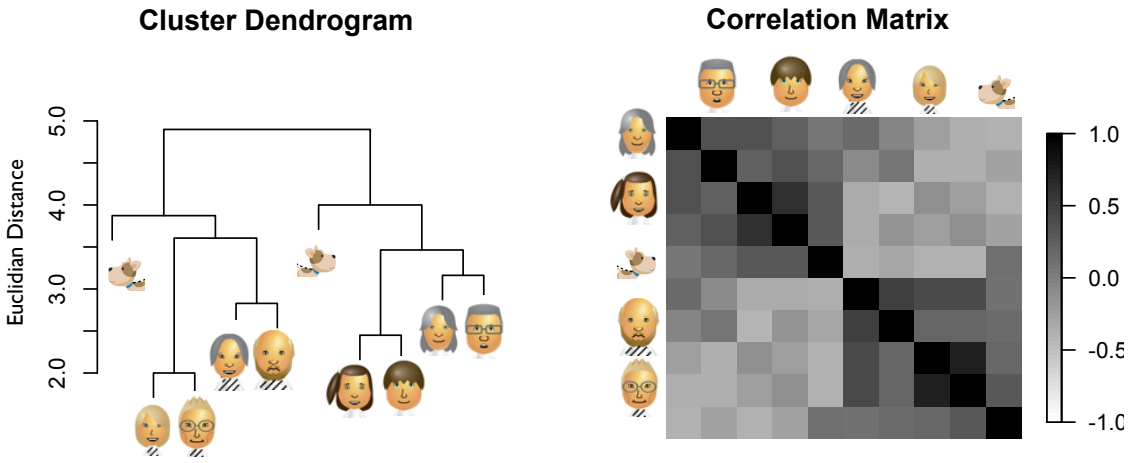

Figure 18. The left panel displays a hierarchical cluster of the training patterns used for Simulation 3. The right panel displays a correlation matrix of these same training patterns. The training patterns are structured such that any given member of either family is most similar to another member of the same family. Additionally, despite mostly non-overlapping inputs and outputs, the families are structured analogously.

Simulations 1-3: The basic model

In the first simulation, the network learns about the Stripes and the Solids, with a single fact omitted about the Solid family. However, the families are otherwise very

similar, with isomorphic relational structures (i.e., they have the same family members in the same relationships to one another). In particular, the daughter of the Stripes both owns the dog and walks the dog. While the network knows that the daughter of the Solids owns their dog, it receives no information about who walks their dog. This network does a good job of learning the facts on which it is trained, but the question of interest is whether it can extend its knowledge to answer a question on which it received no training: Who walks the Solids' dog?

I can contrast two major predictions. Naively, one might think that the network runs on raw association. As the Solids' dog is most similar to the Stripes' dog, the network should conclude that the Stripes' daughter walks the Solids' dog! Alternatively, one might expect that the network will encode the relational structure between the two families, and so will correctly conclude that the person in the appropriate position within the Solid family -- namely, the Solid's daughter -- will be the one who walks their dog. In fact, the network decides that within the Solid family, the daughter walks the dog. A paired t-test contrasting the activation levels of the Stripes' daughter with the Solids' daughter is highly significant, $t[9] = 7.75$, $p < .001$ (see Figure 19).

It might be the case that the network is not driven by the relational similarity between the two families, but rather by some non-obvious feature of the input within the Solid family alone. For example, perhaps the fact that the daughter owns the dog creates enough of an association that the network would also conclude that she walks the dog, even without drawing any analogical inferences from the Stripe family. To control for this, I ran a second simulation, in which the model was trained only on the Solid family, with no information about the Stripe family. In this network, the model does not conclude that the daughter walks the dog. Instead, it defaults to a different kind of mapping on which it has also been trained: the identity mapping. Without any further information, the network decides that the dog walks itself! A paired t-test contrasting the activation levels of the Solids' dog with the Solids' daughter is highly significant, $t[9] = 5.61$, $p < .001$ (see Figure 19).

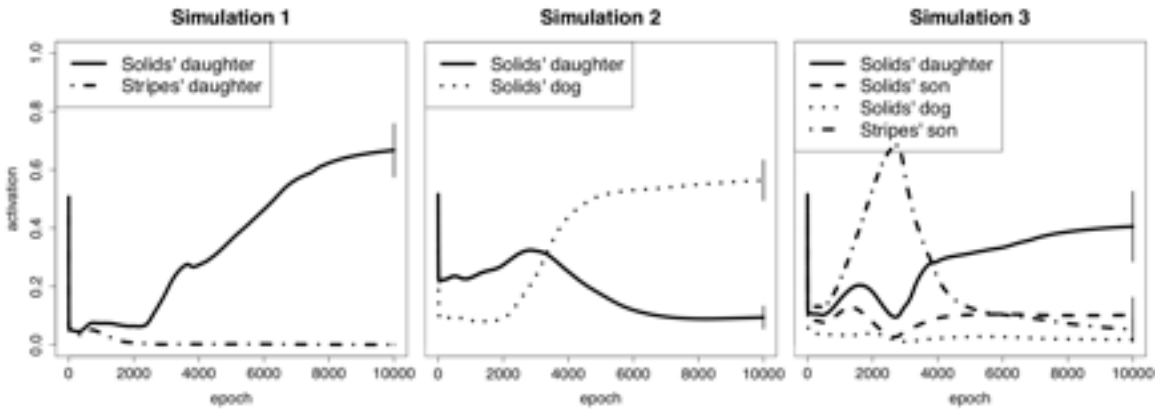

Figure 19. Each of the three panels above displays activation levels for the target units for the first three simulations in response to the inference pattern over the course of learning. The left panel illustrates results from the first simulation when the daughter walks the dog in both the Stripe and Solid family; the middle panel illustrates results from the second simulation when the model only learns about the Solid family; and the right panel illustrates results from the third simulation when the son walks the dog in the Stripe family and the daughter walks the dog in the Solid family.

Simulations 1 and 2 do not, however, distinguish another set of predictions. One possibility is that the network has learned to align the two families with respect to their relational structure, but only in an exact way. On this account, the model may have placed both mothers, both daughters, and both dogs in exactly the same structure, perhaps a tree structure, with a dimension dividing the families from each other, driven by the *overall* relational similarity between the families. Another possibility is that the network learns to associate certain common features, such as the features shared among the analogous members of each family, in order to drive its success on the relational questions. On either of these views, the network should only be able to align the structures between the two families when the correspondence is exact, or nearly so.

On the other hand, perhaps the network has learned the details of the family relations within each family as well as across families. In this case, it could learn a regularity like "whoever owns the dog, walks the dog," which is driven neither by perfect, global structural alignment nor by associations between surface features. This kind of relational binding is closely related to those tasks that previous researchers have argued can only be done using a distinct mapping mechanism operating over explicit

symbols (Gentner & Markman, 1993, 1995; Holyoak & Hummel, 2000; Hummel, 2010; Markman, 1999). Therefore, it would be a surprising and exciting finding if this network were able to succeed in such an abstract relational mapping task.

In order to distinguish between these hypotheses, I ran a third simulation, very similar to the first, except that in the Stripe family, the *son*, not the daughter, both owns and walks the dog. When the network is informed that the daughter of the Solids owns the dog, but is not told who walks it, what should the network conclude? If the network needs to align each member of each family exactly, it should overlay the two dogs in the same place relative to each family, and conclude that the Solids' son walks their dog. Similarly, if overall association of the dog to certain features (perhaps those that the sons share) is driving learning, then the Solids' son should again walk the dog. However, if the network is learning the details of the relational structure, and in particular the regularity between owning a dog and walking it, then the network should succeed in inferring that the Solids' daughter walks the dog.

This is precisely what occurs. Separate tests contrasting the activation level of the Solids' daughter with the activation level of the Stripes' son, $t[9] = 2.58$, $p < .05$, the activation level of the Solids' son, $t[9] = 2.95$, $p < .05$, and the dog, $t[18] = 3.35$, $p < .01$, are all significant (see Figure 19). This demonstrates that raw co-occurrence, or other simple associative processes which are often believed to underlie the performance of error-driven learning models (e.g., Hummel, 2010 in reply to Ramscar, Yarlett, Dye, Denny, & Thorpe, 2010), are not the key to learning in this model. It is, however, interesting to notice that the Stripes' son is the model's choice early in training, suggesting that the network first tends to make judgments predominately based on surface similarity, but over time shifts towards judgments based on relational similarity. This "relational shift" has been widely observed in the literature on the development of analogical reasoning abilities (Gentner, 1988; Goswami, 1992; Ratterman & Gentner, 1998). Intriguingly, this pattern is observed throughout the various simulations presented in this paper.

Simulations 4-6: Extending the model

I have presented a basic set of simulations showing that the model succeeds in performing analogical inference from a family that is fully described to one that is less fully described. In the simulations below, I will extend the basic model in several directions, addressing possible objections to the claim that the model really is succeeding at analogical inference. Each of these models will extend the third simulation, in which the son of one family owns and walks the dog, and the task of the model is to infer that the daughter of the other family, who owns the dog, also walks it.

Inexact match - Can the model align non-isomorphic structures? In the previous simulations, each family had the same number of family members, sitting in the same (or extremely similar) relationships to one another. One can investigate the extent to which the network relies on perfectly overlaying the two families by making the family structures only approximately match. In the fourth simulation, the Stripes have three children, two sons and a daughter, and one of the sons again owns (and walks) their dog. The Solid family still has two children, one son and one daughter, and their daughter owns the dog. The model continues to make the inference that she probably walks the dog as well. A paired t-test contrasting the activation levels of the Solids' daughter with the Stripes' son is highly significant, $t[9] = 4.28$, $p < .01$ (see Figure 20). This demonstrates that the network can learn to draw inferences over structures that are only partially alignable, which has been shown to be important for analogical reasoning in previous work (such as Falkenhainer, Forbus, & Gentner, 1989).

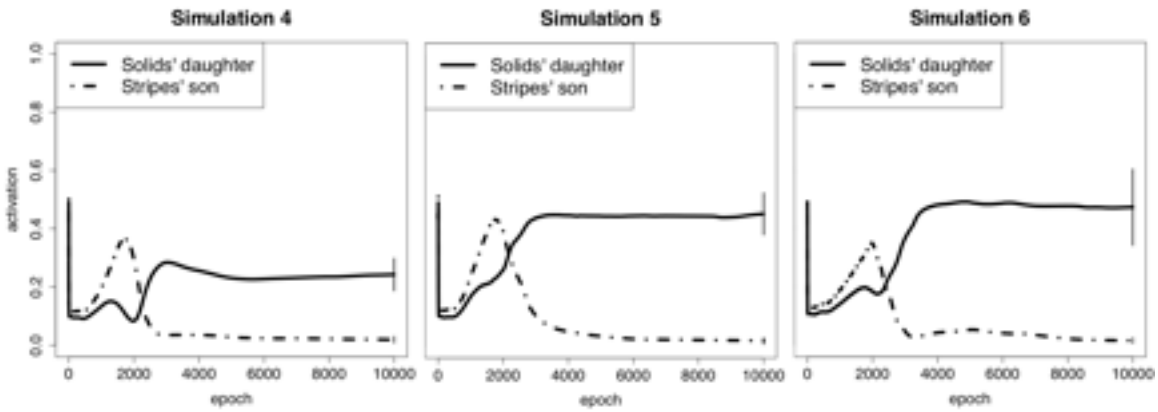

Figure 20. The three panels above display activation levels for the target units for simulations four, five, and six over the course of learning. The left panel illustrates results from the fourth simulation in which the family trees have relational structures that are less readily alignable; the middle panel illustrates results from the fifth simulation in which the network takes distributed input representations of the family members; and the right panel illustrates results from the sixth simulation in which there are no overlapping output units.

Distributed inputs - Does the model rely on implementing symbols? I have claimed that the success of this network depends on its development of distributed, sub-symbolic representations, with which it can integrate the perceptual and the relational information about the family members within a high-dimensional representational space. Others might argue instead that the network is simply implementing symbols, and succeeds by performing some syntax-like transformation on those symbols. Such an argument may point to the localist input units which represent the family members. The localist inputs are a useful simplification, but that focusing on them is a distraction, as the network can never directly exploit these localist units. Instead, it is required to re-represent each item as a pattern of activation over a hidden layer, as described above.

To make this point more clearly, I ran a fifth simulation which had distributed, rather than localist, input representations for the family members. Following a model by Rogers and McClelland (2004), these were simply chosen to be the perceptual features of each family member. For instance, whereas in the first four simulations the mother in the Stripes family was represented by a localist unit corresponding to her identity, in the fifth

simulation she was represented by a collection of units (e.g. “grey hair”, “ponytail”, “brown eyes”) that describe her physical appearance. This should not assist the network in acquiring the relational structure; if anything, it should bias the network towards the surface-level perceptual features for generalization. Nevertheless, the network still infers that the owner of the dog walks it, transferring from the Stripes’ son to the Solids’ daughter. A paired t-test contrasting the activation levels of the Solids’ daughter with the Stripes’ son was highly significant, $t[9] = 6.05$, $p < .001$ (see Figure 20).

Non-overlapping outputs - Does the model require perceptual overlap? On the other hand, one might argue that the architecture is biased in the opposite direction: the more direct overlap between the two families at the feature level (that is, the Output layer), the less work the model needs to do to align their structures. What if only the relational similarity is available, as might be the case when constructing analogical mappings across very different domains of knowledge? This kind of analogy may be critical for explaining how analogy can subserve cognition and reasoning more generally (e.g., Gentner, 2010b).

To test this, I constructed a sixth simulation that had completely non-overlapping output units. The network essentially had two copies of each output property, so that each family’s target representations were totally distinct. To succeed in generalizing the relation between the two families, the network would need to align the structures even in the absence of any surface-level similarity between the two families. And this is precisely what it did. Again, when the network is told that, in the Stripe family, the son owns and walks the dog, it concludes that for the Solids, the owner of the dog -- the daughter -- must also walk it. A paired t-test contrasting the activation levels of the Solids’ daughter with the Stripes’ son was significant, $t[9] = 3.58$, $p < .01$ (see Figure 20).

Discussion

To summarize the results of the above simulations, I have demonstrated that analogical inference can emerge from a domain-general, distributed connectionist model of semantic learning and reasoning. Critically, this analogical inference: (1) is driven by generalization from a source domain to a target domain; (2) relies on abstract relational structure, not surface-level similarities or direct featural associations or co-occurrences; (3) parallels important features of the development of analogy in children; (4) can operate over structures which only approximately match, or which are only partially alignable; and (5) exploits structural similarity even in the absence of explicit feature overlap, allowing the possibility of cross-domain analogical inference in guiding learning.

How is it that a connectionist model that lacks symbolic representations and an explicit mapping mechanism can succeed at this kind of analogical inference task? As I have demonstrated in several variations of the model, it is not due to any direct co-occurrence of features. Neither is it due to any kind of surface-level similarity between the items. In the extreme case, the two families can share absolutely no output units, and the model will still draw on the appropriate relational structure to make novel inferences. I argue that part of the answer involves the progressive differentiation of its representations over the course of development. Initially, all the weights are set to very small random values, so the network essentially treats every family member, and every relation, as being the same. Over the course of training, the model learns to "pull apart" those representations that must be differentiated in order to produce the right answers. However, it only does so in response to erroneous predictions. One crucial constraint on this process that the model embodies is a particular architectural design that forces inputs from each domain to pass through the same sets of weights and hidden units. Any changes to the weights that influence one representation will also tend to affect similar representations. This biases the network to reuse as much representational structure as it can get away with. Alternative architectures that do not enforce the same general constraints fail to capture these patterns of learning (Rogers & McClelland, 2004).

In this particular network, the families share a great deal of structural similarity. If the family members are represented as points in a high-dimensional space, which is one approximate characterization of the network's hidden representations, then the members of each family sit in similar positions relative to the other family (or families). That is, the father and mother in the Stripes bear the same relationship to each other and to their children as do the father and mother in the Solids. An efficient representation of these similarities is to use one dimension to separate the families from each other, and the remaining dimensions to capture the relational structures common to each family (indeed, another model that learns family trees settles on exactly this kind of representation: Hinton, 1986).

As a result, the network's representations of the families become aligned over the course of training, because this allows the network to learn more efficiently (i.e. to reduce error more quickly). The side effect of this representational overlap is that when the network learns a fact about one family (e.g. one dog's owner walks it), the representations of the members of the *other* family (e.g. between that dog and its owner) get to come along for the ride. This is not to say that the model is stuck with its first guess about the structure of the world. As I indicated in the description of simulation 3, and as is visible in other simulations as well, the model undergoes a developmental shift from predominantly perceptual to predominantly relational inference when the environment warrants such a shift (a finding I discuss in more detail below).

We can observe the process of progressive differentiation in this network by looking at a clustering diagram and a correlation matrix across the Subject Representation layer at different points in time for simulation 3 (see Figure 21). These are different ways of visualizing how the model perceives the similarity between items throughout the course of learning. Early in training, the network groups items essentially at random, since the weights are initialized to very small values. Later in training, the network's representations capture both the surface similarities and the relational similarities between items (in contrast to the training patterns, depicted in Figure 21). Progressive

differentiation in semantic networks has been explored more extensively in previous work (Flusberg, Thibodeau, Sternberg, & Glick, 2010; Rogers & McClelland, 2004).

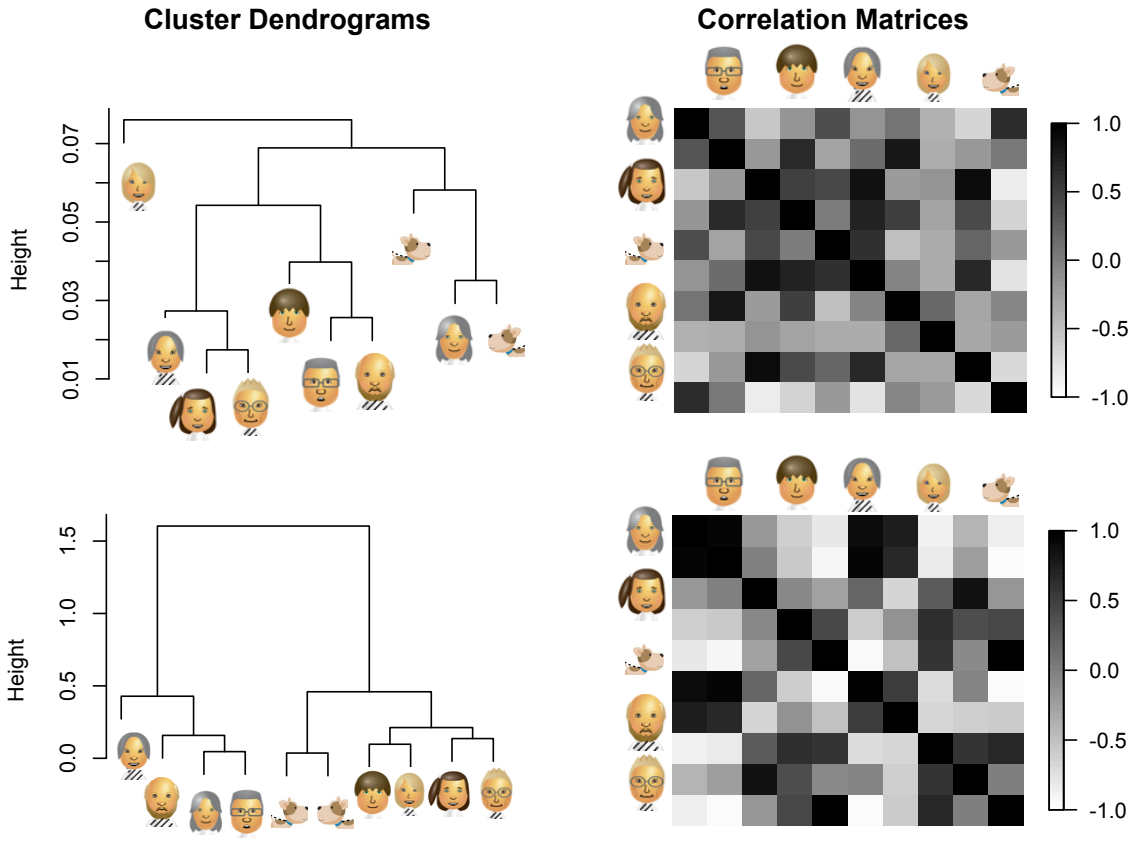

Figure 21. Each of the plots above illustrates the similarity structure of the learned Subject representations in simulation 3. Hierarchical clusters are on the left and correlation matrices are on the right. Early in training (the upper panels), the network does not group individuals by family or relation. Later in training, at 1,300 epochs (the lower panels), the network has aligned the families according to their relational similarity.

Several proponents of the *Structured* approach to analogy have suggested that a defining feature of analogical reasoning is the ability to perform *structured pattern completion* (Gentner & Markman 1993; 1995), which refers to a process whereby "a partial representation of the target is completed based on its structural similarity to the base" (Gentner & Markman, 1995). This is typically defined in contrast to *simple pattern*

completion, which is "based on the vector similarities of the current activation pattern to previously learned patterns" (Gentner & Markman, 1993). Emergent models are frequently criticized for only being able to perform simple pattern completion, which is unable to account for a great deal of sophisticated human behavior (as I detailed in the introduction; see also Gentner & Markman, 1995). As I have shown, however, the model behaves in a way that is perfectly captured by the concept of structured pattern completion: it draws inferences based on information from patterns that do not share strong (or, in some cases, any) similarity in terms of raw vector overlap. Rather, the patterns that come to most strongly influence the model's inferences about novel inputs are those that share the most similar structural relationships with the other training items.

I have argued that this approach reveals that a sub-symbolic model that lacks an explicit mapping mechanism, can, in principle, do some forms of analogical or metaphorical reasoning. In the following section I will use the same model and the same principles to simulate an empirical demonstration of metaphoric reasoning.

A Case Study of Embodied Conceptual Metaphor: Time as Space

One the best documented cases of how abstract thinking can be metaphorically structured by concrete experience comes from the domain of time (Clark, 1973; Lakoff & Johnson, 1980; McGlone & Harding, 1998; Boroditsky, 2000, 2001; Boroditsky & Ramscar, 2002; Gentner et al., 2002; Evans, 2004; Matlock et al., 2005; Casasanto & Boroditsky, 2008). The language we use to talk about time is heavily infused with the language we use to talk about space, as when we talk about a long meeting or two birthdays being close together (Clark, 1973; Lakoff & Johnson, 1980). Indeed, our actual perception of space can influence how we experience and reason about time (Casasanto & Boroditsky, 2008). For example, Casasanto and Boroditsky (2008) found that the length (in spatial extension) of a line on a computer screen affected how long (in temporal duration) it was judged to remain on the screen: the longer the line, the longer the time.

Like many other abstract, complex domains, there is more than one system of metaphor for talking and thinking about time (Clark, 1973; Lakoff & Johnson, 1980; Gentner et al., 2002). For instance, we can imagine ourselves moving forward through time, like when we talk about coming up on the holidays (ego-moving perspective), but we can also imagine remaining stationary as time moves toward us, like when we talk about the holidays fast approaching (time-moving perspective). Some spatial words that we use to talk about temporal events are ambiguous because they can be interpreted differently depending on which metaphorical perspective is adopted. For example, if you are told that Wednesday's meeting has been moved forward 2 days and you had adopted the ego-moving perspective, you would conclude that the meeting is now on Friday.

However, if you had adopted the time-moving perspective you would conclude that the meeting is now on Monday (McGlone & Harding, 1998; Boroditsky, 2000; Boroditsky & Ramscar, 2002). Several experiments have demonstrated that the way people are currently thinking about space directly affects which of these perspectives they select and therefore how they reason about time (Boroditsky, 2000; Boroditsky & Ramscar, 2002). For example, people who are asked the Wednesday's meeting question at an airport are more likely to take the ego-moving perspective (i.e., give the Friday response) because they are about to take a flight (i.e., move through space) than when they are waiting to pick someone up (i.e., someone is approaching them in space; Boroditsky & Ramscar, 2002).

Here, I will use this empirical finding to demonstrate how this *Emergent* approach to metaphorical reasoning can capture an empirical effect of metaphoric reasoning.

General Modeling Framework

In the following two simulations, the network can be thought of as an agent experiencing its world. Over the course of "training" the agent repeatedly experiences events in the world, predicts their outcomes, and learns something about how the actual events differ from its predictions. The environment and the agent are simplified so as to

render the learning process tractable, while still retaining those aspects of environmental structure which are crucial for producing the phenomena the model is supposed to explain, and to make it possible to analyze what the agent has learned (for a discussion of this issue, see McClelland, 2009).

In this model, the environment consists of the various items in the world that the agent experiences in their various relational contexts (collectively forming the input patterns), together with the subsequent states of the world that the network attempts to predict (the target output patterns). The network that comprises the agent is wired up in a strictly feed-forward fashion, as in the previous set of simulations. While in reality agents interact with the world in a dynamic fashion, for simplicity I consider only one portion of this dynamic interaction. On each trial, the agent experiences some portion of the world (e.g., that it is standing in a particular section of space and moving in a particular direction), makes a prediction about what it will experience next (e.g., that it will encounter another particular section of space), and learns about the ways in which it was incorrect, thereby improving future predictions.

Importantly, the network's knowledge is stored in the weights between the layers. When a pattern of activation occurs across one of the layers, that activation propagates forward through the weights to the next layer. The patterns of activation at the input layers are thought of as multimodal sensory-motor input from the environment. In the Item layer, these inputs stand for the experience of physical locations in space and temporally extended events such as the days of the week or a meeting. In the Relation layer, the inputs stand for different kinds of relationships that these items can have to each other; for example, I might ask the network what day is earlier than Wednesday, or what section of space is West of the blue section.

The network is instantiated with small random weights connecting each of the layers. As a result, its internal representations of all items and all relations will be similar, and therefore its predictions about the world will be the same for all inputs.

Whenever the network's output fails to match the target pattern, however, it receives an error signal in proportion to the squared output error. This error signal

informs the network both when it has predicted events that do not occur and when it has failed to predict an event that did occur. In practice, this error signal serves to adjust the weights from the inputs to the outputs in proportion to the error that they caused, using the standard backpropagation learning algorithm.

Since different input patterns predict different events "in the world," the network will gradually learn to differentiate the items from each other, and the relations from each other. This process of differentiation is driven by differences in what the various items predict about what else may happen in the world, not directly by, for example, the degree of overlap in the perceptual inputs (for related views, see Gibson & Gibson, 1955; Rogers & McClelland, 2008; Ramscar et al., 2010). However, wherever there is similarity between different items, these similarities will be encoded in the learned, distributed representations. The "similarity," as I will show, can be similarity either in the explicit overlap between their predictions or in the systematic structural relationships among the various items within a domain. These internal representations therefore capture, in a graded and sub-symbolic fashion, both the similarities and the differences between the items. In the simulations that follow, I examine whether this framework can account for some of the empirical findings from the conceptual metaphor literature.

Simulation 7

In Simulation 7, the network learns about space and time through experience trying to predict how space and time are structured in the model environment (see Table 9 for model specifications). In the simplified environment of the model, space is laid out along a single dimension running from West to East (unlike our own environment, in which space is three-dimensional and also includes north and south, up and down! See Figure 22).

	Sim 7	Sim 8
Units in Item layer	11	10
Units in Relation layer	6	6
Units in Learned Item Representation	7	7
Units in Learned Relation representation	4	4
Units in Integration layer	9	9
Units in Output layer	11	10
Initial weight range	-.05/.05	
Activation function	Sigmoid	
Error measure	SSE	
Learning rate	0.1	
Momentum	0	

Table 9. Network parameters for Simulations 7 and 8.

To make the simulation easier to talk about and understand, space is divided into sections of different colors, going from red to blue to green to yellow as you move toward the East. Throughout the course of training the network will attempt to learn that two relations – East of and West of – structure the spatial arrangement of the colored sections in the environment. Training proceeds by asking the model to predict what color section of space it will "see" if it looks toward the East or West of its current position (and adjusting the weights in proportion to the error of this prediction, as described above). In practice, this works by presenting the network with one or more items along with a relation in the input layers and asking it to generate all appropriate outputs. For instance, if the network were presented with blue and West of it would have to output green and yellow (i.e., what the network would see if it looked toward the East while standing on the blue section: the sections of space that the blue section lies to the West of).

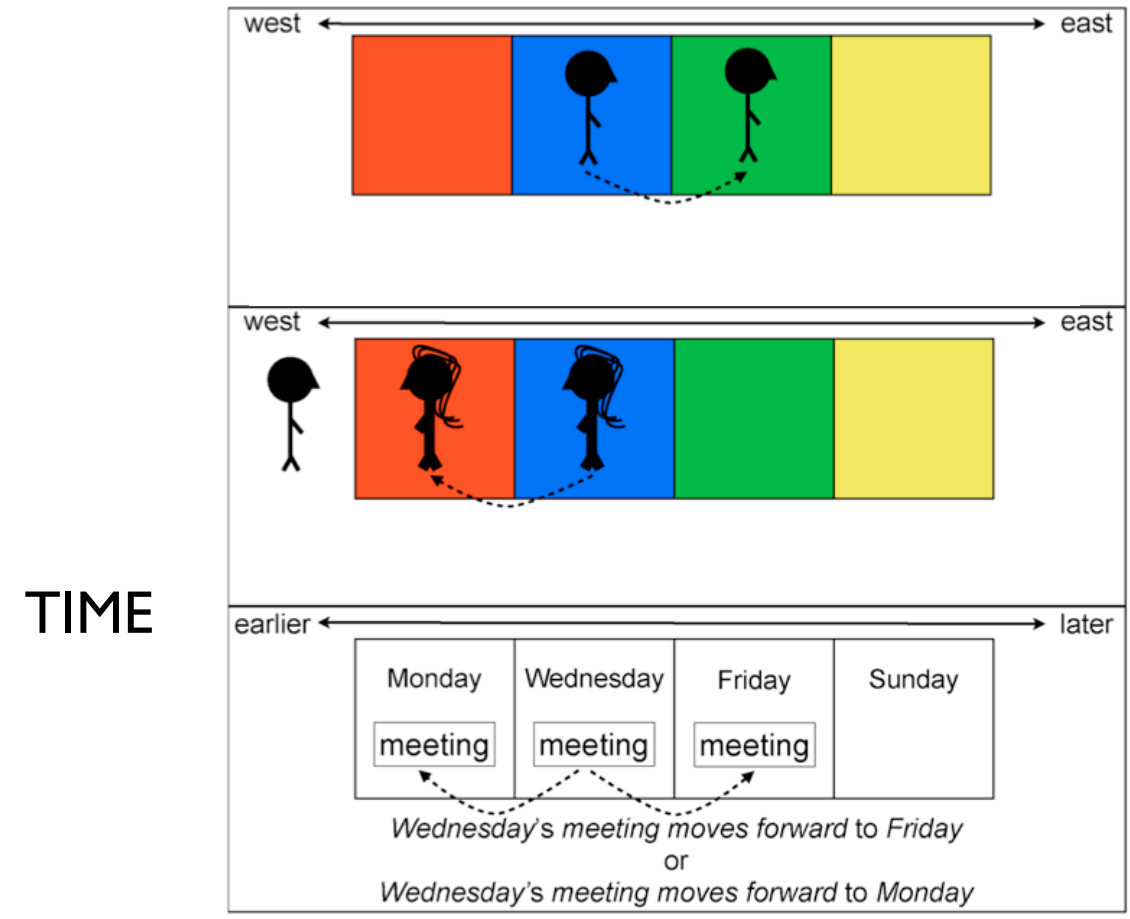

Figure 22. An illustration of the structure of the spatial and temporal relations in the model environment. The network learns about the consequences of both itself and other agents moving in the environment, though movement in the temporal domain is ambiguous.

Time is also laid out in a single dimension from earlier to later events. Time is divided up into distinct moments, the days of the week, which follow a specific temporal sequence (going from Monday to Wednesday to Friday to Sunday as you progress later in time). During training the network will attempt to learn that two relations – earlier than and later than – structure the temporal sequence of the days of the week. The network learns about time in the same way that it learns about space. Thus, if the network were presented with Monday and earlier than as inputs it would have to generate Wednesday, Friday, and Sunday as outputs (i.e., the days of the week that Monday is earlier than).

Crucially, the network enjoys a richer, more structured set of experiences in the spatial domain because it can observe the consequences of its own movements in space (as well as the consequences of the movements of other agents in the environment). Like most mobile organisms, the network has both a front and back and can move forward and backward in space.

To keep things simple, let us imagine that the network is standing on the blue section of space facing toward the East. If the network moves forward, it will move toward the East end up on the green section of space, while if it moves backward it will move toward the West and end up on the red section of space. However, forward and backward movements in space are not simply the same as moving toward the East or West. Now imagine that the network is observing another agent in the environment that is standing on the blue section of space and facing West. If this other agent moves forward, it will move toward the West end up on the red section of space, while if it moves backward it will move toward the East and end up on the green section of space. Thus the network has to learn that the consequences of moving backward and forward in space depends on whether it is attending to its own movements or to the movements of another agent. In practice, this works by including self and other items in the input layer to let the network know whose movements it is observing (see Figure 22). To keep things simple, I assume that the model is always facing toward the East and the other agent in the environment is always facing toward the West.

While the effects of movement in the spatial environment are unambiguous in the presence of either the self or other context, the model's experience of "movement" in the temporal domain is ambiguous in that there is no consistent mapping between forward/backward and earlier/later. The model learns that when a Wednesday meeting moves forward, it sometimes is moved to Monday, and other times it is moved to Friday. The same can occur when a meeting moves backward. Structuring the temporal domain in this way allows us to study to the ambiguity explored in Boroditsky (2000) and Boroditsky and Ramscar (2002). In particular, while the network has no experience with the self/other distinction in the temporal domain, I can examine whether it can use its experience

with the effects of these contexts in the spatial domain to resolve the ambiguity of "movement" in the temporal domain. That is, I can test whether activating a particular spatial frame of reference (i.e., the self or other perspective) in the context of reasoning about a temporal event (i.e., moving the Wednes day meeting forward) will influence the network's expectations about the effects of "movement" in the temporal domain.

Results. In this simulation I explored whether an ambiguity in the temporal domain (i.e., that a Wednes day meeting sometimes moves forward to Monday and sometimes moves forward to Friday) can be resolved by activating a particular spatial frame of reference. That is, even though the model has no experience with the self/other distinction in the temporal domain, I can nevertheless activate one of these spatial frames of reference in the temporal domain when asking the model whether it thinks Wednes day meetings move forward to Monday or Friday. If these reference frames influence the model's interpretation of moves forward in a way that is consistent with empirical results (Boroditsky, 2000; Boroditsky & Ramscar, 2002), one would expect that including self as an input (along with Wednes day, meeting, and moves forward) would yield relatively more activation in the Friday output unit than the Monday output unit. Alternatively, one would expect that including other as an input instead would result in relatively more activation in the Monday output unit than the Friday output unit.

To investigate this, I exposed the network to 10,000 epochs of training in the simplified environment, at which point I froze the weights to prevent further learning and began the testing phase. The statistics reported for both simulations include activation values that have been averaged across 10 instances of the model to ensure that any effects are not the result of a random bias in a particular instance.

First, I tested whether the network had learned the unambiguous spatial and temporal structure of its environment by presenting it with the same input–output pairings that it was trained on. Indeed, the network performed quite well on this test (mean tss = 2.31, *SD* =0.32)[13], demonstrating that it had correctly learned the features of its environment that it had been directly exposed to during training. Next I tested whether

[13] The imposed ambiguity in the temporal domain (i.e., that *Wednesday's* meeting can move both *forward* and *backward* to *Monday* or *Friday*) made it impossible for the network's tss to improve beyond 2.

including a spatial frame of reference (i.e., self vs. other) influenced the network's predictions for the effect of moves forward in the temporal domain. I measured this effect by comparing three test patterns: (1) the ambiguous pattern that the network was trained on in which the Wednes day and meeting items were paired with the moves forward relation, (2) this same pattern with the self item included as an input, and (3) this same pattern with the other item included as an input (instead of self).The model learned that moves forward was ambiguous in the temporal domain when the self and other items were not included as inputs. Specifically, when tested on the ambiguous pattern, the model fully predicted meeting on the output ($M = 0.981$, $SD = 0.0043$) and partially predicted both Monday ($M = 0.500$, $SD = 0.0075$) and Friday ($M = 0.503$, $SD = 0.0069$). No other units had average activations greater than 0.02. A regression model that predicted output activation of the two target units (Monday and Friday) with contrast-coded predictors for Day (Monday, Friday) and Perspective (self, other) as well as a Day s Perspective interaction term, was fit to the two test patterns. Both main effects were significant (Day: $B = 0.062$, $p < 0.05$; Perspective: $B = 0.100$, p 0.01) as was the interaction term ($B = 0.081$, $p < 0.01$), indicating that including the perspective units shifted the degree to which the model predicted that Wednesday's meeting would move forward to Monday or Friday (see Figure 23).

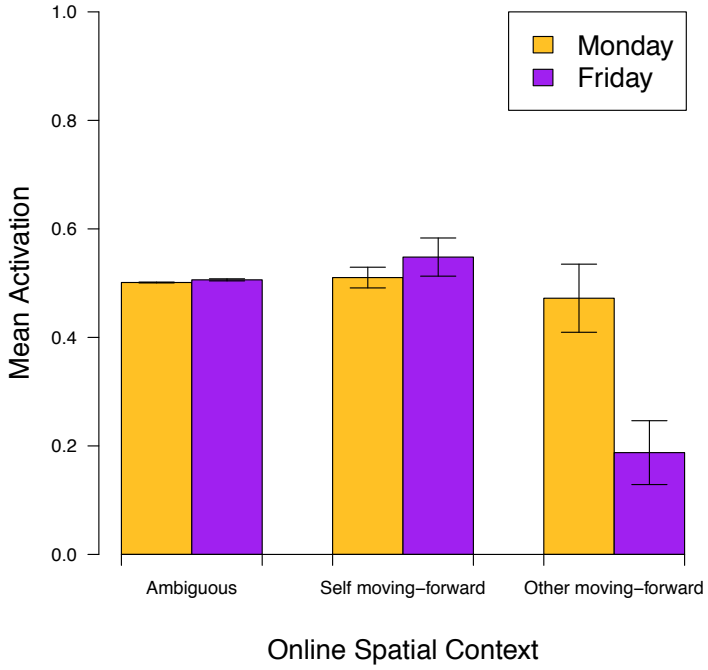

Figure 23. Results of Simulation 7 showing that activating a particular spatial frame of reference biases the network's predictions about movement in the temporal domain.

Simulation 8

The seventh simulation investigated whether the network would learn to metaphorically map its relatively rich experience with space onto the parallel but experientially impoverished domain of time in order to resolve an ambiguous temporal reasoning task.

In Simulation 8, I explore whether the network can learn to map the directionality of time (from earlier to later) onto other spatial cues in the environment (e.g., the directionality of space, from Wes to East). Several studies have demonstrated that culturally specific spatial cues, such as writing direction (Fuhrman & Boroditsky, 2010) and absolute spatial coordinate systems (Boroditsky & Gaby, 2010), can influence and structure how people think about the directional "flow" of time. The model was set up in a very similar manner as in Simulation 7. However, where Simulation 7 included an

ambiguity in the temporal domain, Simulation 8 removes that ambiguity in order to closely align the meanings of the temporal and spatial relations.

In particular, the moves forward relation was made unambiguous in both the spatial and temporal domains, by removing the other item. In the temporal domain, moves forward always predicted that the Wednes day meeting should occur on Friday, never Monday, and moves backward always predicted that Wednesday meetings should occur on Monday, never Friday. This might be interpreted as a culturally specific bias, analogous to the experience of reading temporally sequenced material like calendars and comics from left to right (or even writing direction itself, see Fuhrman & Boroditsky, 2010). In the spatial domain, I removed the patterns in which the other agent moves forward from blue to red and moves backward from blue to green, again rendering the situation unambiguous.

Removing these four patterns, two from the temporal domain and two from the spatial domain, leaves the moves forward relation consistent with the earlier than relation in the temporal domain and with the West of relation in the spatial domain. This can be seen in the predicted outcomes of the events: moves forward from Wednesday predicts Friday, and Wednesday is earlier than Friday, and so on for the spatial domain.

If the model is sensitive to the structural similarities present in this environment, it should learn that the West of and earlier than relations make similar predictions in their respective domains, as do the East of and later than relations. As a result, the learned distributed representations for these pairs of relations should become similar as a function of experience – allowing, for instance, spatial words like East of and West of to be sensibly interpreted in the temporal domain (e.g., Wednesday is East of Monday or Wednesday is West of Friday). The model only ever observes West of and East of in the spatial domain, and earlier than and later than in the temporal domain, so an account based on direct co-occurrence would not generate the same prediction. This would provide a demonstration of how culturally specific features of the environment such as writing direction or dominant spatial coordinate systems could come to organize our representations of abstract domains such as time.

Results. In the eighth simulation, I explored whether the network could in principle learn to map the directionality of time (from earlier to later) onto the directionality of space (from Wes t to East). In order to clearly explore this possibility, I modified the model's environment slightly from that of Simulation 7 so that the moves forward relation was consistent with later than in the temporal domain and East of in the spatial domain (and so that moves backward was consistent with earlier than and Wes t of). If the network is able to take advantage of this similarity in a way that is consistent with empirical findings (e.g., Fuhrman & Boroditsky, 2010; Boroditsky & Gaby, 2010), then it should be able to interpret, for example, the relations East of and West of in the domain of time (i.e., Wednesday is East of Monday andWednesday is West of Friday) even though these relations were never explicitly paired with temporal inputs or outputs in training.

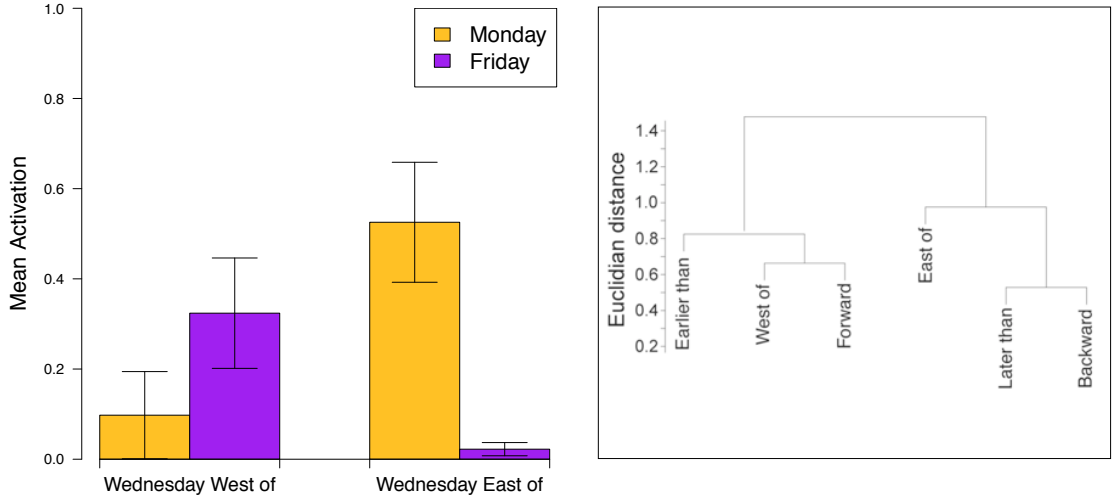

Figure 24. Left panel: the results of Simulation 8 showing that spatial relations that were never directly experienced in the temporal domain can still structure temporal reasoning. Right panel: The structure of the Learned Relation Representation Layer in Simulation 8.

I investigated this issue by exposing the network to 10,000 epochs of training and then freezing the weights. At this point the network had learned to make correct predictions for each of the training patterns (mean tss = 0.0898, *SD* = 0.0375). I then

presented the network with two novel test patterns: one pairing Wednes day with East of on the input, the other pairing Wednes day with Wes t of on the input.

When Wednes day was paired with the East of relation, the model predicted Monday ($M = 0.525$, $SD = 0.421$) more than Friday($M = 0.022$, $SD = 0.046$), whereas when Wednes day was paired with the Wes t of relation, the model predicted Friday ($M = 0.324$, $SD=0.387$) more than Monday ($M=0.098$, $SD=0.306$). A regression model that predicted output activation with contrast-coded predictors for Day (Monday, Friday) and Relation (East of, Wes t of) as well as an interaction term was fit to the two test patterns. Neither main effect was significant (Day: $B = 0.032$, $p = 0.54$; Relation: $B = 0.069$, $p = 0.19$); however, the interaction was significant ($B=0.182$, $p < 0.01$), indicating that these relation units held meaning in the domain of time even though this was the first time the model had encountered them in the temporal context (Figure 4).In order to determine why this effect occurred, I submitted the representations for each of the relations of interest (i.e., West of, East of, earlier, later, moves forward, and moves backward) in the Learned Representation Layer to a hierarchical clustering analysis (shown in Figure 5). This analysis shows that the representation of West of is very similar to the representation of earlier, and the representation of East of is similar to the representation of later. It also shows that moves forward is more similar to West of and earlier, and moves backward is more similar to East of and later.

Discussion

Both model simulations successfully learned a representation of the temporal structure of the world based partially on their experience with space. In Simulation 7, this allowed the network to resolve an ambiguity in the temporal domain by relying on additional structure only present in the spatial domain: a true application of conceptual metaphor to aid cognition. In Simulation 8, the network's representations of the spatial and temporal domains were shaped by a structural homology between the domains: in this case, a "culturally driven" bias to scan from Wes t to East through time. In both

cases, the network's metaphoric concepts were not driven by direct co-occurrence between concepts within the domains (e.g., distance with duration). Rather, the available information for learning the metaphor was the second-order relations between items within each domain (e.g., things move around in space in a similar way to how events can be sequenced in time).

Several theorists have proposed that the grounding of abstract thought in concrete knowledge may be due in part to the direct co-occurrence of certain domains in experience, for example, of time with space or of love with physical warmth (Lakoff and Johnson, 1999; see also the afterward to the 2003 edition of Lakoff and Johnson, 1980). The model presented here demonstrates another, more general, yet equally grounded pathway to metaphor: structural similarity between the target and the source domain (see also Boroditsky, 2000). However, this use of structural similarity is not a distinct online, rule-based algorithm operating over symbolic representations, as in other theories of structural alignment in metaphor comprehension and analogical reasoning (e.g., Falkenhainer et al., 1989); rather, it is a result of the gradual process of differentiation that takes place over the entire course of learning. It is fair to say that the network's knowledge of time is partially constituted by the learned structural relations in the spatial domain. This is demonstrated by the metaphoric remapping between time and space, which, in this model, share almost no input (only the moves forward and moves backward units) and no output units at all. Merely having distributed representations, as most connectionist models do, is not sufficient for this kind of behavior to emerge; the process by which those representations were acquired through experience with the environment is also critical.

To understand why the remapping occurs, recall that the model initially treats all items and relations as equivalent (due to its small and random initial weights) and only discriminates objects as it is forced to do so by the flow of information from the world. Over the course of this differentiation process, the model constructs several high-dimensional and highly overlapping representations for the items, the relations, and the item–relation conjunctions, all passing through the same sets of weights and patterns of

activation over the same sets of units. If the network can reuse certain dimensions of its representations because of similarity in the structural relationships between and among items and contexts, it will tend to do so. Since the spatial and temporal domains share most, though not all, of their respective relational structure, the network learns a partially unified representation for the two domains. These overlapping representations, which are a direct result of the differentiation process, give rise to the influence of the concrete spatial experience on the abstract temporal reasoning task.

This model demonstrates that the structural homology between domains of experience is one aspect of the environment that can drive generalization (or, more properly, partial lack of differentiation) for metaphorical inference. But it is not the only way that metaphors can be learned. As mentioned above, co-occurrence of more abstract with more concrete domains of experience may also cause the learner to build metaphorical semantic representations.

This is because the experience with the abstract domain will often predict properties that are also predicted during experience with the concrete domain, which may drive the representations to become more similar than they would otherwise be. In Simulation 8, it is indeed co-occurrence that drives learning, but it is indirect, not direct, co-occurrence. The moves forward unit in this simulation is unambiguously similar to West of when it occurs in the spatial domain, and to earlier than when it occurs in the temporal domain.

Notice that West of and earlier than never predict similar outputs in a way that would cause them to become similar, so this is not a matter of raw co-occurrence. Still, the model is encouraged to draw its representations of moves forward, West of, and earlier than into a similar semantic structure because these relations must be re-represented in the Learned Representation Layer. In a parallel fashion, moves backward draws together the cross-domain relations East of and later than. The bridging between structures can occur because of similarity in the structural relationships among the items within each domain, or because of some direct or indirect co-occurrence (or co-prediction) in the environment, or both.

Another mechanism that drives metaphorical structuring is language use. There are at least two routes through which language can bring about metaphoric alignment, one slow and one fast. In the current framework, linguistic experience is considered to be another aspect of the environment (this point is discussed in more detail below). The influence of language here would be across the (slow) course of development, serving as additional scaffolding for similar high-order structures (as in Simulation 7) or as an indirect co-occurrence cue or outcome (as in Simulation 8).

On the other hand, language might be used online to point out a novel metaphorical structural mapping, such as "an atom is like the solar system." The agent's task is then to take two existing semantic structures and figure out what aspects of those structures the conversational partner intends to highlight. The model deals with a very slow process of learning and differentiation, but does not have a way of rapidly integrating new information, so this kind of novel metaphorical language is a problem in this model. However, I am not claiming that the model describes the whole story, and any model of learning like ours will eventually need to take into account fast-learning processes as well (McClelland et al., 1995). The model is nevertheless a novel contribution to the literature, as existing models of metaphor (and analogy) that do deal with online structural alignment (e.g., Falkenhainer et al., 1989) do not attempt to slowly integrate structural information over the course of development.

Limitations of the Emergent Approach

While I have argued that this modeling approach has successfully captured a number of important features of human analogical and metaphorical inference, this is not to say that the model can account for all of human analogical and metaphorical reasoning. Far from it! I have only attempted to capture the sort of analogical or metaphorical inference that emerges as a result of a gradual learning process over the course of development. This may include any task that requires making inferences based on shared relational structure that has been stored in long term memory, including phenomena like

reasoning with conceptual metaphors. Other researchers have captured additional aspects of analogical reasoning in theoretically related *Emergent* models. For example, Leech and colleagues (2008) used a recurrent connectionist architecture to simulate analogical reasoning in tasks that take the classic "A:B::C:D" form.

That being said, many of the analogy tasks used in previous work, which models like SME and LISA can capture so well, seem to rely on cognitive processes which I did not even attempt to simulate (Bowdle & Gentner, 1997; Clement & Gentner, 1991; Markman & Gentner, 1997; Morrison et al., 2004). For example, I would agree that some analogy tasks may rely on strong working memory and cognitive control processes, as well as one-shot learning and episodic memory (e.g. Gick & Holyoak, 1980). Though I have highlighted the fact that the model performs structured pattern completion, this is not the only defining feature of analogy (Genter & Markman, 1993; 1995). For instance, mature analogical reasoning is highly flexible, and allows for multiple interpretations of a single item in different comparisons, as well as multiple interpretations of any given comparison (Gentner & Markman, 1995). This sort of flexibility lies outside the scope of the particular model and the task I set out to simulate, and it may in fact require additional, online cognitive capacities like working memory (Morrison et al., 2004).

Moreover, some aspects of analogical reasoning may demand much richer linguistic abilities than I implemented in the model. As noted above, I treat relational language as an environmental cue, encoding a certain kind of statistical structure that is then used to shape semantic representations. While this is one important role of language in analogical reasoning, it may not be the only one; the ability to verbally re-describe a situation to oneself, for example, is an important tool in many higher-level reasoning tasks (Williams & Lombrozo, 2010).

These issues point to a possible heterogeneous view of analogy. While some forms of analogical inference may naturally emerge over the course of development due to the operation of low-level learning mechanisms, other forms of analogical reasoning may only be possible by exercising additional processes like cognitive control and working memory. Therefore, I would like to suggest that one major unsolved problem in

the field is the integration of the kind of slow-learning semantic cognition model described in this paper with the online, structurally explicit models already in place. The extensive and valuable work on models such as SME and LISA over the past twenty years, no less than the connectionist models I have implemented, must be used to guide future research into analogical processing across development, in behavior, and in the brain.

Summary of Simulations 1-8

In eight simulations I have grounded analogical and metaphorical reasoning in a sub-symbolic account of learning general semantics in an attempt to provide an *Emergent* approach to analogical inference. This may seem surprising, as analogical inference has been thought to require explicit symbolic representations as well as a distinct mapping mechanism operating over those symbols (Gentner, 2010a; Gentner & Markman, 1993, 1995; Holyoak & Hummel, 2000; Hummel, 2010; Markman, 1999).

As I have alluded to already, there are several strengths of this approach. First, it offers a more parsimonious account of many behavioral findings because it does not require positing additional complex cognitive machinery (Rogers & McClelland, 2004). Second, it allows us to naturally address the development of cognitive functioning and to link this up with what we know about biological and neural development more generally (Elman et. al, 1996; Leech et. al, 2008; Spencer, Thomas, & McClelland, 2009). In fact, *Structured* models of analogy have been criticized for failing to account for how the mechanisms they instantiate could possibly develop in the first place (Leech et. al, 2008). Finally, the use of distributed and graded representations more accurately reflects the graded and quasi-regular nature of both the environment we live in and human behavior in general (Rogers & McClelland, 2008; Spivey, 2007). Even tasks that appear to require discrete responses are often supported by more graded and dynamic cognitive mechanisms (Spivey, 2007).

Chapter 5
Conclusions

Metaphor is a powerful tool for reasoning about abstract and complex domains. Previous work has shown that deliberately using metaphors and analogies can powerfully shape the way we learn and think in a variety of domains (Blackwell, Trzeniewski, & Dweck, 2007; Boyd, 1979; Gentner & Jeziorski, 1993; Green, 1993; Kuhn, 1979; Mayer, 1993; Petrie & Oshlag, 1979; Pylyshyn, 1979; Sticht, 1993).

In this dissertation, I investigated whether the conventional metaphors that suffuse natural language, and often go unnoticed, also influence the way we reason about important issues. Across several experiments, I found that that these fleeting and unnoticed metaphors instantiate complex knowledge structures and influence people's reasoning in a way that is similar to the role that schemas (Rumelhart, 1975; Rumelhart & Ortony, 1977), scripts (Abelson, 1981; Schank & Abelson, 1975), and frames (Minsky, 1975) have been argued to play in reasoning and memory (Bartlett, 1932; Bransford & Johnson, 1972; Liberman, Samuels, & Ross, 2004; Ortony, Schallert, Reynolds, & Antos, 1978).

In Chapter 2 I presented five experiments in which the virus and beast metaphors (a) provided people with a structured framework for understanding crime in Addison, (b) influenced the inferences that they made about the crime problem, and (c) suggested different causal interventions for solving the problem. This was true even though the metaphors themselves did not strike the participants as particularly influential.

Further, in Chapter 3, I presented six experiments that uncovered some of the factors and processes that support metaphoric reasoning. Specifically, I found that the metaphor (i.e., "Crime is a virus/beast"), and not the lexical units "virus" or "beast," led people towards different conceptions of the crime problem. I also found that the influence of these metaphors depended on their being processed early, rather than late, in a stream of processing. As a result, I speculated that the way in which these metaphors work is by shading how people constructed an on-line representation of the crime problem. For instance, words like "defense systems" and "vulnerable" may take on one interpretation after reading that crime is a virus and a different meaning after reading that crime is a beast, leading people to conceptualize the crime problem differently on the two framings.

These different ways in which people conceptualize the crime problem then invite different solutions to the crime problem.

Together, these experiments suggest that metaphors dynamically affect reasoning. The meaning that a metaphor takes on is shaped by its surrounding context. And a metaphor's surround context will be influenced by the meaning it takes on.

In Chapter 4, implemented this theory in a connectionist model to further explore how metaphors come to influence reasoning. I found evidence that an overlapping, distributed scheme for conceptual representation was capable of supporting some forms of metaphorical reasoning without recourse to a specific mapping or binding mechanism. Given the neural evidence for overlapping, distributed representations (Dilkina, McClelland, & Plaut, 2008; Rogers & McClelland, 2004; Rogers et al., 2004), these simulations suggest that metaphor may be a natural byproduct of our system of conceptual representation.

This work also has important implications that go beyond cognitive science. It suggests that the metaphors we use to talk about social issues are vitally important to the way we reason about, attempt to solve, and behave in relation the fundamental issues facing the country and world today. In short, good metaphors lead to effective problem solving; bad metaphors make problems worse.

References

Abelson, R.P. (1981). Psychological status of the script concept. *American Psychologist, 36,* 715-729.

Adorno, T.W., Frenkel-Brunswik, E., Levinson, D.J., & Sanford, R.N. (1950). *The Authoritarian Personality.* New York: Harper and Row.

Bartlett, F.C. (1932). *Remembering: A study in experimental and social psychology.* Cambridge, England: Cambridge Univ. Press.

Benet-Martinez, V., & John, O.P. (1998). Los cinco grandes across cultures and ethnic groups: Multitrait-multimethod analyses of the Big Five in Spanish and English. *Journal of Personality and Social Psychology, 75,* 729-750.

Blackwell, L.S., Trzeniewski, K.H., & Dweck, C.S. (2007). Implicit theories of intelligence predict achievement across an adolescent transition: A longitudinal study and an intervention. *Child Development, 78,* 246-263.

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

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

Boroditsky, L. (2000). Metaphoric Structuring: Understanding time through spatial metaphors. *Cognition, 75,* 1-28.

Boroditsky, L. (2001). Does language shape thought? English and Mandarin speakers' conceptions of time. *Cognitive Psychology, 43,* 1-22.

Boroditsky, L., & Gaby, A. (2010). Absolute spatial representations of time in an aboriginal Australian community. *Psychological Science, 11,* 1635-1639.

Boroditsky, L., & Ramscar, M. (2002). The Roles of Body and Mind in Abstract Thought. *Psychological Science, 13,* 185-188.

Bowdle, B.F., & Gentner, D. (1997). Informativity and asymmetry in comparisons. *Cognitive Psychology, 34,* 244–286.

Bowdle, B.F., & Gentner, D. (2005). The career of metaphor. *Psychological Review, 112,*

193-216.

Boyd, R. (1979). Metaphor and theory change: What is 'metaphor' a metaphor for?. In A. Ortony (ed.). *Metaphor and Thought.* New York: Cambridge University Press.

Bransford, J.D., & Johnson, M.K. (1972). Contextual prerequisites for understanding: Some investigations of comprehension and recall. *Journal of Verbal Learning and Verbal Behavior, 11,* 717-726.

Buhrmester, M.D., Kwang, T., & Gosling, S.D. (2011). Amazon's Mechanical Turk: A new source of inexpensive, yet high-quality, data?. *Perspectives on Psychological Science, 6,* 3-5.

Cacioppo, J.T., & Petty,R.E. (1982). The need for cognition. *Journal of Personality and Social Psychology, 42,* 116-131.

Cacioppo, J.T., Petty, R.E., & Kao, C.F. (1984). The efficient assessment of need for cognition. *Journal of Personality Assessment, 48,* 306-307.

Carey, S. (2009). *The Origin of Concepts.* Oxford University Press.

Casasanto, D., & Boroditsky, L. (2008). Time in the mind: using space to think about time. *Cognition, 106,* 579–593.

Charteris-Black, J. (2011). *Politicians and Rhetoric: The Persuasive Power of Metaphor.* New York: Macmillan Publishers.

Chen, Z. (1996). Children's analogical problem solving: the effects of superficial, structural, and procedural similarity. *Journal of Experimental Child Psychology, 62,* 410-431.

Christie, S. & Gentner, D. (2010). Where hypotheses come from: Learning new relations by structural alignment. *Journal of Cognition and Development, 11,* 3, 356-373.

Clark, E. V. 2003. *First language acquisition*. Cambridge: Cambridge University Press.

Clark, H.H. (1973). “Space, time, seman- tics, and the child,” in *Cognitive Development and the Acquisition of Language*, ed. T. E. Moore (New York, NY: Academic Press), 27–63.

Clark, H.H. (1996). *Using Language.* New York: Cambridge University Press.

Clement, C. & Gentner, D. (1991). Systematicity as a selection constraint in analogical

mapping. *Cognitive Science, 15*, 89–132.

Currie, J. (2005). The marketization of depression: The prescribing of SSRI antidepressants to women. *Women and Health Protection,* May, 1-27.

Den Boer, E. (1998, August). The frequency of original metaphors in literary and nonliterary texts. Paper presented at the biannual conference of the Internationale Gesellschaft fur Empirische Literaturwissenschaft, Utrecht University, Utrecht, The Netherlands.

Dilkina, K., McClelland, J. L., & Boroditsky, L. (2007). How language affects thought in a connectionist model. In McNamara, D. S. and Trafton, J. G., editors, *Proceedings of the 29th Annual Cognitive Science Society,* pages 215–220, Austin, TX. Cognitive Science Society.

Doumas, L., Hummel, J., & Sandhofer, C. (2008). A theory of the discovery and predication of relational concepts. *Psychological Review, 115,* 1–43.

Elman J.L. (2004). An alternative view of the mental lexicon. *Trends in Cognitive Science, 8,* 301-306.

Elman, J.L., Bates, E.A., Johnson, M.H., Karmiloff-Smith, A., Parisi, D. & Plunkett, K. (1996). *Rethinking innateness: A connectionist perspective on development.* Cambridge, MA: MIT Press.

Evans, V. (2004). *The Structure of Time: Language, Meaning and Temporal Cognition.* Amsterdam: John Benjamins.

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

Federal Bureau of Investigation. Uniform crime report. (2008). - http://www.fbi.gov/ucr/cius2008/data/table_01.html

Flusberg, S.J., Thibodeau, P.H., Sternberg, D.A., & Glick, J. J. (2010). A connectionist approach to embodied conceptual metaphor. *Frontiers in Psychology,* 1(0):12.

Fuhrman, O., & Boroditsky, L. (2010). Cross-cultural differences in mental representations of time: Evidence from an implicit nonlinguistic task. *Cognitive Science, 34,* 1430–1451.

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

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

Gentner, D. (1988). Metaphor as structure mapping: The relational shift. *Child development, 59,* 47–59.

Gentner, D. (2010a). Bootstrapping the mind: Analogical processes and symbol systems. *Cognitive Science, 34,* 752–775.

Gentner, D. (2010b). Psychology in cognitive science: 1978–2038. *Topics in Cognitive Science, 2,* 328–344.

Gentner, D. & Jeziorski, M. (1993). The shift from metaphor to analogy in Western science. In A. Ortony (ed.). *Metaphor and Thought.* New York: Cambridge University Press.

Gentner, D., & Forbus, K.D. (2011). Computational models of analogy. *WIREs Cognitive Science, 2*, 266-276.

Gentner, D., & Gentner, D. R. (1983). Flowing waters or teeming crowds: Mental models of electricity. In D. Gentner & A. L. Stevens (Eds.), *Mental models* (pp. 99-129). Hillsdale, NJ: Lawrence Erlbaum Associates. (Reprinted in M. J. Brosnan (Ed.), *Cognitive functions: Classic readings in representation and reasoning.* Eltham, London: Greenwich University Press).

Gentner, D., & Markman, A. (1993). Analogy-watershed or waterloo? Structural alignment and the development of connectionist models of analogy. In S.J. Hanson, J.D. Cowan, & C.L. Giles (Eds.), *Advances in Neural Information Processing Systems 5* (pp. 855-862). San Mateo, CA: Morgan Kaufmann Publishers.

Gentner, D., & Markman, A. (1995). Analogy-based reasoning in connectionism. In M. Arbib (Ed.), *The Handbook of Brain Theory and Neural Networks* (pp. 91–93). MIT Press, Cambridge, MA.

Gentner, D., & Markman, A. B. (1993). Analogy - watershed or waterloo? structural

alignment and the development of connectionist models of cognition. In *Advances in Neural Information Processing Systems 5* (pp. 855–862). San Francisco, CA, USA. Morgan Kaufmann Publishers Inc.

Gentner, D., Bowdle, B., Wolff, P., & Boronat, C. (2001). Metaphor is like analogy. In D. Centner, K.J. Holyoak, & B.N. Kokinov (Eds.) T*he analogical mind: Perspectives from cognitive science* (pp. 199-253). Cambridge MA, MIT Press.

Gentner, D., Imai, M., & Boroditsky, L. (2002). As time goes by: evi- dence for two systems in processing space → time metaphors. *Language and Cognitive Processes, 17,* 537–565.

Gentner, D., Imai, M., & Boroditsky, L. (2002). As time goes by: Evidence for two systems in processing space time metaphors. *Language and Cognitive Processes, 17,* 537-565.

Gentner, D., Rattermann, M.J., & Forbus, K.D. (1993). The roles of similarity in transfer: separating retrievability from inferential soundness. *Cognitive Psychology, 25,* 524–75.

Gibbs, R. (1980). Spilling the beans on understanding and memory for idioms in conversation. *Memory & Cognition, 8,* 149–156.

Gibbs, R. (1994). *The poetics of mind: Figurative thought, language, and understanding.* New York: Cambridge University Press.

Gibbs, R.W., & O'Brien, J.E. (1990). Idioms and mental imagery: The metaphorical motivation for idiomatic meaning. *Cognition*, 36, 35-68.

Gibson, J., & Gibson, E. (1955). Perceptual learning: differentiation or enrichment? *Psychological Review, 62,* 32–41.

Gick, M.L., & Holyoak, K.J. (1980). Analogical problem solving. *Cognitive Psychology, 12,* 306-355.

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

Goswami, U. (1992). *Analogical reasoning in children.* New York: Psychology Press.

Graesser, A., Long, D., & Mio, J. (1989). What are the cognitive and conceptual

components of humorous texts? *Poetics, 18,* 143-164.

Green, T.F. (1993). Learning without metaphor. In A. Ortony (ed.). *Metaphor and Thought.* New York: Cambridge University Press.

Hall, R.P. (1989). Computational approaches to analogical reasoning: A comparative analysis. *Artificial Intelligence, 39,* 39-120.

Harris, R. (1976). Comprehension of metaphor: A test of a two stage processing model. *Bulletin of the Psychonomic Society, 8,* 321–324.

Hinton, G. (1986). Learning distributed representations of concepts. In *Proceedings of the Eighth Annual Conference of the Cognitive Science Society,* 1-12. Amherst, MA.

Holyoak, K., & Hummel, J. (2000). The proper treatment of symbols in a connectionist architecture. In E. Dietrich & A. Markman (Eds.). *Cognitive dynamics: Conceptual and representational change in humans and machines* (pp. 229–264). Lawrence Erlbaum Associates, Hillsdale, NJ.

Holyoak, K., & Hummel, J. E. (2008). No way to start a space program: Associationism as a launch pad for analogical reasoning. *Behavioral and Brain Sciences, 31,* 388-389.

Holyoak, K.J., & Koh, K. (1987). Surface and structural similarity in analogical transfer. *Memory & Cognition, 15,* 332-340.

Hummel, J.E. (2010). Symbolic versus associative learning. *Cognitive Science, 34*, 958–965.

Hummel, J.E. (2011). Getting symbols out of a neural architecture. *Connection Science, 23*, 109-118. Hummel, J.E. & Holyoak, K. (1997). Distributed representations of structure: A theory of analogical access and mapping. *Psychological Review, 104*, 427–466.

Hummel, J.E., & Holyoak, K.J. (1997). Distributed representations of structure: A theory of analogical access and mapping. *Psychological Review, 104,* 427-466.

Hummel, J.E., & Holyoak, K.J. (2003). A symbolic-connectionist theory of relational inference and generalization. *Psychological Review, 110,* 220-264.

Hummel, J.E. & Holyoak, K.J. (2005). Relational reasoning in a neurally plausible

cognitive architecture. *Current Directions in Psychological Science, 14*, 153-157.

Jeffers, R.J., & Lehiste, I. (1979). *Principles and methods for historical linguistics.* Cambridge: MIT Press.

John, O.P., Donahue, E.M., & Kentle, R.L. (1991). *The Big Five Inventory--Versions 4a and 54.* Berkeley: University of California at Berkeley, Institute of Personality and Social Research.

John, O.P., Naumann, L.P., & Soto, C.J. (2008). Paradigm shift to the integrative Big Five trait taxonomy: History, measurement, and conceptual issues. In O.P. John, R.W. Robbins, & L.A. Pervin (Eds.), *Handbook of Personality: Theory and Research* (pp. 114-158). New York: Guilford.

Kahle, D. (February 3, 2012). From timy seeds of ideas, a mighty economy grows. The Register-Guard. Retrieved from http://projects.registerguard.com/web/opinion/27555782-47/eugene-grass-kahle-money-seeds.html.csp

Kelling, G. (1991). Crime and metaphor: Toward a new concept of policing. *City Journal, 1.*

Kemp, C., & Tenenbaum, J.B. (2008). Structured models of cognition. *Behavioral and Brain Sciences, 31,* 717-718.

Keysar, B., & Bly, B. (1995). Intuitions of the transparency of idioms: Can one keep a secret by spilling the beans? *Journal of Memory and Language, 34,* 89–109.

Keysar, B., & Bly, B. (1999). Swimming against the current: Do idioms reflect conceptual structure? *Journal of Pragmatics, 31,* 1559–1578.

Keysar, B., Shen, Y., Glucksberg, S., & Horton, W.S. (2000). Conventional language: How metaphorical is it? *Journal of Memory and Language, 43,* 576-593.

Kuhn, T.S. (1979). Metaphor in science. In A. Ortony (ed.). *Metaphor and Thought.* New York: Cambridge University Press.

Lakoff, G. (2002). *Moral Politics: How Liberals and Conservatives Think.* Chicago: University of Chicago Press.

Lakoff, G. (2004). *Don't think of an elephant!: Know your values and frame the debate-the essential guide for progressives.* White River Junction, VT: Chelsea Green

Publishing Company.

Lakoff, G. (2007). *Whose freedom?* New York: Picador.

Lakoff, G. (2008). *The Political Mind: Why you can't understand 21st century politics with an 18th century brain.* New York, NY: Penguin Books.

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

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

Lakoff, G., & Nuñez, R. (2001). *Where mathematics comes from: How embodied mind brings mathematics into being.* New York: Basic Books.

Leech, R., Mareschal, D., & Cooper, R. (2008). Analogy as relational priming: A developmental and computational perspective on the origins of a complex cognitive skill. *Behavioral and Brain Sciences, 31,* 357–378.

Liberman, V., Samuels, S.M., & Ross, L. (2004). The name of the game: Predictive power of reputations versus situational labels in determining prisoner's dilemma game moves. Personality and Social Psychology Bulletin 30, 1175-1185.

Liptak, A. (April 23, 2008). Inmate count in U.S. dwarfs other nations'. *New York Times.*

Loftus, E.F. (1975). Leading questions and the eyewitness report. *Cognitive Psychology, 7,* 560-572.

Marcus, G.F., (2001). *The Algebraic Mind: Integrating Connectionism and Cognitive Science.* Cambridge, MA: MIT Press.

Marcus, G.F., & Keil, F.C. (2008). Concepts, correlations, and some challenges for connectionist cognition. *Behavioral and Brain Sciences, 31,* 722-723.

Markman, A.B. (1999). *Knowledge Representation.* Mahwah, NJ: Lawrence Erlbaum Associates.

Markman, A.B. & Gentner, D. (1997). The effects of alignability on memory. *Psychological Science, 8*, 363–367.

Markman, A.B. & Laux, J.P. (2008). Analogical inferences are central to analogy. Commentary on Leech, Mareschal & Cooper. *Behavioral and Brain Sciences, 31,* 390-391.

Marr, D. (1982). Vision. San Francisco: Freeman.

Matlock, T., Ramscar, M., & Boroditsky, L. (2005). On the experiential link between spatial and temporal lan- guage. *Cognitive Science, 29,* 655–664.

Mayer, R.E. (1993). The instructive metaphor: Metaphoric aids to students' understanding of science. In A. Ortony (ed.). *Metaphor and Thought.* New York: Cambridge University Press.

McClelland, J.L. (2009). The place of modeling in cognitive science. *Topics in Cognitive Science, 1,* 11-38.

McClelland, J.L., & Rogers, T. (2003). The parallel distributed processing approach to semantic cognition. *Nature Reviews Neuroscience, 4,* 310-322.

McClelland, J.L., McNaughton, B.L., & O'Reilly, R.C. (1995). Why there are complementary learning systems in the hippocampus and neocortex: Insights from the successes and failures of connectionist models of learning and memory. *Psychological Review, 102*, 419–37.

McGlone, M., & Harding, J. (1998). Back (or forward?) to the future: the role of perspective in temporal language comprehension. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 24,* 1211–1223.

Mial, D.S. (1987). Metaphor and affect: The problem of creative thought. *Metaphor and Symbolic Activity, 2*, 81-96.

Milbank, D. (April 18, 2012). Democrats battle back against Republican 'war on women'. The Washington Post. Retrieved from http://www.washingtonpost.com/opinions/democrats-battle-back-against-republican-war-on-women/2012/04/18/gIQAX1llRT_story.html

Minsky, M. (1975). A framework for representing knowledge. In P.H. Winston (Ed.) *The psychology of computer vision.* New York: McGraw-Hill.

Morrison, R., Krawczyk, D., Holyoak, K., Hummel, J., Chow, T., Miller, B., & Knowlton, B. (2004). A neurocomputational model of analogical reasoning and its breakdown in frontotemporal lobar degeneration. *Journal of Cognitive Neuroscience, 16*, 260–271.

Murphy, G. L. (1996). On metaphoric representations. *Cognition, 60,* 173–204.

Murphy, G. L. (1997). Reasons to doubt the present evidence for metaphoric representation. *Cognition, 62,* 99–108.

Mutafchieva, M. & Kokinov, B. (2007). Does the Family Analogy Help Young Children To Do Relational Mapping. In *Proceedings of the European Cognitive Science Conference.*

Myers, R.H. (1990). *Classical and Modern Regression with Applications*. New York: Duxbury Press.

Nayak, N.P., & Gibbs, R.W. (1990). Conceptual knowledge in the interpretation of idioms. *Journal of experimental psychology: General, 119,* 315 – 330.

Nersessian, N. (1992). How do scientists think? Capturing the dynamics of conceptual change in science. In R.N. Giere (Ed.), *Cognitive Models of Science: Minnesota Studies in the Philosophy of Science, XV.* Minneapolis, MN: University of Minnesota Press.

Obama, B.H. (2010, September 8). *Remarks by the President on the economy in Parma, Ohio.* Retrieved from http://www.whitehouse.gov/the-press-office/2010/09/08/remarks-president-economy-parma-ohio

Opfer, J.E., & Doumas, L.A.A. (2008). Analogy and conceptual change in childhood. *Behavioral and Brain Sciences, 31,* 723-724.

Ortony, A. (1975). Why metaphors are necessary and not just nice. *Educational Theory, 25,* 45-53.

Ortony, A. (1978). Remembering, understanding, and representation. *Cognitive Science, 2,* 53-69.

Ortony, A., Schallert D.L., Reynolds R.E., & Antos S.J. (1978). Interpreting metaphors and idioms: Some effects of context on comprehension. *Journal of Verbal Learning and Verbal Behavior, 17,* 465-477.

Petrie, H.G., & Oshlag, R.S. (1979). Metaphor in education. In A. Ortony (ed.). *Metaphor and Thought.* New York: Cambridge University Press.

Pinker, S. (2007). *The stuff of thought: Language as a window into human nature.* New

York: Penguin Books.

Pollio, H. R., Barlow, J. M., Fine, H. J., & Pollio, M. R. (1977). *Psychology and the poetics of growth: Figurative language in psychology, psychotherapy, and education.* Hillsdale, NJ: Erlbaum.

Putnam, H. (1973). Meaning and reference. *The Journal of Philosophy, 70,* 699-711.

Pylyshyn, Z.W. (1979). Metaphorical imprecision and the 'top-down' research strategy. In A. Ortony (ed.). *Metaphor and Thought.* New York: Cambridge University Press.

Radu, P.T., McClure, S.M. (In press). Neuroeconomics of addiction: Integrating neuroscience, cognition and behavioral economics. In: *The Encyclopedia of Addictive Behaviors* (P. Miller, Ed.), Elsevier Press: Oxford, UK.

Ramscar, M., Yarlett, D., Dye, M., Denny, K., & Thorpe, K. (2010). The effects of feature-label-order and their implications for symbolic learning. *Cognitive Science, 34,* 909–957.

Rattermann, M. & Gentner, D. (1998). More evidence for a relational shift in the development of analogy: Children's performance on a causal-mapping task. *Cognitive Development, 13*, 453–478.

Rogers, T.T. & McClelland, J.L. (2004). *Semantic Cognition.* MIT Press, Cambridge, MA.

Rogers, T.T. and McClelland, J.L. (2008). Précis of semantic cognition: A parallel distributed processing approach. *Behavioral and Brain Sciences, 31*, 689–749.

Rumelhart, D.E. (1975). Notes on a schema for stories. In D.G. Bobrow & A.M. Collins (Eds.), *Representaion and understanding: Studies in cognitive science.* New York: Academic Press.

Rumelhart, D.E. (1979). Some problems with the notion of literal meanings. In A. Ortony (Ed.), *Metaphor and Thought,* (pp. 71 – 82). Cambridge, England: Cambridge University Press.

Rumelhart, D.E. (1990). Brain style computation: Learning and generalization. In S. Zornetzer, J.L. Davis, & C. Lau (Eds.), *An introduction to neural and electronic networks,* (pp. 405–420). New York: Academic Press.

Rumelhart, D.E. & Todd, P. (1993). Learning and connectionist representations. A*ttention and performance XIV: Synergies in experimental psychology, artificial intelligence, and cognitive neuroscience,* pp. 3–30.

Rumelhart, D.E., & Ortony, A. (1977). The representation of knowledge in memory. In R.C. Anderson, R.J. Spiro, & W.E. Montague (Eds.), *Schooling and the acquisition of knowledge.* Hillsdale, NJ: Lawrence Erlbaum Associates.

Rumelhart, D.E., Hinton, G.E., & McClelland, J.L. (1986). A framework for parallel distributed processing. In J.L. McClelland & D.E. Rumelhart (Eds.) *Parallel Distributed Processing: Explorations in the Microstructure of Cognition, VII.* Cambridge, MA: MIT Press.

Salhani, C. (May 5, 2006). Tango diplomacy: Is war with Iran inevitable? Lebanon Wire. Retrieved from http://khaleejtimes.com/kt-article-display-1.asp?section=opinion&xfile=data/opinion/2006/may/opinion_may17.xml

Schank, R.C., & Abelson R.P. (1975) Scripts, plans and knowledge. In *Advance Papers of the Fourth International Joint Conference on Artificial Intelligence.* Tbilisi, Georgia: USSR, 151 -157.

Schön, D. (1979). Generative metaphor: A perspective on problem-setting in social policy. In A. Ortony (Ed.) *Metaphor and Thought.* Cambridge, UK: University Press.

Searle, J. (1979). Metaphor. In A. Ortony, *Metaphor and thought,* (pp. 83-111). Cambridge, UK: Cambridge University Press.

Shakur, T. (1996). How do you want it. *All Eyez on Me.* Los Angeles, CA: Death Row Records.

Smith, M. K., Pollio, H. R., & Pitts, M. K. (1981). Metaphor as intellectual history: Conceptual categories underlying figurative usage in American English from 1675-1975. *Linguistics, 19,* 911-935.

Spencer, J.P., Thomas, M.S., & McClelland, J.L. (Eds.). (2009). *Toward a new grand theory of development? Connectionism and dynamic systems theory re-considered.* New York: Oxford University Press.

Spivey, J. (2007). *The Continuity of Mind.* Oxford, UK: Oxford University Press.

Sternberg, R.J., Tourangeau, R., & Nigro, G. (1978). Metaphor, induction, and social policy: The convergence of macroscopic and microscopic views. In A. Ortony (ed.). *Metaphor and Thought.* New York: Cambridge University Press.

Sticht, T.G. (1993). *Educational uses of metaphor.* In A. Ortony (ed.). Metaphor and Thought. New York: Cambridge University Press.

Thibodeau, P.H., & Durgin, F.H. (2008). Productive figurative communication: Conventional metaphors facilitate the comprehension of related novel metaphors. *Journal of Memory and Language, 58,* 521-540.

Tversky, A., & Kahneman, D. (1974). Judgment under uncertainty: Heuristics and biases. *Science, 185,* 1124-1131.

U.S. Bureau of Justice Statistics website. Available: http://bjs.ojp.usdoj.gov/content/glance/tables/corr2tab.cfm. Accessed 2011 Jan 18.

Verbrugge, R. R. (1977). Resemblances in language and perception. In R. Shaw & J. Bransford (Eds.), *Perceiving, acting and knowing: Toward an ecological psychology* (pp. 365–389). Hillsdale, NJ: Erlbaum.

Walmsley, R. (2009). *World prison population list* (eighth edition). London, UK: King's College of London.

Williams, J. & Lombrozo, T. (2010). The role of explanation in discovery and generalization: evidence from category learning. *Cognitive Science, 34*, 776–806.

Zaltman, G., & Zaltman, L.H. (2008). *Marketing Metaphoria: What Deep Metaphors Reveal about the Minds of Consumers.* Boston: Harvard Business School Publishing.

Appendix A

What do people recognize as a metaphor?

The following passage is excerpted from a 2010 speech on the economy, "Remarks by the President on the economy in Parma, Ohio," by President Obama.

> It was an America where you didn't buy things you couldn't afford; where we didn't just think about today -– we thought about tomorrow. An America that took pride in the goods that we made, not just the things we consumed. An America where a rising tide really did lift all boats, from the company CEO to the guy on the assembly line.
>
> That's the America I believe in. That's what led me to work in the shadow of a shuttered steel plant on the South Side of Chicago when I was a community organizer. It's what led me to fight for factory workers at manufacturing plants that were closing across Illinois when I was a senator. It's what led me to run for President -– because I don't believe we can have a strong and growing economy without a strong and growing middle class.
>
> Now, much has happened since that election.

Retrieved from: http://www.whitehouse.gov/the-press-office/2010/09/08/remarks-president-economy-parma-ohio

Appendix B

Bootstrapping *p*

Step 1

Randomly assign participants into two groups, X and Y, such that group X has as many participants as were in the virus condition and group Y has as many participants as were in the beast condition.

Step 2

Compute the difference in the proportion of enforcement-oriented suggestions in groups X and Y.

Step 3

Repeat Step 2 10,000 times to establish a distribution of the difference in the proportion of responses that emphasize enforcement that might come about by chance.

Step 4

Compute the difference in the proportion of enforcement-oriented suggestions among participants who read the crime-as-beast and crime-as-virus framings.

Step 5

Use the bootstrapped distribution in Step 3 to test whether the difference attributable to the frame could have come about by chance using the formula below.

In Mathematical Notation

$$d = p(enforce \mid beast) - p(enforce \mid virus)$$
$$d_i' = p(enforce \mid Y) - p(enforce \mid X)$$
$$z = \frac{d - mean(d_i')}{sd(d_i')}$$

Appendix C

Personality Questionnaires

BFI-10

On the BFI-10, people are asked "How well do the following statements describe your personality?" and to respond using a 5-point Likert scale with the following values: Disagree strongly, Disagree a little, Neither agree nor disagree, Agree a little, Agree strongly.

I see myself as someone who...

1. ...is reserved
2. ...is generally trusting
3. ...tends to be lazy
4. ...is relaxed, handles stress well
5. ...has few artistic interests
6. ...is outgoing, sociable
7. ...tends to find fault with others
8. ...does a thorough job
9. ...gets nervous easily
10. ...has an active imagination

Scoring the scales:

Extraversion: 1R, 6; Agreeableness; 2, 7R; Conscientiousness: 3R, 8; Neuroticism: 4R, 9; Openness: 5R, 10 (R = item is reverse-scored).

Need for Cognition Scale

On the NCS, people are asked "For each of the statements below, please indicate whether or not the statement is characteristic of you or of what you believe." They respond on a 5-point Likert scale with the following values: extremely uncharacteristic of me, somewhat characteristic of me, uncertain, somewhat characteristic of me, extremely characteristic of me.

1. I prefer complex to simple problems.
2. I like to have the responsibility of handling a situation that requires a lot of thinking.
3. Thinking is not my idea of fun.
4. I would rather do something that requires little thought than something that is sure to challenge my thinking abilities.
5. I try to anticipate and avoid situations where there is a likely chance that I will have to think independently about something.
6. I find satisfaction in deliberating hard and for long hours.
7. I only think as hard as I have to.
8. I prefer to think about small daily projects to long term ones.
9. I like tasks that require little thought once I've learned them.
10. The idea of relying on thought to make my way to the top appeals to me.
11. I really enjoy a task that involves coming up with new solutions to problems.
12. Learning new ways to think doesn't excite me very much.
13. I prefer my life to be filled with puzzles I must solve.
14. The notion of thinking abstractly is appealing to me.
15. I would prefer a task that is intellectual, difficult, and important to one that is somewhat important but does not require much thought.
16. I feel relief rather than satisfaction after completing a task that requires a lot of mental effort.
17. It's enough for me that something gets the job done; I don't care how or why it works.
18. I usually end up deliberating about issues even when they do not affect me personally.

Scoring the scale: items 3, 4, 5, 7, 8, 9, 12, 16, & 17 are reverse-coded.

The Fascism Scale

On the f-scale, people are asked "For each of the statements below, please indicate whether or not the statement is characteristic of you or of what you believe." They respond on a 5-point Likert scale with the following values: extremely uncharacteristic of me, somewhat characteristic of me, uncertain, somewhat characteristic of me, extremely characteristic of me.

1. Obedience and respect for authority are the most important virtues children should learn.
2. If people would talk less and work more, everybody would be better off.
3. Science has its place, but there are many important things that can never be understood by the human mind.

4. Every person should have complete faith in some supernatural power whose decisions he obeys without question.
5. Young people sometimes get rebellious ideas, but as they grow up they ought to get over them and settle down.
6. What this country needs most, more than laws and political programs, is a few courageous, tireless, devoted leaders in whom the people can put their faith.
7. Nowadays more and more people are prying into matters that should remain personal and private.
8. People can be divided into two distinct classes: the weak and the strong.
9. Human nature being what it is, there will always be war and conflict.
10. The true American way of life is disappearing so fast that force may be necessary to preserve it.

Scoring the scale: no items are reverse coded.

Appendix D

Other Domains Framed as a Virus or Beast

Cheating

Cheating is a {beast preying on / virus infecting} Addison High School. Five years ago the school was in good shape, with no obvious vulnerabilities. Unfortunately, in the past five years the school's defense systems have weakened, and it has succumbed to a major cheating problem. During the past year, there were more than 400 incidents involving cheating – up by more than 300 incidents from 2009. Indeed, it is no longer merely students falling victim to the temptation to cheat; a recent investigation has revealed that teachers have also been cheating to protect their jobs: it was uncovered that teachers had been changing their students' grades on standardized tests so that the classes could meet national testing criterions. School officials fear that if they do not get the problem under control soon, even more serious issues may start to develop.

Forced-choice options

1. Increase penalties for teachers and students implicated in cheating scandals. (enforcement)
2. Reform the testing and evaluation systems for students and teachers. (reform)

Homelessness

Homelessness is a {beast preying on / virus infecting} the city of Addison. Five years ago Addison was in good shape, with no obvious vulnerabilities. Unfortunately, in the past five years the city's defense systems have weakened, and it has succumbed to a major homelessness problem. In the past five years, the homeless population has risen from 2,000 to over 6,000 and has {assailed / plagued} many different neighborhoods. Authorities suspect that people have been coming to Addison from across the country to be homeless because it is an affluent city with a mild climate. This trend concerns longtime residents of the city, who are increasingly worried about their safety and the safety of the growing homeless community. They note that a large number of the homeless suffer from mental illness and are emotionally and behaviorally unstable. There is a worry that if the city does not do something soon, even more serious problems may start to develop.

Forced-choice options

1. Work to discourage people from coming to Addison to be homeless and offer incentives to homeless people already in Addison to return to their hometown. (enforcement)

2. Invest in mental health facilities, shelters, and jobs programs to integrate the growing homeless population into society. (reform)

Terrorism

Terrorism is a {beast preying on / virus infecting} society. While terrorism has always been a threat, it has raged out of control in recent years. By some estimates, the number of victims of terror attacks has increased from an average of about 1,000 per year in the 1990s to over 5,000 per year in recent years. Countries around the world have found themselves vulnerable in ways that they could not anticipate. There is a worry that if the international community does not regain its strength soon, even more serious problems may start to develop.

Forced-choice options

1. Tighten the borders and increase travel restrictions to make it harder for potential terrorists to enter the United States. (enforcement)
2. Reduce conflict by withdrawing armed forces from Middle East, including Afghanistan and Iraq. (reform)

Appendix E

Other Domains and Frames

Economy

The US economy's {battery is dead / engine has died}, and it is {stalled / broken down} in the middle of the road. In recent years, unemployment has reached record highs and remains greater than 8%, GDP has fallen, and the national debt is over 15 trillion dollars. Increasing frustration with the current economic situation has yielded a number of social movements like the Tea Party and Occupy Wall Street, and it appears that economic issues will dominate the 2012 election cycle.

Follow-up Questions

1. What is the likelihood that you would support another round of large-scale stimulus spending?
2. What is the likelihood that you would support a program that aimed at upgrading the nation's infrastructure and retraining the US workforce?
3. To what extent do you think that US Treasury Secretary Timothy Geithner and Federal Reserve Chairman Ben Bernake know what is wrong with the economy?

Education

Schools are {factories / gardens}, attempting to {mold/nurture} young minds in preparation for the future. Historically, the educational system in the United States has been excellent. It has {assembled/yielded} some of the most innovative and talented leaders and professionals in the world today. Unfortunately, in recent years, few students have been {produced/grown} to meet the standards that we hope to achieve. In a recent assessment, the US had fallen behind countries like Canada, Germany, and Japan. Students are dropping out of high school at an alarming rate (nearly 1.2 million students drop out each year) and 68% of students cannot read at grade level by 6th grade.

Follow-up Questions

1. Which of the following two statements is a better reflection of your view of what a school should strive to achieve?
    a. Schools should strive to prepare students to get a good job.
    b. Schools should strive to develop students into well-rounded adults.

2. In your opinion, how effective are standardized tests at measuring student achievement?

3. Who is most at fault for the breakdown in our schools?
    a. politicians
    b. school administrators
    c. teachers
    d. students

4. To what extent do you think a student's surroundings (his or her peers, extracurricular activities, home life, etc) impact his or her ability to succeed in school?

Internet Privacy

Google is one of the most powerful { remote controls / encyclopedias } that the world has ever seen. With a {click of a button / turn of the page}, people can find information on almost anything they want, including information that some people may want to keep private. In a recent court case, for instance, Mr. P argued that Google's power to seek out personal information had caused him direct harm because a Google search of his name returns a faulty arrest record. While Mr. P. was in fact arrested by the police a few years ago, it was later revealed that a mistake had been made and he was exonerated of any wrong doing. The arrest was erased from his police record; however, it is still available on a Google search of his name. Mr. P believes that this prevented him from getting the job and as a result, has sued, arguing that Google has become {a remote control / encyclopedia} that is too powerful.

Follow-up Questions

1. To what extent do you think Google should be held liable in Mr. P's claim?
2. To what extent do you think people are responsible for the pictures and information they post on the internet, including on social media websites like Facebook, MySpace, and Twitter?