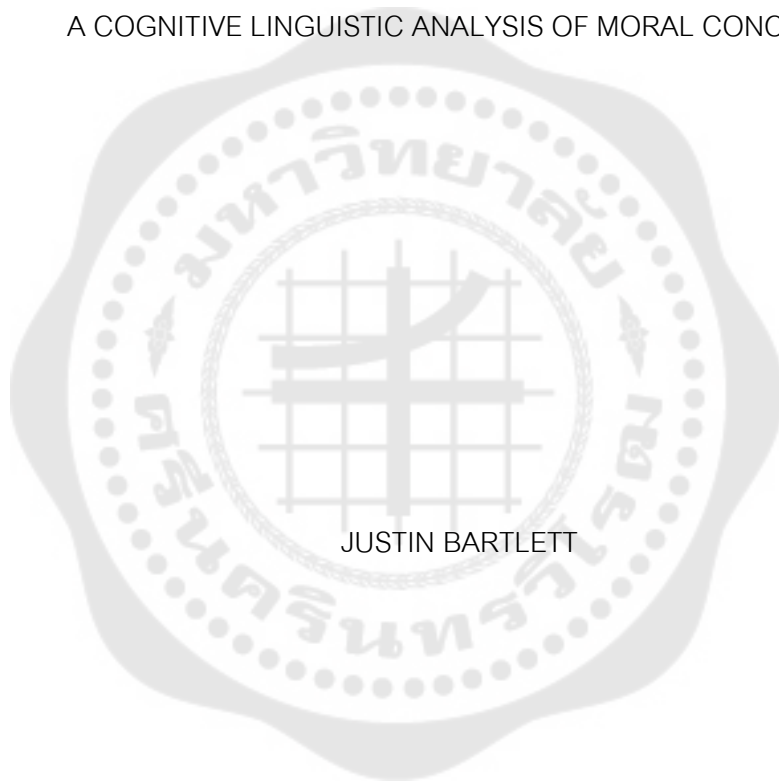




MORALITY, MIND AND MEANING:
A COGNITIVE LINGUISTIC ANALYSIS OF MORAL CONCEPTS



JUSTIN BARTLETT

Graduate School Srinakharinwirot University

2021



ปริญญาบัตรนี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตร
ปรัชญาดุษฎีบัณฑิต สาขาวิชาภาษาและการสื่อสารสากล
คณะมนุษยศาสตร์ มหาวิทยาลัยศรีนครินทรวิโรฒ
ปีการศึกษา 2564
ลิขสิทธิ์ของมหาวิทยาลัยศรีนครินทรวิโรฒ

MORALITY, MIND AND MEANING:
A COGNITIVE LINGUISTIC ANALYSIS OF MORAL CONCEPTS



A Dissertation Submitted in Partial Fulfillment of the Requirements
for the Degree of DOCTOR OF PHILOSOPHY
(Ph.D. (Language and Global Communication))
Faculty of Humanities, Srinakharinwirot University

2021

Copyright of Srinakharinwirot University

THE DISSERTATION TITLED
MORALITY, MIND AND MEANING:
A COGNITIVE LINGUISTIC ANALYSIS OF MORAL CONCEPTS

BY
JUSTIN BARTLETT

HAS BEEN APPROVED BY THE GRADUATE SCHOOL IN PARTIAL FULFILLMENT
OF THE REQUIREMENTS FOR THE DOCTOR OF PHILOSOPHY
IN PH.D. (LANGUAGE AND GLOBAL COMMUNICATION) AT SRINAKHARINWIROT
UNIVERSITY

(Assoc. Prof. Dr. Chatchai Ekpanyaskul, MD.)
Dean of Graduate School

ORAL DEFENSE COMMITTEE

..... Major-advisor Chair
(Assoc. Prof. Dr.Sugunya Ruangjaroon) (Asst. Prof. Dr.Nuntana Wongthai)

..... Committee
(Asst. Prof. Dr.Anchalee Jansem)

..... Committee
(Dr.Sakulrat Worathumrong)

..... Committee
(Dr.Attasith Boonsawasd)

Title	MORALITY, MIND AND MEANING: A COGNITIVE LINGUISTIC ANALYSIS OF MORAL CONCEPTS
Author	JUSTIN BARTLETT
Degree	DOCTOR OF PHILOSOPHY
Academic Year	2021
Thesis Advisor	Associate Professor Dr. Sugunya Ruangjaroon

This work aims to elucidate the *structure* and *form* of moral concepts in a multidisciplinary manner, combining both philosophical and empirically-driven praxes. Loosely situated within the field of Cognitive Linguistics, the guiding assumption made is that the analysis of language can be instrumental in uncovering both conceptual structure and form. The work is comprised of three main studies which explore distinct facets of moral concepts: 1. The first study, in aiming to elucidate conceptual structure, argues that moral concepts RIGHT and WRONG are emotionally embodied and reducible to *emotion-dispositional* concepts. 2. The second study examines the effectiveness of Conceptual Metaphor Theory as a framework through which the structure of the conceptual domain MORALITY and its constituent concepts RIGHT and WRONG can be uncovered. 3. The third study explores the conceptual form of RIGHT and WRONG as binary opposite concepts, endeavouring to understand whether they are mentally represented as mutually exclusive or gradable via analysis of the antonyms *right/wrong*.

Keyword : Moral Concepts, Moral Language, Metaethics, Cognitive Linguistics, Conceptual Metaphor Theory

ACKNOWLEDGEMENTS

This work would not have been possible without the unwavering help and support of my friend, colleague and advisor Sugunya Ruangjaroon. I wish to sincerely thank her for being with me every step of the way.

I would also like to extend my thanks to the management and administration of Srinakharinwirot University who bestowed on me the SWU 70-Year Anniversary Scholarship which enabled the research presented herein; to the PhD committee members for their time and comments; and to the reviewers and editors from various academic journals – most notably *Axiomathes* – who's comments on earlier versions of the arguments presented herein have helped to vastly improve the quality of the work.

Rights and Privileges

The work from various sections of Chapter 4 is reproduced here courtesy of Springer Nature and was first published in: Bartlett, J. J. & Ruangjaroon, S. (2022). A War of Words: Dissecting the Foundational Claims of CMT. *Axiomathes*. DOI: 10.1007/s10516-021-09612-0.

JUSTIN BARTLETT

TABLE OF CONTENTS

	Page
ABSTRACT	D
ACKNOWLEDGEMENTS.....	E
TABLE OF CONTENTS.....	F
LIST OF TABLE.....	K
LIST OF FIGURE.....	L
TERMINOLOGY.....	1
CHAPTER ONE BACKGROUND & INTRODUCTION.....	2
1. Background & Introduction.....	2
1.1 Background and Overview.....	2
1.2 Language and Thought.....	3
1.3 Objectives.....	5
1.4 Scope.....	7
1.5 Overview of the Arguments.....	7
1.5.1 The Emotional Structure of Right & Wrong.....	7
1.5.2 Conceptual Structure & Moral Metaphor.....	8
1.5.3 The Conceptual Form of Right & Wrong.....	9
1.6 Significance.....	10
1.7 Theoretical Framework.....	10
1.7.1 Epistemic Claim: Encyclopaedic Knowledge.....	12
1.7.2 Concepts, Frames and Domains.....	13
1.7.3 A Theory of Mind: Embodied Cognition.....	14

1.7.4 Concepts and Language	16
CHAPTER TWO LITERATURE REVIEW.....	17
2.Literature Review.....	17
2.1 Introduction.....	17
2.2 Constructing Concepts	18
2.3 Concepts as Definitions	19
2.4 Prototype Theory	21
2.5 Exemplar Theory	22
2.6 The Knowledge Theory	24
2.7 Defining Concepts.....	26
CHAPTER THREE METHODOLOGY	27
3.General Methodology	27
3.1 Conceptual and Logical Analysis	28
3.2 Intuition and Introspection.....	29
3.3 Secondary Data.....	30
3.4 The Lexical Approach	30
CHAPTER FOUR ANALYSIS	31
4. Analysis	31
4.1 The Emotional Construction of Moral Concepts:	31
4.2 The Metaphorical Structure of Moral Concepts	32
4.2.1 Conceptual Metaphor Theory & Moral Concepts.....	32
4.2.2 The Foundations of CMT	33
4.2.3 Under Attack	36

4.2.4 The Multiple Source Domain Problem	38
4.2.5 The Moral High Ground	39
4.2.6 Doing the Dirty	42
4.2.7 The Dark Side.....	44
4.2.8 Paying the Price	45
4.2.9 Deplorable Acts	46
4.2.10 The Problem of Divergent Source Domains	50
4.2.11 Building Conceptual Layers	52
4.2.12 The Scope of Metaphor	52
4.1.13 From the Ground Up	54
4.1.14 A One-Way Relationship.....	56
4.1.15 Directionality.....	59
4.2.16 Constituents of Multi-Level Mappings	60
4.2.17 Doing Good.....	61
4.2.18 Summary	68
4.3 The Form of Moral Concepts.....	68
4.3.1 Lexical Relations	70
4.3.2 The Conceptual Nature of Antonymy.....	72
4.3.3 Antonymic Form	74
4.3.4 Gradable Antonyms and Gradable Concepts.....	74
4.3.5 Complementary Antonyms and Mutually Exclusive Concepts	75
4.3.6 Conceptual Representation and Grammar	76
4.3.7 Grammatical Analysis	76

4.3.8 Interpretations	79
4.3.9 Context and Conceptual Form.....	79
4.3.10 Mutually Exclusive Concepts.....	80
4.3.11 Gradable Concepts	83
4.3.12 Grounding Right and Wrong.....	87
4.3.13 Summary	88
CHAPTER FIVE CONCLUSIONS & IMPLICATIONS	89
5. Conclusions and Implications.....	89
5.1 The Emotional Constitution of Moral Concepts.....	90
5.2 Conceptual Metaphor.....	90
5.3 Conceptual Form	91
5.4 Further Insights & Implications.....	92
5.4.1 Kinds of Concept	92
5.4.2 Building A Theory of Conceptual Kinds	93
5.4.3 Entity Concepts and Predicate Concepts	95
5.4.4 Compositionality and Emergence.....	98
5.4.5 Empirical Evidence for Divergent Conceptual Kinds	99
5.4.6 The Flexibility of Words and Concepts	101
5.4.7 Other Conceptual Kinds	103
5.4.8 Summary	106
5.5 In Conclusion.....	107
5.5.1 Final Thoughts.....	109
REFERENCES.....	111

APPENDIX 131

APPENDIX A 132

APPENDIX B 152

VITA 170



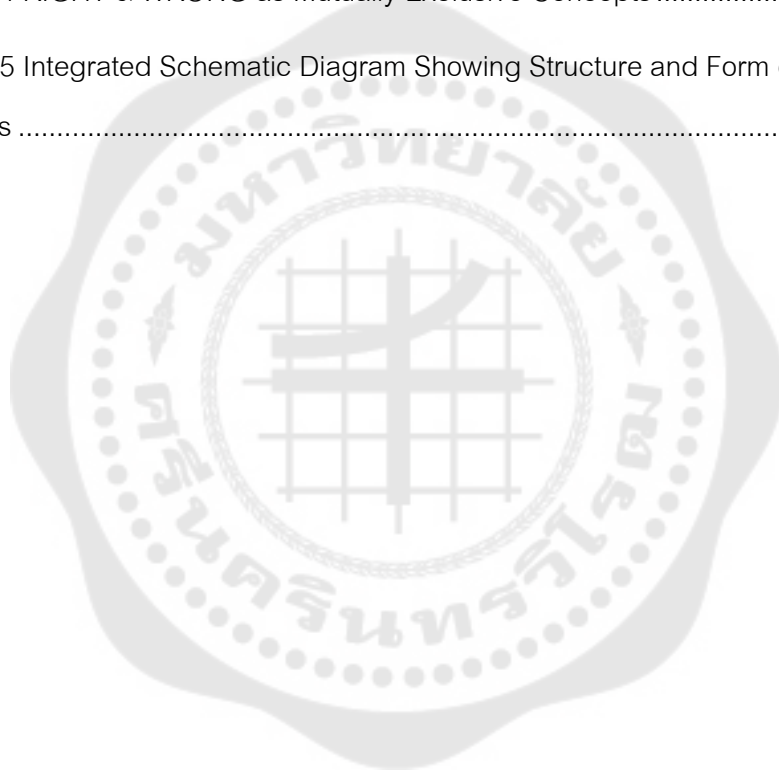
LIST OF TABLE

	Page
Table 1: Multi-layer Mapping of MORALITY (Three Levels)	58
Table 2 Multi-layer Mapping of MORALITY (Six Levels with Positive Concepts).....	63
Table 3 Multi-layer Mapping of MORALITY (Eight Levels with Positive Concepts)	64
Table 4 Multi-layer Mapping of MORALITY (Eight Levels with Negative Concepts)	65



LIST OF FIGURE

	Page
FIGURE 1 Three-level Multi-layer Mapping for MORALITY	60
FIGURE 2 Seven-level Multi-layer Mapping for MORALITY	68
FIGURE 3 <i>RIGHT & WRONG as Gradable Concepts</i>	88
FIGURE 4 RIGHT & WRONG as Mutually Exclusive Concepts	88
FIGURE 5 Integrated Schematic Diagram Showing Structure and Form of Moral Concepts	108



TERMINOLOGY

ABBREVIATIONS

CMT: Conceptual Metaphor Theory

ECMT: Extended Cognitive Metaphor Theory

CL: Cognitive Linguistics

LOC: Lexical Items as Ontologies and Construals

LCCM: Lexical Concepts and Cognitive Models

PS: The Principle of Superordinacy

PR: Principle of Structural Relation

PC: The principle of Constituency

TYPOGRAPHIC CONVENTIONS

Small Capital Letters: Small caps will be used when representing conceptual domains such as MORALITY and concepts within those domains such as RIGHT and WRONG.

Lowercase Italics: Lowercase italics will be used to represent antonym pairs e.g. *hot/cold*, and quotes will be used to refer to linguistic forms such as words, phrases and sentences.

CHAPTER ONE

BACKGROUND & INTRODUCTION

1. Background & Introduction

1.1 Background and Overview

This work attempts to examine and elucidate the structure and form of the moral concepts *RIGHT* and *WRONG*. Understanding moral thought has been a preoccupation of mine for several years and, no doubt, will remain so into the future. It ought to be stated from the outset, however, that my interest is not in trying to find a virtuous path in life, or to determine the ways in which righteous people should behave, although these are noble goals. My aim is, in contrast, metaethical; which is to say that I am concerned with elucidating the way in which humans understand moral concepts, such as *RIGHT* and *WRONG*, *MORAL* and *IMMORAL*, *GOOD* and *EVIL*, or the concept of *MORALITY* itself.

There are many varied questions asked in metaethics including: Why do humans label some actions 'right' and others 'wrong'? Why does morality motivate us? How do people reason when making moral decisions? These questions may seem to the layperson to have simple answers; "It's obvious!", they may exclaim, "My religion and culture teaches me the rules, and I follow them!". They may even go as far as to claim that "religion is what makes people good". However, when faced with such claims, we might still legitimately ask our imaginary interlocutor: What does 'good' mean and who decides its meaning? We might also remind them about those devout believers who frequent the temple or the church, but who nevertheless readily commit immoral acts. Or, conversely, of the unbelievers who never subscribe to any religion, but who steadfastly strive to help others and stand up against evil and corruption in society. Such easy-to-make observations should lead us to carefully reconsider the simplistic views of morality which, unfortunately, are still pervasive in modern society.

In an attempt to tackle the metaethical problem of understanding moral concepts, the forgoing work analyses the *structure* and the *form* of the moral concepts *RIGHT* and *WRONG* through methods of conceptual and linguistic analysis. Accordingly, the following chapters will contain an in-depth exploration into the link between moral

language and moral thought, or more specifically, between moral *words* and moral *concepts*. In light of this goal, it will be instructive to start by appreciating the wonder that is language and its link to thought before moving on with the proposed analysis.

1.2 Language and Thought

The phenomenon of language is often taken for granted, or overlooked, as a mundane and simple feature of everyday life. Indeed, there is evidence for this when one talks of language in friendly conversation; it is often assumed when one says 'I study language' that one means *languages* and not *language* – that is, various foreign languages, or dialects, and not language itself. It is, I believe, commonly asked of linguists: How many languages do you speak? Such a question is testament, I claim, to the common belief that language is simply a matter of knowing the right words and grammar, and that the only interesting thing about it is that there are other exotic languages spoken in distant lands. The study of foreign languages is, of course, an extremely worthy and enlightening endeavour, but it is certainly not the only one, and arguably not the most profound.

Making a similar point, the Austrian philosopher Ludwig Wittgenstein once claimed: '*Die für uns wichtigsten Aspekte der Dinge sind durch ihre Einfachheit und Alltäglichkeit verborgen*' (The aspects of things that are most important for us are hidden because of their simplicity and familiarity) (Wittgenstein, 2009) . Although Wittgenstein means this as a comment about all worldly experience, language is clearly on his mind, and, in regards to language, this claim is extremely poignant, as what seems, on the face of it, the most simple and familiar of things, is in fact an extremely complex and enigmatic phenomenon.

To give an example, as you are reading this sentence, take a moment to appreciate the incredible achievement it is to understand it (assuming you do understand it). The intricate shapes (letters), the complex arrangements in which they appear (words) and the even more convoluted structures which they are used to build (sentences) all work together to encode information, and you are able to decode this information and understand it almost instantaneously. Remarkably, the data encoded in the sentences on

this page emanates from the most private of spaces: my mind. Language, therefore, is capable of capturing and transmitting one's thoughts, encoding them and storing them so that, at a later date, others can read what has been written, and translate those words into their own thoughts.

Viewed in this way, language can be understood as connecting one mind with another, and depending on the content of the information transmitted, even changing one's thoughts, enriching one's knowledge, altering one's intentions or invoking one's emotions. So, as can be seen from this brief illustration, language, both written and spoken, is a device for connecting people's minds and, what's more, it can do so in the present moment, or over vast periods of time, so that we are able to understand not only the thoughts and opinions of our contemporaries, but also those of our ancestors; Plato, Aristotle, Descartes.

The relationship between language and thought has concerned philosophers for millennia, at least since Aristotle who noted in *On Interpretation* that 'Spoken words are the symbols of mental experience' (Aristotle, 1963). However, it is arguably due to the, *linguistic turn* of analytic philosophy during the 20th Century, that the focus on language and its relation to thought became an area of intense study taken up not only by philosophers, but by, psychologists, neuroscientists and linguists alike. Today, the study of language and how it relates to the mind is a truly interdisciplinary endeavor making use of both conceptual analyses undertaken from the philosopher's proverbial armchair as well as empirical evidence garnered from the neuroscientist's lab (Castroviejo et al., 2018). The foregoing study adds to this body of work in the same interdisciplinary manner by undertaking the analysis of moral language as a means of answering psychological and philosophical questions.

Given the above, we can make a preliminary broad assumption: that if language is a tool for communicating thoughts and concepts, then the study of language will be instrumental, to some extent, in understanding the mind. It must be stressed, however, that I don't mean merely the content of the mind – this is too trivial an assertion – but I allude instead to the form and structure of our concepts and the intricate

machinations of human cognitive processes. The present work, then, in general terms, aims to penetrate the mundane façade with which language presents us and explore the deep structuring and functioning of the human conceptualization processes at its foundations, through the analysis of language itself. And, specifically, to understand moral concepts through analysis of moral language.

Before doing this, it will first be necessary to set the stage by outlining the main objectives of the foregoing work, before moving on to explicate in detail the specific claims put forward in this study, as I will now proceed to do.

1.3 Objectives

As a whole, this work forms a multidisciplinary and multifaceted analysis of moral language which aims at exposing the conceptual nature of the moral concepts *RIGHT* and *WRONG*. In doing this, the study will cover the divergent linguistic phenomena of *predication*, *metaphor* and *binary opposition* (or antonymy) in moral language. The arguments presented here function as stand-alone, but complementary, analyses; each supporting the other in such a way as to render the complete work a defence of a single thesis, which can be stated as follows:

Thesis Statement: The moral concepts *RIGHT* and *WRONG* are abstract and emotionally grounded; such that being able to experience moral emotions is a necessary condition for grasping moral concepts. The conceptual form of *RIGHT* and *WRONG* is flexible due to their abstract nature and can, therefore, be mentally represented in divergent forms as either mutually exclusive or gradable opposites.

The specific aims of the work, in supporting this thesis are as follows:

- i To elucidate the emotional structure of the moral concepts *RIGHT* and *WRONG*.
- ii To understand whether Conceptual Metaphor Theory is an effective framework through which to study the structure of moral concepts.
- iii To show the divergent conceptual forms which *RIGHT* and *WRONG*, as binary opposites, assume in the mind.

These aims are related in that they all concern the nature of moral concepts; however, they refer to two divergent conceptual phenomena viz., *conceptual structure* (i, ii) and *conceptual form* (iii). Analysing the *structure* of a concept is akin to identifying the various necessary *parts which constitute* the concept, and the *form* of a concept is analogous to finding the *shape* in which a fully-constituted concept takes. In approaching these aims, I will present three separate studies (§ 4.1, 4.2 & 4.3) which analyse these phenomena through varied linguistic analyses. The studies are motivated by the following foundational assumption, which guides the analyses presented herein:

Foundational Assumption: Given that language is used to express conceptual material, the form and structure of moral concepts will be reflected, to some extent, in the form and structure of moral language.

This assumption is in line with the general enterprise of Cognitive Linguistics, within which the foregoing work is loosely situated. Cognitive Linguistics is not a single theory or framework, but, rather, a branch of cognitive science which understands language to be a product of more general cognitive functions – as opposed to emanating from a specific language faculty – and, therefore, sees the study of language as a legitimate means by which to study said functions. More succinctly, Cognitive Linguistics assents to the following broad claim:

[L]anguage is not an autonomous cognitive faculty. The basic corollaries of this hypothesis are that the representation of linguistic knowledge is essentially the same as the representation of other conceptual structures, and that the processes in which that knowledge is used are not fundamentally different from cognitive abilities that human beings use outside the domain of language (Croft & Cruse, 2004).

The above outlines the general outlook and foundational assertions made in the field of Cognitive Linguistics, from the perspective of which, myriad linguistic phenomena are studied. The Cognitive approach which started to develop in the works of Charles Fillmore, George Lakoff and Mark Johnson, Ron Langacker, and Leonard Talmy (Geeraerts & Cuyckens, 2007), represents a marked shift from traditional views in the philosophy of language and linguistics, from which language was often seen as a

cognitively autonomous, or even externally situated, phenomenon, to one where it is viewed as a function of the mind's general cognitive processes; as dependent upon, and structured by, more fundamental conceptual and perceptual mechanisms.

In the past forty years, research in the field of Cognitive Linguistics and cognitive psychology has become increasingly intertwined, especially in relation to the embodiment hypothesis (Rohrer, 2010), with theoretical work from linguists being used in conjunction with experimental results in cognitive science. Such a perspective, I hope to show, provides us with a lens through which we can examine the conceptual nature of moral concepts *RIGHT* and *WRONG* via analysis of moral language, as, from such a perspective, the structure and functioning of moral language is seen as reflective of the structure and functioning of moral thinking.

1.4 Scope

In having as its goal the explication of moral concepts, and by doing so through the lens of Cognitive Linguistics, this work will inevitably involve the exploration of moral language. As mentioned above, the divergent linguistic phenomena of predication, metaphor and binary opposition will be examined in order to reach an understanding of the conceptual structuring and form of *RIGHT* and *WRONG*. The following paragraphs will outline the specific scope of these studies and provide an overview of the three main arguments to be presented in Chapter 4.

1.5 Overview of the Arguments

The proceeding work is comprised of three studies which all have as their aim the examination and elucidation of the moral concepts *RIGHT* and *WRONG*. The first two studies of *predication* and *metaphor* concern the structure of *RIGHT* and *WRONG* and the third investigation of *binary opposition* (antonymy) explores their form. The following paragraphs provide an outline of each study.

1.5.1 The Emotional Structure of Right & Wrong

The first study, (§ 4.1) presents a theoretical argument, supported by secondary empirical findings, which posits that moral concepts are emotionally structured. In supporting this claim, I survey the extant psychological and neurological evidence that suggests strong influence of emotional structures of the brain in moral

thought. I then present an analogical argument which draws a parallel between moral properties and concepts and what I refer to as *emotion-dispositional* properties and concepts which are identified by adjectives such as '*annoying*', '*frightening*,' or '*disgusting*' and appear to be uncontroversial and common examples of mistakenly predicating subjective emotions to mind-independent states of affairs; a conceptual and linguistic mistake which I have dubbed *attribution error* (Bartlett, 2020). I develop a reading of attribution error which allows us to explain why it happens by alluding to *figure and ground construal* and showing that emotional stimuli are foregrounded due to their cognitive salience, thereby construing grammatical form in such a way as to commit the attribution error. I then show that moral language functions in an analogous way to emotion-dispositional language, thus giving reason to believe that moral actions are emotional in nature. In offering further support, I draw attention to the fact that moral judgements are often explicitly expressed in emotion-dispositional terms. It is ultimately concluded, based on this analysis, along with the convergent experimental data, that moral properties and concepts can be reduced to emotion-dispositional properties and concepts.

1.5.2 Conceptual Structure & Moral Metaphor

The second study (§ 4.2) aims at assessing the effectiveness of the influential framework Conceptual Metaphor Theory (CMT), in further elucidating the structure of moral concepts. I will give a critical assessment of CMT both in itself and in light of the conclusions drawn in the first study. To this aim, I pose the question:

Is the framework of Conceptual Metaphor Theory an effective means for discovering the conceptual structure of moral concepts?

In answering this question, I give a CMT-based analysis of the conceptual domain *MORALITY* and its constituent concepts *RIGHT* and *WRONG* in order to assess whether the framework can exert any leverage on the question of *MORALITY*'s conceptual structure. This examination exposes a series of divergent source domains apparently structuring the target. I then explicate how this is potentially problematic for CMT, as it raises the question of whether all source domains are necessary for the structure of the target, or not.

Furthermore, *EMOTION* does not appear as a salient source domain in any previous CMT-based analyses which, in light of the overwhelming evidence that implicates emotional regions of the brain in moral thought, raises doubts about the efficacy of the framework. I sketch a hierarchical approach that tries to coherently integrate the divergent source domains assumed to be necessary to the understanding of *MORALITY*, but I ultimately conclude that this supplementation of CMT does not help the theory to overcome deeper problem unearthed here.

1.5.3 The Conceptual Form of Right & Wrong

The third study (§ 4.3), which examines the conceptual *form* (as opposed to the *structure*) of *RIGHT* and *WRONG*, analyzes the use of the antonyms *right/wrong* in moral discourse in order to draw conclusions about the form in which their binary opposition is mentally represented. The specific question under investigation is:

What form of binary opposition – mutually exclusive, or gradable – do the moral concepts *RIGHT* and *WRONG* take?

The foundational assumption guiding this investigation claims that the relation between conceptual representation and language is such that, the form in which concepts are represented in speakers' minds is encoded by lexical items and determines, therefore, their lexical form, along with the grammatical form of the utterances in which they appear. Based on this assumption, I explore how moral antonyms *right/wrong* are used in moral discourse in order to discern their lexical status as either complementary or gradable contrary antonyms and, therefore, the form of the related concepts *RIGHT* and *WRONG*. I hypothesize that if an agent conceptually represents *RIGHT* and *WRONG* as mutually exclusive concepts, then they will use *right/wrong* as complementary antonyms and, conversely, when represented as gradable, they will use *right/wrong* as gradable contrary antonyms, which will be evident from the grammar of their utterances. I conclude that the form of representation which moral concepts exhibit is, not fixed, but flexible; that in some instances agents mentally represent *RIGHT* and *WRONG* as mutually exclusive, and in others as gradable.

I go on to demonstrate how the form of conceptual representation that RIGHT and WRONG take – as either mutually exclusive or gradable – is contingent upon the states of affairs to which they are applied. In utilitarian-style moral judgements, based on maximizing or minimizing outcomes, for instance, RIGHT and WRONG are commonly understood as gradable in nature. Deontological judgements, in contrast, with their foundations in categorical ‘black and white’ moral rules, tend to force RIGHT and WRONG into a mutually exclusive form. I present linguistic evidence in support of a conceptual link between the form of RIGHT and WRONG and the physical form of the states of affairs to which they are applied.

1.6 Significance

The current work is deemed to be significant in presenting a multidisciplinary study which hopes to integrate the fields of Cognitive Linguistics and analytic metaethics. Although linguistic analysis has been traditionally used in metaethics, frameworks from Cognitive Linguistics, to my knowledge, have not previously been applied to the philosophical study of moral concepts. The work, therefore, represents a novel multidisciplinary investigation into the structure of moral concepts and what this tells us about moral thought and action.

1.7 Theoretical Framework

As stated above, the current work is multidisciplinary in nature, but takes a broadly ‘Cognitive Linguistic’ approach to the study of moral concepts through the examination of divergent linguistic phenomena. I will now proceed to explicate this theoretical framework in more detail.

The Cognitive Linguistic approach grew out of the work of several theorists including Charles Fillmore (1968, 1976), Leonard Talmy (1975), George Lakoff and Mark Johnson (1980) and Ronald Langacker (1987). This approach has been taken further in the work of others (Cruse & Toggia, 1996; Evans, 2009b; Goldberg; Adele E., 1995; Steven. Jones et al., 2012; Kövecses, 2000) to mention just a few. The Cognitive Linguistics enterprise encompasses only very broad foundational assumptions, or commitments, to the study of language and cognition and does not constitute a fully-fledged theory of mind and language, as Vyvyan Evans explains:

It is important to note that cognitive linguistics is best described as an “enterprise” precisely because it does not constitute a single closely articulated theory. Rather, it represents an approach that has a number of core commitments and guiding principles, which have led to a diverse range of complementary, overlapping (and sometimes competing) theories. The cognitive linguistics enterprise is characterized by two fundamental commitments: the Generalization Commitment and the Cognitive Commitment (Evans, 2009b)

The ‘Generalization Commitment’ can be understood as the assumption that explanations of certain linguistic phenomena are generalizable to other phenomena in language. This means that an explanation of semantics, for example, should be generalizable and applicable in the study of grammar, syntax and other areas of linguistics. The Generalization Commitment represents a marked difference in perspective from that of traditional schools of thought within linguistics which see the different aspects of language as qualitatively distinct. The second of the two foundational assumptions, the ‘Cognitive Commitment’, states that the characterization of language should be in line with our best understanding of the mind, and cognitive functioning, as understood in other scientific disciplines such as psychology, neuroscience and philosophy of mind (Lakoff, 1990).

Although the specific approaches taken, and linguistic phenomena studied, by the authors working within the field vary, they all assent to the abovementioned assumptions. Cognitive linguists, therefore, assume that there is an intrinsic link between cognition and language, such that language does not rely on a specialized or autonomous cognitive process, but instead emerges from more general cognitive functioning. This assertion, as I have noted elsewhere, is “grounded in the fact that no specific structure in the brain has been found to be responsible for language (Anderson & Lightfoot, 2002). Cognitive linguists argue that conceptualization underlies not *only* semantic representation but that syntax, morphology, and phonology are also conceptually grounded” (Bartlett, 2020). The assumption that general cognitive processes such as conceptualization underpin language use, grounds the assumption that the analysis of

language can be instrumental in uncovering and understanding these cerebral processes. It is

In addition to the above commitments, the current work makes several further assumptions about mind and language. These foundations are important theoretical pre-requisites which depart, sometimes radically, from classical theories of mind and language. It will, therefore, be necessary to elucidate these assumptions before moving on.

1.7.1 Epistemic Claim: Encyclopaedic Knowledge

The current work makes the epistemic claim that concepts are situated within, and are understood against the background of, complex and coherent knowledge structures which are 'encyclopaedic' in nature. The so-called 'encyclopaedic' view of knowledge sees conceptual information as follows:

The structured body of non-linguistic knowledge to which a linguistic unit such as a word potentially provides access. Encyclopaedic knowledge is modelled in terms of a number of constructs including the domain [...], the cognitive model and the idealised cognitive model (Evans, 2007).

Understanding conceptual knowledge as being encyclopaedic in nature is in opposition to the traditional view of knowledge as being dictionary-like. That is to say, our knowledge of a particular concept or domain is understood as part of a complex of integrated knowledge about the world – much like an encyclopaedia – and that it cannot be defined or understood apart from this rich integrated system of background knowledge.

In contrast, classical views see individual concepts as discrete conceptual symbols which can be totally and clearly defined without making reference to any further background knowledge (I will survey theories of concepts in Chapter 3). Hence, to put it poetically, the fabric of our concepts is our knowledge of the world and this knowledge is a rich and detailed tapestry which is the mental representation of the totality of one's stored conceptual information. This tapestry can be imagined as being

one long landscape where all concepts are linked together in a coherent scene, as opposed to understanding concepts as being separate unconnected entities.

1.7.2 Concepts, Frames and Domains

With a basic characterization of encyclopedic knowledge now in-hand, we are able to reach an adequate understanding of the theoretical constructs *concepts*, *frames*, *domains* and *domain matrices*. I will explicate these fundamental ideas, as they are to be understood in the current work, in turn:

Concepts: Mental representations which can vary in form and degree of abstractness and represent the smallest units of encyclopedic knowledge and highlight individual pieces of conceptual knowledge, such as *BALL*, for instance, from a complex background of encyclopedic knowledge.

Frames: Complex mental representations which encode a scene or section of background knowledge (as opposed to an individual part of the scene). Frames contain many constituent concepts which are understood to represent the entities and relations between the constituents of a particular context such as *FOOTBALL* which includes the concepts *BALL*, *PLAYER*, *GOAL*, *OFF-SIDE* etc.

Domains: More complex mental representations which refer to a whole domain of experience. Domains are large knowledge structures, such as *PHYSICAL ACTIVITY*, which contain frames and their constituent concepts such as $FOOTBALL = \{BALL, PLAYER, GOAL\dots\}$ and $JUDO = \{THROW, ARMLOCK, DEFEND\dots\}$.

Domain matrix: the largest unit of mental representation consisting of domains, frames and concepts. A domain matrix is assumed to link multiple conceptual domains with each other; such that the understanding of one domain presupposes others. For instance, *PHYSICAL ACTIVITY* presupposes *GOAL-DIRECTED BEHAVIOR* and *PHYSICAL OBJECTS*, thus the domains are linked in a matrix.

The conceptual structures explicated above all relate to mental representations which can be activated in one's mind upon encountering certain stimuli. Understood in this way, stimuli serve to highlight particular sections of knowledge in our encyclopedic understanding of the world. For example, the word '*ball*' is a linguistic

stimulus which activates the concept BALL, as is the word *'football'* which activates the frame FOOTBALL, along with its constituent concepts BALL, PLAYER, GOAL etc.

1.7.3 A Theory of Mind: Embodied Cognition

Along with the encyclopedic view of knowledge, the current work embraces a *weakly embodied* theory of mind. The view that cognition is 'weakly embodied' needs to be understood against the background of a *strong* embodiment and this, in turn, in contrast to classical theories of mind.

Classical theories of cognition, such as the Representational Theory of Mind (RTM) and the Computational Theory of Mind (CTM), view the mind as being an abstract symbol manipulation system that computes concepts in an amodal (non-sensory-specific) fashion and which does not extend beyond the confines of the brain. In contrast, (and in reaction) to this theory, many contemporary cognitive scientists have proposed and supported a strongly 'embodied' view of cognition which sees cognition as a process that is multi-modal and perceptually grounded; such that sensory-motor systems are necessary for the comprehension of various concepts. As such, the strongly embodied view sees cognition as an activity which is distributed across the whole of the body, and not as an activity that happens exclusively in the brain (or in certain regions of the brain). There is empirical evidence which appears to support the embodiment hypothesis (Barsalou, 2007; Gallese et al., 1996; Gallese & Lakoff, 2005), and this is the predominant view among cognitive linguists. As Vyvyan Evans explains, embodied cognition is:

One of the guiding principles of cognitive semantics and at the heart of much research in cognitive linguistics. This thesis holds that the human mind and conceptual organisation are a function of the way in which our species-specific bodies interact with the environment we inhabit. In other words, the nature of concepts and the way they are structured and organised is constrained by the nature of our embodied experience. (Evans, 2007)

It is important to note that such a characterization also includes what is sometimes referred to as 'situated' or 'embedded' cognition – the thesis that our cognition is dependent to some extent on our experience of being situated within, or interacting

with, the physical world. The theory of *embedded cognition* emphasizes the importance of being environmentally situated in our cognitive functioning (Hutchins, 1995; Suchman, 1987). Thus, the theory described by Evans above (which is commonly held by cognitive linguists) is a view which can be called *embodied embedded cognition* (EEC) (Clark, 1999). There is, indeed, empirical support for the EEC thesis which includes observations that humans are able to perform certain cognitive tasks more efficiently when using hand gestures, suggesting that certain cognitive functions can be 'sub-contracted out' to other parts of the body (Donald, 1991) and also that gesturing helps to facilitate the comprehension and processing of language (McNeil, 1992).

Another related non-classical theory of mind is the *extended mind thesis* which claims that the mind extends beyond the body and into the physical environment; such that physical entities in the world can actually take on the cognitive functions of an individual, thereby reducing the processing load (Clark & Chalmers, 1998). The theories of embedded and extended cognition are, in principle, compatible with embodied cognition, however, strong embodiment does not necessarily entail either embedded or extended cognition.

Against this background, we can understand the 'weak embodiment' theory which the current work adopts. A weakly embodied theory of mind holds that cognition can be both embodied and multimodal as well as disembodied and amodal. Such 'hybrid' approaches are starting to develop as a way to account for experimental evidence that threatens the soundness of both classical views of cognition and strongly embodied views (Dove, 2011, 2015; Goldinger et al., 2016). Thus, the embodied view assumed here is a more nuanced perspective and should not be understood as endorsing the strong view that *all* cognition is embodied, but claims, instead, that some concepts may be embodied and some may not. This amounts to claiming that the mind is able to perform both embodied modal conceptualization as well as amodal abstract symbol manipulation, thus emphasizing human cognitive flexibility. Moreover, weak embodiment does not entail either embedded or extended cognition.

1.7.4 Concepts and Language

There is undoubtedly a close relationship between thought and language which many have tried to understand. Some have claimed that there is a one-to-one relationship between language and thought such that words and concepts are interchangeable (Humphreys et al., 1999; Vigliocco & Vinson, 2012). Others have challenged this view showing that there are many concepts which do not become lexicalized, meaning that we might have more concepts than words (Murphy, 2002). Whether there is a one-to-one mapping or a many-to-one mapping of concepts with lexical items, is still an open question which need not occupy us here, however.

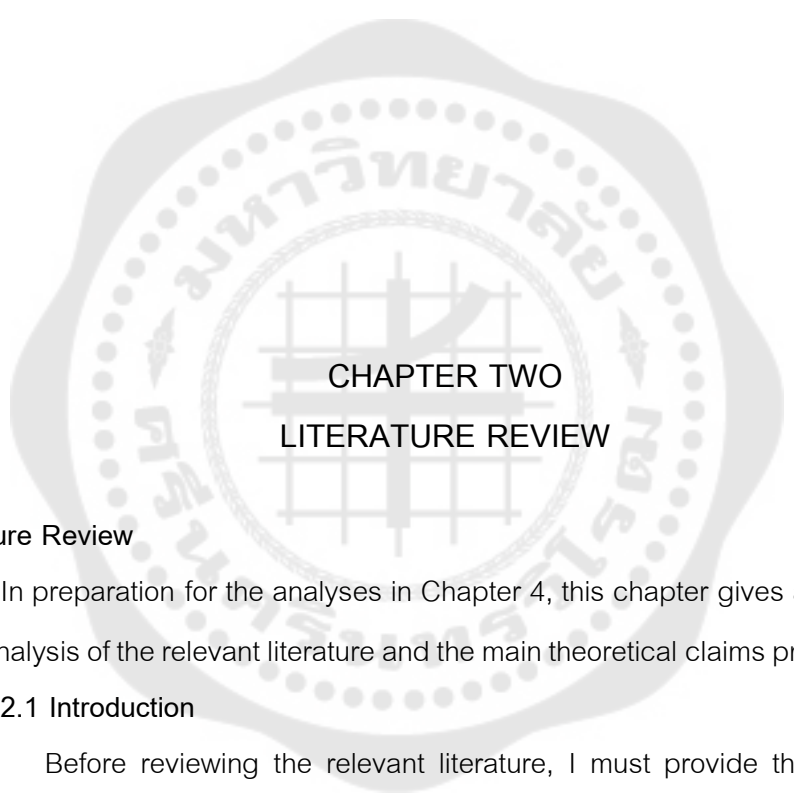
It is uncontroversial that language is a means of expressing conceptual content, but that's about all that is uncontested about the link between language and thought. Despite the huge amount of attention paid to language and its connection to the mind by theorists across myriad fields, the precise nature of this connection has proven to be notoriously tricky to describe and is still fiercely debated. This is partly, I conjecture, due to the fact that concepts themselves have proven so difficult to understand.

In the foregoing work, I make the assumption that lexical items serve to encode and express conceptual data, as mentioned above. Concepts and words, however, are not understood as being one and the same. That is to say, this claim should not be seen as the ideational metasemantic assumption that the meanings of words are defined by the concepts to which they relate, but, rather, the more modest claim that, in serving to express conceptual material, linguistic form and structure must be apt to encode conceptual form and structure.

If this assumption is true, the analysis and understanding of linguistic functioning and form will provide insight into conceptual structure and form, and it is this assumption which motivates the current work. The most compelling reason for taking such a view is that a great many conceptual phenomena – which I will proceed to survey in Chapter 2 – are evident in language. Categorization effects and hierarchical structure (Federmeier & Kutas, 1999), typicality effects and basic-level naming effects (Lin et al.,

1997; Rosch et al., 1976; Wisniewski & Murphy, 1989) have all been observed in language use and comprehension.

With the main aims and theoretical assumptions established, I will proceed, in the following chapter, to survey of the relevant literature in preparation for the main analyses to come in Chapter 4.



CHAPTER TWO LITERATURE REVIEW

2.Literature Review

In preparation for the analyses in Chapter 4, this chapter gives an overview and critical analysis of the relevant literature and the main theoretical claims presented therein.

2.1 Introduction

Before reviewing the relevant literature, I must provide the reader with a disclaimer. Due to this work being multidisciplinary in nature, and undertaking distinct, but related, analyses which concern divergent conceptual and linguistic phenomena, it is necessary to refer to a wide range of background literature which spans the fields of psychology, philosophy, linguistics and neuroscience. Given this, having a dedicated literature review section is, perhaps, not the best format to follow, but in an effort to address both structural formalities and also to cover all the relevant literature, I will introduce only the very foundational works which have been crucial to the historical development of the understanding of concepts in this dedicated chapter, and will add to

this by including and critiquing the more specific studies and secondary data relevant to each argument alongside the individual analyses in Chapter 4.

2.2 Constructing Concepts

It must be said that a deep and committed dive into the literature on concepts leaves one quickly struggling for breath and wondering which way is up. That is to say, it is an extremely difficult and controversial question as to what a concept is, and even if there are such things at all. Given this, it will be impossible, with limited space and scope, to cover all the relevant theories. I will, then, not even attempt such a feat and will instead try to give a brief, but theoretically adequate, overview of the main ideas so as to clear the ground ready for the analyses presented in Chapter 4.

Concepts¹ have been the focus of much work in philosophy, psychology and Cognitive Linguistics, and there are many divergent views which try to define what they are (Carey, 2009; Dennett, 1987; Gibson, 1966; Hampton & Moss, 2003; G. Murphy, 2002; Vigliocco et al., 2006; Vigliocco & Vinson, 2012). The psychologist Gregory Murphy said that 'Concepts are the glue that holds our mental world together' (Murphy, 2002), and when we consider the importance of concepts, as Murphy invites us to do, it is not surprising that the literature on conceptualization is immense and spans many fields (Margolis & Laurence, 1999, 2015; G. Murphy, 2002). Indeed, the moniker 'cognitive science' is now routinely applied to the study of the human mind and its processes undertaken from divergent disciplines and methodological perspectives, all of which have the same goal: to understand the human mind. 'Cognitive scientists' trained in diverse fields are working to reach an understanding of what concepts are and how they function (Bermúdez, 2014).

Before explaining the content and functioning of specific concepts we need to start by addressing the ontological status of concepts themselves. This question, put in its simplest form is: How are concepts built or structured? Many theories have tried to provide an answer to this question and I will proceed to give an overview of the most

¹ The concept literature is vast, see (Margolis & Laurence, 1999, 2015) for an overview of the field.

relevant ones, their related claims and theoretical weaknesses in the following paragraphs.

As mentioned in Chapter 1, the current work adopts an embodied view of mind – which understands cognition to be grounded in the physiological structures of the whole organism, as opposed to it being a disembodied abstract symbol manipulation process which happens exclusively in self-contained areas of the brain – along with what can be broadly called an ‘encyclopaedic knowledge view’ of concepts. Such claims, however, cannot be adequately understood or assessed in isolation, and should not simply be assumed without argument. I will therefore proceed to introduce the contending theories of concepts which have led to the development of the aforementioned views and this will act as both a discussion of the literature on concepts and also as an argument in support of the foundational assumptions made here.

2.3 Concepts as Definitions

In the early psychological literature, the approach to describing concepts can be called the ‘definitional’ approach. This early approach sought to describe concepts as definitions; such that a concept could be precisely defined by sufficient and necessary conditions. A certain set of sufficient and necessary conditions, it was assumed, could be used to assess whether a particular object fell into a certain category and that the specific criteria for inclusion defined the concept. As an example, to be a member of the category `COMPUTER` an object needs to meet all the relevant criteria e.g. it is an electronic device, it runs software, it has a monitor, it has a hard drive and so on. The full set of criteria, the definitional view claims, perfectly describes a computer. It is held, therefore, on this view, that one’s understanding of the criteria for identifying an object as a computer is what it means to understand or possess the concept `COMPUTER`. Hence, from the definitional perspective, it is one’s knowing the sufficient and necessary conditions for category inclusion which determines their understanding of a given concept. This view can be seen, either implicitly or explicitly, in the majority of the work on concepts from the early to mid-Twentieth Century². Clark Hull’s (1920) work is an example of an early definitional

² See Murphy (2002) for an excellent overview of these theories.

view from psychology. Hull's view assumes, briefly, that for every example of a concept, there is a necessary element unique to it which makes a concept distinct from others. This suggestion was rejected by Kenneth Smoke who also assented to a definitional view of concepts, but did not accept the existence of a single essential element unique to a concept, which Hull had proposed previously. Instead, Smoke claimed that a concept is defined by the full range of elements which are necessary and sufficient for the understanding of a concept (Smoke, 1932). Hence, although Smoke advocated a different version, it was still a definitional theory of concepts. A further example of such a view can be found in the work of Inhelder and Piaget, who also assumed a definitional stance towards concepts which saw them as being defined logically and precisely (Inhelder & Piaget, 1964). Such views were once so abundant and readily assumed in the psychological literature that this theory has become known as 'the classical view' of concepts.

However popular the definitional view of concepts once was, it has all but died in contemporary theory. There are many reasons why the definitional view provides an inadequate model of concepts, but one of the most important and strongest arguments against such a view comes from Ludwig Wittgenstein (although this problem had been known since antiquity and is classically formulated in the Sorites Paradox (Oms & Zardini, 2019). In the *Philosophical Investigations* (1953/2009), Wittgenstein challenges the reader to define the concept *GAME* in terms of necessary and sufficient conditions which, at first, seems like a relatively easy task, but proves to be extremely difficult, if not impossible. Indeed, if you try to define most concepts in this manner, you will soon see that reaching a precise definition which includes all things that, for example, you call a game but excludes all things which you do not conceive of as games, cannot be done. Upon acknowledging this problem, Wittgenstein concludes that concepts (and therefore word meanings) cannot be defined in such a way; that we recognize and categorize things by way of seeing 'family resemblances', or similarities, and not by clearly defined criteria. It could be, therefore, that we only recognize approximate similarities between typical elements of all members in a particular category, and based on this rough similarity we

are able to categorise things. It is this philosophical argument, along with experimental evidence which eventually led to the demise of the definitional theory of concepts. Indeed, the empirical data is heavily stacked against the theory. One of the important discoveries made by psychologists which shows the definitional theory to be wrong was that of so-called *typicality effects*. An example is seen in the work of Rips, Shoben, and Smith (1973) who showed that, when completing categorization tasks, participants found it easier to assign things if they were more 'typical' of that category. For example, people are much quicker at judging that a robin is a bird than they are with a chicken, which is a less typical bird. These typicality effects were also famously shown in the work of Rosch and Mervis (Rosch & Mervis, 1975). With the theoretical strain put on the classical view by Wittgenstein, along with the empirical evidence showing that typicality judgments – and not precise criteria – are what allow people to categorize things and form concepts, the definitional theory fell out of favour and theorists started to adopt a new theory of concept formation which was based on the evidence of typicality effects. This became known as the prototype theory.

2.4 Prototype Theory

The prototype theory is the view that categorization and concept formation are based on how typical an example is of a particular class of entities. Categorization, then, is seen as a comparison process in which the similarity of the constituents of a category is computed to form a 'prototype' concept which acts somewhat like a template. Things which have more features similar to the norm and, therefore, closely resemble the prototype, are more easily understood as belonging to the category in question (Rosch et al., 1976; Rosch & Mervis, 1975; Wattenmaker et al., 1986).

The prototype theory is preferable to the definitional view as it, of course, avoids the problems posed by strict definitions and explains why precise definitions are so difficult to construct. It also explains the typicality effects observed in the experiments mentioned above, as it was designed to do. This view also has its limitations, however. The prototype theory appears to explain how we apply concepts when making quick automatic judgements, but has trouble accounting for more reflective judgements. Some

judgements seem to show that humans often categorize things in a manner that suggests typicality is not the only criterion used in categorization tasks (Gelman, 2007). To illustrate, I will give an example from personal experience. When I was a child, I remember my friend's mother baking a batch of cupcakes for us which were coloured and moulded into the shape of a dog's droppings. Despite her ingenious and skilled crafting of the treats, which left them resembling an extremely life-like and very typical specimen of dog poop, we nevertheless did not conceive them as such; we understood what they were and categorized them as being, cupcakes, despite them being extremely atypical examples. Hence, it appears in such judgements, we are not only relying on typicality in categorization, and prototype theory has a difficult time explaining this.

Another problem for the prototype theory is that it cannot seem to explain the compositionality of concepts. In many cases, when combining concepts to form a compound or complex concept, the resulting concept has a different typicality structure to those of its components (Osherson & Smith, 1981). To give a famous example (Fodor & Lepore, 1996), a typical PET FISH has different prototype form to those of typical FISH or typical PETS. A typical PET is most likely to resemble a dog or cat, and a typical FISH is medium-sized and grey, whereas a typical PET FISH is likely to be imagined as small and colourful – so, nothing like a typical PET or a typical FISH. Thus, compound concepts often have emergent typicality profiles which cannot be predicted from the typical structures of their parts and this also poses a threat to the prototype theory.

2.5 Exemplar Theory

Another competing view of concepts is called exemplar theory (Medin & Schaffer, 1978). Originally called a 'context theory', exemplar theory explains categorization and conceptual formation as being a process which involves agents accessing and using information stored in memory to categorize a particular entity or stimulus. Medin and Schaffer explain their model as follows:

The general idea of the context model is that classification judgments are based on the retrieval of stored exemplar information. Specifically, we assume that a probe stimulus functions as a retrieval cue to access

information stored with stimuli similar to the probe. This mechanism is, in a sense, a device for reasoning by analogy inasmuch as classification of new stimuli is based on stored information concerning old exemplars.

(Medin & Schaffer, 1978)

On this view, it is one's stored memories gathered through experience, as opposed to a list of necessary and sufficient conditions (referential theory), or typical features (prototype theory) which allows us to recognize things and build concepts. Thus, my concept *COMPUTER*, for instance, is made up of memories of all the computers I have encountered in my life and upon seeing a new computer, which might be a novel design, I am able to recognize it by comparing it to my previous memories of other computers and, based on how well the new stimulus matches the *COMPUTER* exemplar, apply my concept *COMPUTER* in order to classify it. Murphy explains this view as follows:

In the exemplar view, the idea that people have a representation that somehow encompasses an entire concept is rejected. That is, one's concept of dogs is not a definition that includes all dogs, nor is it a list of features that are found to greater or lesser degrees in dogs. Instead, a person's concept of dogs is the set of dogs that the person remembers. In some sense, there is no real concept (as normally conceived of), because there is no summary representation that stands for all dogs. (Murphy, 2002)

An exemplar, then, is not a prototypical concept or ideal template, but a composite impression in memory, such that my concept *COMPUTER* is built of a range of knowledge of computers I have consolidated from past experience. Categorization, therefore, involves a process where one accesses all previous memories of computers which may include many different sizes and varieties and compares them to find similarities. Hence, upon seeing a computer I am able to retrieve this information, compare it to the new machine and categorize it based on its similarity to stored exemplars information. Hence, concepts

are constructed of stored information which can be compared against a new example or stimulus.

Exemplar theory avoids the problems of defining characteristics and it is also able to neatly explain typicality phenomena, as the most typical members of a category will naturally be the ones which are most similar to a large number of category members and they will, therefore, be categorized more quickly than less typical items (Nosofsky & Alfonso-Reese, 1999; Nosofsky & Palmeri, 1997b, 1997a). One major disadvantage to the exemplar theory is that it has difficulty in accounting for the hierarchical organization of concepts and would also be an extremely energy-intensive cognitive process if one had to recall and go through all of one's previous memories in order to categorize something (Murphy, 2016).

2.6 The Knowledge Theory

The last theory of conceptual structure I will mention here is the knowledge approach, or sometimes referred to as the '*theory theory*' (Morton, 1980). This approach was proposed in order to develop a more theoretically sound and empirically supported theory of concepts than had previously been posited. This theory was originally framed in terms of the intuitive and naive theoretical knowledge that people have about the world, hence the '*theory theory*' label. As Murphy and Medin (1985) explain:

Our claim is that representations of concepts are best thought of as theoretical knowledge or, at least, as embedded in knowledge that embodies a theory about the world. (Murphy & Medin, 1985)

On this view, it is assumed that the knowledge we have about the world forms a coherent and interconnected whole which contains information about entities, properties, actions and their relations to one another, and it is this knowledge structure which is the raw conceptual material of our concepts:

The knowledge approach argues that concepts are part of our general knowledge about the world. We do not learn concepts in isolation from everything else (as is the case in many psychology experiments); rather, we learn them as part of our overall understanding of the world around us. (Murphy, 2002)

Hence, the formation of concepts and our ability to classify things is influenced by what we already know. Our prior knowledge is not simply used to remember and operationalize specific exemplars to be used in classifying individual concepts, but, instead, all background knowledge is used to make inferences about things and aid categorization. The knowledge view, therefore, emphasizes the use of logical reasoning processes in conceptualization. To give an example of how this works, upon seeing an adult holding the hand of a child in the supermarket, it would be natural to categorize the child as the son or daughter of that adult. In such a case, we make a reasoned inference based on our background knowledge and experience of the world. It is important to note that it is not just our knowledge of a prototype SON or DAUGHTER which is used to make this inference, but also the knowledge that, for example, young children are smaller than their parents, parents often hold hands with their children to keep them safe, if the child was being taken care of by a teacher, they would probably not bring them to the supermarket, and so on. All of this knowledge licences the inference that the child is probably the son or daughter of the adult, even though we have never seen these people before and have no specific information regarding their relationship.

There is much empirical data showing the presence of so-called 'knowledge effects' which provides support for this theory. It has been found, for example, that people are able to memorize lists of things when the things in the list relate to a coherent realm of background knowledge, as opposed to unrelated random objects which people find it more difficult to memorize. This suggests that agents access and make use of background knowledge in order to more easily perform memorization and categorization tasks (G. Murphy, 2002; G. L. Murphy & Allopenna, 1994; Spalding & Murphy, 1996).

Finally, the knowledge approach allows for our concepts to be updated by new experiences and to develop over time. Indeed, this is an advantage of the knowledge approach, as it helps to explain why people's concepts develop and become more sophisticated as they increase in age and experience.

2.7 Defining Concepts

The present work's understanding of concepts is broadly in line with the knowledge view outlined above, following Murphy. Accordingly, a concept is defined as:

A nonlinguistic psychological representation of a class of entities in the world. This is your knowledge of what kinds of things there are in the world, and what properties they have. (Murphy, 2002)

As can be seen from the above, theories of conceptualization have tended to concentrate on categorization. It is this focus which has, plausibly, led to concepts being generally treated as being of a single uniform kind. It is this assumption which I want to challenge. In light of the above considerations, then, the current work assumes an understanding of concepts which is broadly in line with the knowledge view, due to its relative theoretical and empirical plausibility, but which also acknowledges the possible diversity of conceptual kinds. Accordingly, concepts are to be understood here as:

Concepts: Concepts are mental representations constituted from one's coherent knowledge of the world which display divergent forms that are apt to denote entities, properties and relations between them.

I understand 'mental representations', here, to be mental images; the pictures and understandings which can be brought to mind with, or without, linguistic stimuli and which are accessible to introspection.

This characterization is deliberately broad so as not to pre-suppose a particular theory of mind or cognition as either embodied, grounded, modal, amodal etc. – on this point, I aim to stay neutral. I will consider only the phenomenologically self-evident and undisputable mental phenomena experienced when imagining and thinking in the following analysis. That is, the representations we are consciously aware of and which manifest before one's mind's eye during introspection and contemplation or which arise when encountering linguistic stimuli such as '*black coffee*', for example. These are the imaginative phenomena I refer to as 'concepts'. The neurological architecture and cognitive functioning involved in the generation of such images remains to be shown, but

I take it as self-evident and indisputable that there are such things. It is from this foundation that I will proceed to argue for a qualitative distinction between conceptual kinds.

In sum, many theories of conceptualization have neglected to emphasize clear distinctions between various conceptual kinds. I endorse a broadly knowledge-based understanding of concepts as theoretically and empirically plausible, but with the qualification that there exist distinct conceptual kinds. In the following section I will explicate the nature of the link between concepts and language assumed by the foregoing work in preparation for the analyses to follow.



CHAPTER THREE METHODOLOGY

3. General Methodology

The current study is multidisciplinary in nature, not only in the sense that the phenomena of language and mind are of interest to a multitude of academic disciplines, but also due to the fact that it employs varied forms of analysis in exploring these phenomena. In general, this work takes the data and findings of Cognitive Linguistics – a thoroughly scientific field – and applies to them traditionally philosophical praxes of logical and conceptual analysis. Although the fundamentalists in their disciplines often stick hard and fast to the standardly accepted techniques within their fields, it is my view that whether in science or philosophy, we all have the same ultimate goal: to discover and clarify the truth. It is in light of this shared goal that I believe researchers should be open to exploring the methodologies from parallel disciplines in reaching this aim. The physicist and

philosopher Friedrich Waismann noted the potential importance of employing the traditionally philosophical methods of logical and conceptual analysis in the sciences:

It has often been noted that philosophy and science express two very different types of attitude of the human mind. The scientific mind searches for knowledge, i.e., for propositions which are true, which agree with reality. On a high level, it rises to the construction of a theory which connects the scattered and in their isolation unintelligible facts and in this way explains them. But the philosopher cannot be satisfied with this. The very nature of knowledge and truth becomes problematic to him; he would like to get down to the deeper meaning of what the scientist does. Now what can be gained through philosophy is an increase in inner clarity. The results of philosophical reflection are not propositions but the clarification of propositions. (Waismann, 1977)

It is in the spirit of Waismann's words that I bring logical and conceptual analysis into the realm of Cognitive Linguistics and hope, in applying these methods to the data gathered and theories constructed, to scrutinize and clarify the foundational assumptions of the field. Such a program may strike many as distasteful, especially as Cognitive Linguistics seems to have evolved as a reaction to the formalism of traditional logical semantics and generative approaches. However, it would be myopic of theorists working in this field, I submit, to let disagreements with the substantive content of the theories and claims of formalist approaches lead us to disregard their methodologies *in toto*. Indeed, the formal methods of analysis – perhaps exemplified in formal semantics – were founded on thousands of years of development in formal logic from Aristotle to Frege and Russell as a deductive system designed to offer precise and necessary conclusions; we discard this at our peril!

3.1 Conceptual and Logical Analysis

The method of logical analysis – sometimes simply 'analysis' – from which analytic philosophy takes its name, is often difficult to define and in fact consists of many varied (Beaney, 2021)³ elucidations and approaches which stem from the works of

³ See (Beaney, 2021) for a discussion on the differing forms of logical and conceptual analysis employed in analytic philosophy.

Gottlob Frege (1879) and Bertrand Russell (Weitz & Russell, 1961) through the early works of Wittgenstein (1922). In the current work the method of analysis should be understood as consisting in:

The methodical elucidation of the concepts employed by a given claim or proposition and a subsequent logical analysis of the entailments of the clarified propositions in order to assess the formal validity of such propositions in the employment of an argument.

In line with the above characterization, such analysis consists of two parts: the elucidation of the concepts or terms used within a proposition (conceptual analysis) and the subsequent evaluation of the proposition as a whole and the logical entailments of the proposition, and validity of arguments which employ the proposition, which follow as a result of the conceptual clarification (logical analysis).

3.2 Intuition and Introspection

In addition to the aforementioned methods of conceptual and logical analysis, I will also be making use of introspection as a way of accessing and understanding conceptual material and assessing conceptual claims. This method, of course, is not so alien to the field of Cognitive Linguistics, and linguistics in general. Leonard Talmy endorses and explains this method in the following:

The issue of methodology is raised by the fact that cognitive semantics centers its research on conceptual organization, hence, on content experienced in consciousness. That is, for cognitive semantics, the main object of study itself is qualitative mental phenomena as they exist in awareness. Cognitive semantics is thus a branch of phenomenology, specifically, the phenomenology of conceptual content and its structure in language. What methodology, then, can address such a research target? As matters stand, the only instrumentality that can access the phenomenological content and structure of consciousness is that of introspection. (Talmy, 2000)

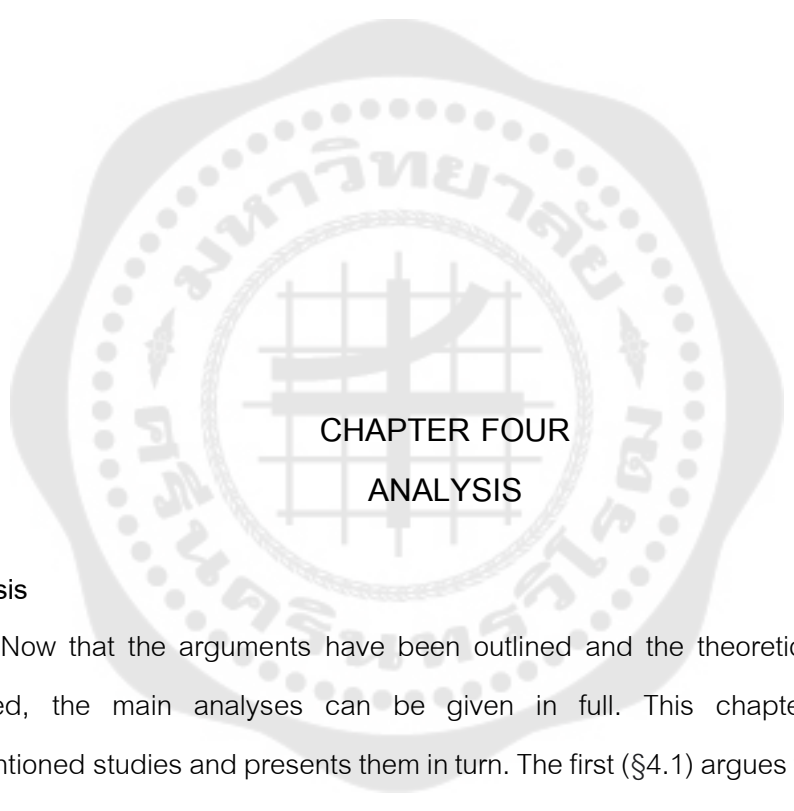
Hence, introspection is used as a means of gathering data about concepts and language to be further analysed. Such a technique, when applied to language, relies on the intuition of the researcher as a native speaker in being able to assess whether an utterance is grammatical or ungrammatical. In the context of psychology and conceptualization, the method makes use of the researcher's intuitive understanding of concepts, their representation and their associations with divergent, but connected concepts. It is of course always possible that intuition, although affording us direct access to conceptual knowledge, is distorted by the influence of cognitive bias and so this is why the employment of such data must be done with discipline and rigor – as is clearly stipulated by Talmy (2000). This consideration necessitates the use of conceptual and logical analysis in the application of such data.

3.3 Secondary Data

As well as intuitively collected data, secondary data will also be collected and analyzed in the same manner in order to assess the validity of the foundational claims concepts and frameworks currently employed.

3.4 The Lexical Approach

Lastly, in addition to the linguistic data collected through introspection, the following study will also make use of what can be called the 'lexical approach' (Kövecses, 2020) which involves collecting examples of language taken from English language published articles, webpages, books and dictionaries. This approach is a standard way of collecting linguistic data from various sources and has been used throughout the various arguments presented here and used in conjunction with the above-mentioned methods in an interdisciplinary way.



CHAPTER FOUR ANALYSIS

4. Analysis

Now that the arguments have been outlined and the theoretical assumptions elucidated, the main analyses can be given in full. This chapter contains the aforementioned studies and presents them in turn. The first (§4.1) argues for the emotional construction of moral concepts, the second (§4.2) critiques Conceptual Metaphor Theory as a means of probing the conceptual structure of moral concepts, the third (§4.3) examines the binary opposite nature of RIGHT and WRONG.

4.1 The Emotional Construction of Moral Concepts:

Understanding what morality is and why we classify some actions as 'right' or 'wrong' are foundational concerns in metaethics and moral psychology. An interesting and oft-made observation is that there appears to be a *prima facie* link between moral thought and emotion. But what is the nature of this link? Is the connection between morality

and emotion merely a reactionary one, or is it something stronger; a necessary one? I claim the answer is the latter.

There are many theoretical frameworks which have been employed in trying to understand the composition and structure of moral concepts. One of the most dominant of these in the cognitive science literature is Conceptual Metaphor Theory (CMT). In the following section I will attempt a CMT-based analysis of moral concepts, against the background of the preceding emotionist study in order to see if the framework is able to shed more light on our *analysandum*.

4.2 The Metaphorical Structure of Moral Concepts⁴

As shown above, there is ample evidence pointing to the emotional construction of moral concepts. But in order to develop a deep and comprehensive understanding of concepts and how they are structured, it is necessary to analyse them from various perspectives, employing diverse techniques. In the current section, then, I will proceed to examine the conceptual domain MORALITY along with its constituent concepts RIGHT and WRONG through the lens of Conceptual Metaphor Theory, in order to see if the results from the previous section can be reconciled with the data garnered from a CMT-based analysis.

Intriguingly, however, in light of the wide range of evidence implying the role of emotion in moral thinking, it is surprising that the probable emotional structuring of moral concepts has not been drawn out or emphasized in the CMT literature. This leads to the question of whether CMT is effective as a framework for probing the nature of moral concepts, which I will proceed to explore in the following paragraphs.

4.2.1 Conceptual Metaphor Theory & Moral Concepts

CMT is, therefore, a theory about conceptual structure which holds that conceptual structure can be probed through the analysis of metaphorical utterances. The main aim of this section is to test the effectiveness of Conceptual Metaphor Theory in discovering the conceptual structure of the domain MORALITY and its constituent concepts

⁴ Work from this section has, in line with university requirements, been previously published and is reproduced here courtesy of Axiomathes and Springer Nature. See (Bartlett & Ruangjaroon, 2022).

RIGHT and WRONG through the analysis of linguistic metaphors. The guiding question, then, is:

Is the framework of Conceptual Metaphor Theory an effective means for discovering the conceptual structure of moral concepts?

In supplying an answer to this question, it will be shown that a CMT-based analysis – although able to give interesting insights into the way English speakers describe moral actions and have implicit conceptual associations – is a tool unfit for the purpose of discovering the structure of moral concepts.

4.2.2 The Foundations of CMT

Although *prima facie* CMT seems intuitively appealing, previous challenges, along with the oppositions to be presented here, I contend, should be taken seriously. Logical analysis of the main theoretical claims of CMT uncovers flaws which should lead theorists to seriously reconsider the ways in which they approach the study of the human conceptual system henceforth. In order to demonstrate this, I will proceed by initially introducing the foundational claims of CMT in preparation for the criticisms to follow.

In the decades since it was first proposed, CMT has seen much development and elaboration by various scholars (Gibbs, 2017; Kövecses, 2020; Lakoff, 1996), but the foundational assumptions of the framework remain the same. It will, therefore, be instructive to start this study by first outlining and examining the theory as it was originally put by George Lakoff and Mark Johnson more than four decades ago. In Lakoff and Johnson's *Metaphors We Live By* (1980), they begin by emphasizing the importance of understanding the machinations of human conceptualization and, vitally, claim that the human conceptual system is 'metaphorical' in nature. In doing so, they start to lay the foundations of Conceptual Metaphor Theory. Lakoff and Johnson state that:

"Our ordinary conceptual system, in terms of which we both think and act, is fundamentally metaphorical in

nature. The concepts that govern our thought are not just matters of the intellect. They also govern our

everyday functioning, down to the most mundane details. Our concepts structure what we perceive, how we get around in the world, and how we relate to other people. Our conceptual system thus plays a central role in defining our everyday realities. If we are right in suggesting that our conceptual system is largely metaphorical, then the way we think, what we experience, and what we do every day is very much a matter of metaphor.” (Lakoff and Johnson, 1980, p3)

Lakoff and Johnson state the foundations of the theory more concisely by explaining that ‘The essence of metaphor is understanding and experiencing one kind of thing in terms of another’ (Lakoff & Johnson, 1980). This cornerstone of CMT is still espoused in the contemporary literature. As Zoltan Kövecses puts it: ‘In the cognitive linguistic view, metaphor is defined as understanding one conceptual domain in terms of another conceptual domain’ (Kövecses 2010, p4). To elaborate, CMT states that the human conceptual system uses a ‘metaphorical’ structuring process in order to aid the understanding of abstract concepts. This means that relatively abstract conceptual domains (target domains) are understood in terms of more concrete conceptual domains (source domains); that in order for humans to grasp the target concepts, our minds must structure them by making use of simpler source concepts. The mental constructs born from this cognitive process are dubbed ‘conceptual metaphors’.

To give an example, a now classic and well-documented case of a conceptual metaphor originally proposed by Lakoff and Johnson (1980) is called *ARGUMENT IS WAR*. Hence, it is claimed that the way in which we understand arguments is necessarily underpinned by our conceptual knowledge of war – we understand arguments *in terms of war*. CMT, therefore, makes a claim about the way in which the human conceptual system functions. It states that the mental representations of abstract concepts are determined by the form of the underlying concrete concepts to which they are mapped and that these mappings are the necessary cause of linguistic metaphor.

Lakoff and Johnson make this psychological hypothesis based on linguistic evidence gathered intuitively from everyday utterances that appear to make manifest widespread and systematic metaphors in certain areas of discourse. It seems, at first glance, apparent from analysis of figurative and metaphorical utterances that many natural language expressions make visible, or express, the underlying conceptual structures (conceptual metaphors) hypothesized. To offer an example, Lakoff and Johnson's original evidence of the proposed conceptual metaphor ARGUMENT IS WAR consists of the following collection of English language phrases:

ARGUMENT IS WAR

Your claims are *undefensible*.

He *attacked every weak point in* my argument.

His criticisms were *right on target*.

I *demolished* his argument.

I've never *won* an argument with him.

You disagree? Okay, *shoot!*

If you use that *strategy*, he'll *wipe you out*.

He *shot down* all of my arguments.

(Lakoff and Johnson, 1980, p4)

It is posited that the above utterances tell us something about the way in which we understand the concept ARGUMENT: that there is an underlying conceptual process which allows us to grasp the concept ARGUMENT by mapping it onto our concept of WAR. It is

held that the conceptual metaphor *ARGUMENT IS WAR* is made explicit by linguistic metaphors and figurative language, such as those listed above, and that analysis of language, therefore, gives us a way of examining the constitution of abstract concepts; that conceptual structure can be 'read off' from our figurative utterances. Lakoff and Johnson explain this link between conceptual structure and language as follows: 'The concept is metaphorically structured, the activity is metaphorically structured, and, consequently, the language is metaphorically structured' (Lakoff & Johnson, 1980, p5). It is argued that there is a systematic relationship between conceptual structure and linguistic structure, such that figurative utterances provide a window to the mind. Such utterances are deemed to be the products of conceptual metaphors, this is made clear by Lakoff and Johnson:

Our conventional ways of talking about arguments [in terms of war] pre-suppose a [conceptual] metaphor we are hardly ever conscious of. (Lakoff & Johnson, 1980: p5)

Conceptual metaphor theorists have subsequently devoted much time to the analysis of metaphorical and figurative language in order to discover the underlying conceptual structures of the mind. This has led to the putative discovery of a multitude of conceptual metaphors at varying levels of generality, or schematicity, which include such examples as *ARGUMENT IS WAR*; *RELATIONSHIPS ARE JOURNEYS* and *SEEING IS UNDERSTANDING* – to mention just a few prominent examples.⁵ CMT does indeed have much intuitive appeal, but, beyond this, how solid are its foundational claims?

4.2.3 Under Attack

As we have seen above, metaphorical utterances are claimed to be made manifest by hypothesised conceptual metaphors in the mind, and the data presented in support of these conceptual metaphors consists of the very same linguistic data which led to the hypothesis (Vervaeke & Kennedy, 1996). This attack is justified, as such an argument is viciously circular. Raymond Gibbs explains:

⁵ See Kovecses (2010) for a comprehensive overview of various conceptual metaphors.

[A]n abundance of experimental evidence to support the claim that conceptual metaphors are an essential part of verbal metaphor use, but that future research requires greater sophistication regarding the degree to which conceptual metaphors influence speaking and understanding metaphor. (Gibbs, 2017)

Notwithstanding the non-linguistic psychological evidence to the contrary, the fact that such experimental data is needed to support CMT should tell us something about which methodologies would be effective in discovering and describing conceptual structure and also serves to remind us what CMT is actually a theory about. As Vyvyan Evans points out, CMT is not a theory about language, but human psychology:

Despite the importance of Conceptual Metaphor Theory in terms of accounting for deeply ingrained systematicities in conceptual structure, it is not a theory about language, nor about figurative language understanding. Rather, Conceptual Metaphor Theory primarily provides an account of knowledge representation. Indeed, in spite of its success, it fails to adequately account for systematicities in language, for instance within a single language, nor in terms of accounting for detailed differences in figurative expression that emerge cross-linguistically. (Evans, 2009)

A pertinent question to ask here is: If specific cross-domain conceptual mappings were to be conclusively proven as cognitively real via experimental psychological and neurophysiological means, would theorists thereby be justified in continuing to use linguistic metaphors as evidence for conceptual metaphors? I claim not. If the objection from circular reasoning, and the means by which CMT theorists have tried to avoid it, has taught us anything, surely it is that we should be more careful about our approach to the study of the conceptual system; that methodologies employing the analysis of linguistic metaphors in inferring the existence of conceptual metaphors might be inadequate.

In sum, CMT has come under attack by a number of theorists who claim that the methodologies and reasoning employed in discovering cross-domain conceptual mappings are flawed. Some of these objections have apparently been met by pointing to non-linguistic evidence as external support for the theory which, CMT theorists claim, justifies the continued study of metaphorical and figurative utterances in the search for conceptual metaphors. Here one gets the feeling of being told, in light of the empirical evidence, “look! We were right all along!”.

4.2.4 The Multiple Source Domain Problem

In addition to the issues mentioned above, I will proceed to cast more doubt on the adequacy of CMT by raising two novel objections to the framework which focus not on its methodological flaws *per se*, but on the theoretical coherence of its foundational claims, examining the conceptual domain MORALITY as a case in point. It will become apparent, during the course of the foregoing, that problems arise for CMT when an analysis of figurative language shows MORALITY to be structured by multiple divergent source domains. This is problematic as it raises the question:

Which of these various source domains is *necessary* to the conceptual structure of MORALITY?

This issue needs to be addressed, as if it is claimed that multiple source domains are necessary for structuring MORALITY – or, indeed, any given target concept – it can be argued – counter to the foundational claims of CMT – that conceptual metaphors are not sufficient for the production of linguistic metaphors and figurative language, due to the fact that a given figurative utterance typically only makes apparent a single source domain. This observation raises doubts not only about the conceptual structuring process it posits, but also about the methods one employs in inferring the existence of conceptual metaphors from the analysis of figurative language. In light of this observation, I propose a theoretical work-around that seeks to solve this puzzle, but ultimately conclude that, even with the plausible supplementary theory, CMT does not escape the main theoretical and methodological issues brought to light. The suggested enrichment, therefore, fails.

In addition, I present an argument which exposes more clearly the theoretical flaws in the foundational claims of CMT. The claims that justify the standard method of analysing figurative language in uncovering conceptual metaphors are shown to be problematic, thus leading to the conclusion that CMT, as it stands, is not an effective tool with which to examine the structure of moral (or any other) concepts.

4.2.5 The Moral High Ground

As explained above, CMT is assumed to provide theorists with a means of examining conceptual structure through the study of metaphor. In order to test this foundational assumption, and the effectiveness of CMT in uncovering the conceptual structure of MORALITY, I will proceed, now, to show how a CMT-based investigation into the domain MORALITY might look, by incorporating data from previous studies and original examples, before elucidating how such an analysis is unsatisfactory.

There are several existing analyses of moral metaphor (Johnson, 1993; Lakoff, 1996b; Lakoff & Johnson, 1980). In the following sections I will attempt, as the aforementioned authors have done, an analysis of figurative utterances which are assumed by CMT theorists to expose the main conceptual metaphors which structure the abstract domain MORALITY.

The first analysis of metaphor in moral discourse undertaken by Lakoff and Johnson (1980) showed MORALITY as being structured by the domain of VERTICAL ORIENTATION. This orientational metaphor structures morality in such a way as to map the moral concepts RIGHT and WRONG with UP and DOWN, respectively. Lakoff and Johnson originally called this particular mapping VIRTUE IS UP; DEPRAVITY IS DOWN (1980):and presented the following linguistic data as evidence for this mapping:

VIRTUE IS UP; DEPRAVITY IS DOWN

He is *high-minded*.

She has *high standards*.

She is *upright*.

She is an *upstanding citizen*.

That was a *low trick*.

Don't be *underhanded*.

I wouldn't *stoop to that*.

That would be *beneath me*.

He fell into *the abyss of depravity*.

That was a *low-down* thing to do. (Lakoff & Johnson, 1980)

The vertical orientation metaphor shown here does, indeed, seem to be widespread and consistent throughout English moral discourse. In order to show how prevalent this way of speaking is in English, here are some further examples of the same mapping which I will call 'MORALITY IS VERTICAL ORIENTATION', as this represents the most schematic level of this mapping:

MORALITY IS VERTICAL ORIENTATION

- (1) You have (can claim) the *moral high ground*.
- (2) That was *morally base*.
- (3) He has hit *rock bottom*.
- (4) You have *gone down* in my expectations.
- (5) Don't *lower the tone*.
- (6) He is the *lowest of the low*.

- (7) You're on a *slippery slope*.
- (8) I won't take *the fall* for you.
- (9) I'm between the *devil* and *the deep blue sea*.
- (10) I would never *sink so low/to such depths*.
- (11) *Lower than a snake's belly!*

Linguistic metaphors, such as the ones shown above, appear to make apparent the general (or schematic) level mapping of the source domain VERTICAL ORIENTATION with the target domain MORALITY, which gives rise to the more specific conceptual metaphors RIGHT IS UP and WRONG IS DOWN. In English, we can see that moral virtue is often spoken of in terms of upwards direction, movement or high places, and, conversely, moral vice is expressed in terms of down, low or under (Orientational metaphors for MORALITY are also observed in Chinese as shown by Yu (2016)).

This link between VERTICAL ORIENTATION and MORALITY is also observed in Judaeo-Christian cosmology. In Christianity, *the fall of man*, which describes the transition of humankind from an innocent and virtuous state to one of guilt and sin, is perhaps an exemplary case of this metaphor in religious culture. Moreover, in religious traditions generally, one very often sees depictions of heaven as above and hell as below, correlating, of course, with the RIGHT IS UP; WRONG IS DOWN mapping (Jones, 2019).

I am, here, starting to undertake a traditional intuitive analysis of MORALITY through the framework of CMT, whereby figurative language examples are compiled, either intuitively or lexically, and it is inferred from them that a particular conceptual structure exists in the mind of the speaker. Accordingly, the analysis above has seemingly allowed us to discover the conceptual metaphor MORALITY IS VERTICAL ORIENTATION. In the following section, I will proceed to repeat this process with more examples in order to expose other source domains which appear to structure the target MORALITY.

4.2.6 Doing the Dirty

As pervasive as speaking of MORALITY in terms of VERTICAL ORIENTATION is in English, it appears from a CMT-style analysis that MORALITY is not understood exclusively in terms of VERTICAL ORIENTATION, but can also be expressed in other ways; that is, the target domain MORALITY appears to be structured by more than just the single source domain VERTICAL ORIENTATION.

Indeed, when looking at figurative language in moral discourse, many other common mappings will become apparent to the CMT theorist. One among them is MORALITY IS CLEANLINESS – meaning we speak and conceive of RIGHT as CLEAN and WRONG as DIRTY, which has been noted previously by other theorists (Huangfu et al., 2021; Lizardo, 2012) Below are some linguistic examples which, CMT theorists might claim, make this conceptual metaphor explicit:

MORALITY IS CLEANLINESS

- (12) She has *clean hands*.
- (13) He is *Mr Clean!*
- (14) I am going to *make a clean breast* of it.
- (15) I have a *clean conscience*.
- (16) *Wash your hands* of it.
- (17) They *did the dirty* on me.
- (18) Don't get your *hands dirty*.
- (19) They were *playing dirty*.

- (20) He *dished the dirt*.

As with the MORALITY IS VERTICAL ORIENTATION metaphor, MORALITY IS CLEANLINESS is also assumed to be a common mapping made evident by a range of linguistic metaphors, as shown by (12-20) and can be found to have a long history in English. *The Thesaurus of Traditional English Metaphors* (Wilkinson, 2002), for instance, contains several entries of figurative expressions which make manifest the same metaphorical mapping:

- (21) Muddy springs will have muddy streams: The outcome of bad parentage or upbringing will also be bad.
- (22) Soft as shit and twice as nasty [sEng]: Applied by country folk to visitors from the cities, especially those with loose morals.
- (23) Mud-bath: Period of moral depravity.
- (24) Folk often get a good meal out of a dirty dish: Unclean or immoral persons are nonetheless capable of good deeds.
- (25) He that falls in the dirt, the longer he lies the dirtier/fouler he is: An encouragement not to resign yourself to a deplorable condition; clear yourself of slander quickly. (Wilkinson, 2002)

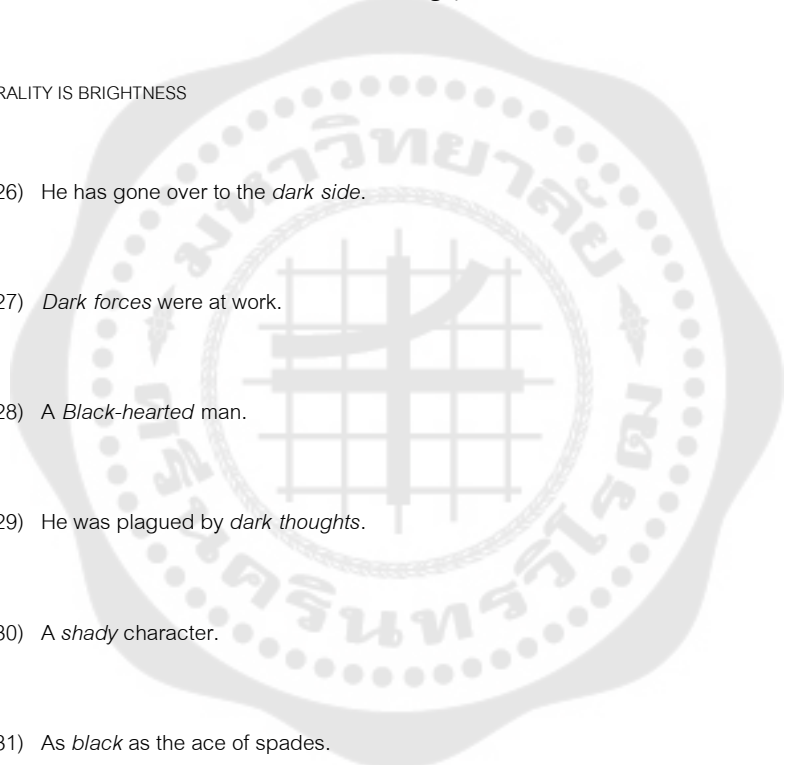
Assuming the truth of the foundational claims made by CMT, we might well conclude that the examples of figurative language presented above are evidence for a corresponding conceptual metaphor in the mind of the speaker; that due to the fact that moral depravity (WRONG) is spoken of in terms of dirtiness, and moral superiority (RIGHT), in terms of cleanliness, that our understanding of the source domain CLEANLINESS is necessary

to our understanding of the target domain MORALITY. Thus, we have, seemingly, exposed another conceptual metaphor: MORALITY IS CLEANLINESS.

4.2.7 The Dark Side

As well as the above putative conceptual metaphors of MORALITY IS VERTICAL ORIENTATION and MORALITY IS CLEANLINESS, another commonly discussed cross-domain mapping is that of MORALITY with BRIGHTNESS. This mapping gives rise to the specific metaphors RIGHT IS LIGHT and WRONG IS DARK which are easily found throughout moral discourse (Kövecses, 2010, p. 21). And can be observed in the following phrases:

MORALITY IS BRIGHTNESS

- 
- (26) He has gone over to the *dark side*.
 - (27) *Dark forces* were at work.
 - (28) A *Black-hearted* man.
 - (29) He was plagued by *dark thoughts*.
 - (30) A *shady* character.
 - (31) As *black* as the ace of spades.
 - (32) I was lost in *darkness*, but now I have seen the *light*.
 - (33) They tried to *whitewash* their actions.
 - (34) She saw *the light*.
 - (35) A *white lie*.

(36) She was *whiter than white*.

Again, we have found a conceptual metaphor: MORALITY IS BRIGHTNESS.

4.2.8 Paying the Price

Another frequently discussed metaphor in relation to morality is that of economic transactions (Johnson, 1993; Kövecses, 2010; Lakoff, 1996b; Lakoff & Johnson, 1980). This is normally called the 'moral accounting' metaphor. Mark Johnson explains it as follows:

[I]n the moral domain we understand our actions metaphorically as commodities exchanged, and we expect their (metaphorical) values to balance out in the end. If I perform good acts, I build up a form of moral credit. If I harm you, then you deserve a certain restitution or payback that balances out the harm done. (Johnson, 1993, p. 45)

I will refer to this mapping as MORALITY IS FAIR TRANSACTION and, in doing so, I am referring to the most schematic form of the accounting metaphor whereby it encompasses all value-based moral metaphors. This is because the moral accounting metaphor appears to be fairly complex and includes several distinct, but genetically related, concepts, as Lakoff explains:

The general metaphor of Moral Accounting is realized in a small number of basic moral schemes: reciprocity, retribution, restitution, revenge, altruism, and others. Each of these moral schemes is defined using the metaphor of Moral Accounting [...] (Lakoff, 1996a)

Some examples of how this mapping is made manifest in language are shown below:

MORALITY IS FAIR TRANSACTION

- (37) I *owe* you
- (38) You will *pay* for what you did
- (39) I am *in your debt*
- (40) Is the punishment *worth* it?
- (41) Imprisonment is a *high price to pay*
- (42) The *scales* of justice
- (43) One good turn *deserves* another
- (44) *Two wrongs* don't make a right

And there we have it, another conceptual metaphor: MORALITY IS FAIR TRANSACTION.

4.2.9 Deplorable Acts

In spite of the above being an interesting record of figurative language in moral discourse, and the previous work on moral metaphors apparently allowing for a deeper understanding of the way in which MORALITY is conceptually structured, the work above is significant for what it is lacking.

CMT-based accounts of moral concepts have failed to draw an explicit connection between emotion and morality, which, as I have shown in (Section 4.1), is clearly there (even if the exact nature of that connection is disputed). There is much reason, then, as I have argued, to suspect that EMOTION plays an intrinsic, and potentially necessary, role in our understanding of MORALITY and so it might be seen as a failure of previous analyses (or indeed CMT as a framework) that a 'MORALITY IS EMOTION' conceptual metaphor has not been 'discovered'.

This putative metaphor of MORALITY IS EMOTION appears to be absent from the CMT literature. Zoltan Kövecses tangentially mentions an apparent connection between morality and emotion in reference to the A'ara people of the Solomon Islands (Kövecses, 1990) and also observes how morality and emotion can be conceptualized in terms of force (2000). Here, Kövecses points out the differences and similarities between the EMOTION AS FORCE and MORALITY AS RESISTING A PHYSICAL FORCE metaphor, but does not speak of a MORALITY IS EMOTION metaphor (Kövecses, 2000). The closest Kövecses comes to explicitly formulating this metaphor is in his book *Where Metaphors Come From: Reconsidering Context in Metaphor* (2015), where he claims:

If we examine the content of the idealized cognitive models associated with emotion or other emotion concepts, we find that they greatly overlap with Harré's rules of emotion. According to Harré (1994), in the course of the appropriate use of emotion words in different cultures people observe certain "local rules." The rules are of four kinds, "classified by reference to what is criterial for their correct usage": (a) "appropriate bodily feelings," (b) "distinctive bodily displays," (c) "cognitive judgments," and (d) "moral judgments" and the "social acts" corresponding to them (p. 7). (Kövecses, 2015)

Although Kövecses makes a vague connection between EMOTION and MORALITY, a conceptual metaphor connecting the two is not explicitly formulated. Given the psychological data confirming the connection between morality and emotion, one would expect the MORALITY IS EMOTION metaphor to be a psychologically real and extremely salient cross-domain mapping. The fact that this connection has largely gone unnoticed casts doubt on the effectiveness of CMT in analysing the structure of MORALITY, or, indeed, any concept.

This failure might, I conjecture, be down to the fact that MORALITY IS EMOTION metaphors are easily overlooked, or interpreted wrongly; let's take an example from Lakoff. In Lakoff's (1996) analysis, he places MORALITY as a target which has CLEANLINESS as its

source and gives the example of ‘that was a *disgusting* thing to do’ (Lakoff, 1996a, p. 92) as evidence for this mapping. However, although, as noted above, the mapping between the two domains appears to exist, I believe this particular analysis to be mistaken. In Lakoff’s example, we see the conflation, I claim, of two separate source domains: EMOTION and CLEANLINESS. This is due to the word ‘disgusting’, in reference to cleanliness, is already metaphorical. In saying that something dirty is ‘disgusting’ (it causes disgust), we are firmly in the domain of EMOTION.

The concept DISGUSTING belongs properly to the domain of EMOTION, but is applied to CLEANLINESS insofar as dirty or rotten things can induce disgust – an emotional response – in the observer. This emotional concept can, therefore, be applied to the target of CLEANLINESS. Hence, an utterance such as ‘the floor is *disgusting*’ meaning ‘the floor is *dirty*’ is, in fact, figurative language. Therefore, the correct analysis of the moral utterance ‘that was a *disgusting* thing to do’ should be that it makes evident the conceptual metaphor of MORALITY IS EMOTION and not MORALITY IS CLEANLINESS. This conflation, however, flies very easily under the radar as the connection between the two concepts DIRTY and DISGUSTING is embodied, with disgust being an emotion-dispositional concept (see Section 4.1).

There are, indeed, figurative utterances which, seemingly, make the MORALITY IS EMOTION mapping manifest:

MORALITY IS EMOTION

- (45) Their actions were *disgusting*.
- (46) They committed truly *heinous acts*.
- (47) His crimes were *disturbing*.
- (48) She is *contemptible*.
- (49) They were kept in *deplorable* conditions

(50) Knowledge of the *abhorrent/repulsive* acts made me *sick to the stomach*.

(51) You make *me sick/repel me!*

(52) His conduct was *despicable*.

(53) It was a *crime of passion*.

(54) That was *shameful*.

(55) I have a *guilty conscience*.

The above examples, a CMT theorist might claim, show the conceptual mapping between EMOTION and MORALITY. Interestingly, the moral metaphors above appear to track different facets of emotional reactions such as physiological responses, as in ‘you make me *sick*’, and psychological affective components from both self-blame and other-blame perspectives as with ‘his conduct was *despicable*’, which I have explicated in Section 4.1⁶.

It should also be noted that it is often the case that employing emotional terms in moral discourse is necessary for conveying the perceived severity of people’s actions. For instance, ‘It was a *dirty trick*’ or ‘It was *underhanded*’ are not as powerful as ‘It was a *despicable crime*’. When describing extremely immoral acts, emotional terms are sometimes the most appropriate. When describing the torture that occurred in Nazi death camps, for example, it is not enough to say that such actions were ‘base’, ‘dirty’, ‘bad’ or even ‘wrong’ – these adjectives simply do not capture the full extent of the immoral behavior. Using terms such as ‘*heinous*’, ‘*deplorable*’ or ‘*disgusting*’ are necessary to convey the true level of the moral transgression in such cases.

It appears, then, that in English a wide range of emotional adjectives such as *heinous*, *despicable*, *deplorable*, *abhorrent*, *appalling*, *awful*, *disgusting*, *loathsome*,

⁶ See Bartlett (2020) for discussion.

obscene, odious, repulsive, shocking and sickening are commonly used in place of moral adjectives⁷, and, in being so used, appear to be metaphorical⁸. Such terms, therefore, may be understood as representing the conceptual metaphor of MORALITY IS EMOTION, a CMT theorist might claim.

It is problematic, then, that CMT has not found this link and hints at its inadequacy as a framework, especially when the link between emotional and moral thinking has been extremely well documented in the psychological and neurological literature. It should lead CMT theorists to ask: Why? It might be due, simply, to the fact that no one has picked up on this kind of language before; or it might be because the method of picking out figurative utterances is particularly dependent on the theorist's pre-conceived notions of morality (bias), or even that the method employed is unreliable for some reason, such as it relying on a shaky or open-to-interpretation distinction between literal and figurative utterances i.e. should such words as 'shameful' or 'deplorable', when used to express moral indignation, be considered as literal or figurative?⁹

4.2.10 The Problem of Divergent Source Domains

Putting the failure to find the MORALITY IS EMOTION metaphor aside, we are now faced, in light of the above, with the task of making sense of the fact that multiple source domains can structure the same target domain. This, I claim, is problematic for CMT because it raises the issue of how this is possible, given the foundational assumptions of the theory. It is vital that this issue is resolved, as it potentially undermines the claim that abstract target domains are necessarily structured by concrete source domains; that the target concept is necessarily understood in terms of the source concept.

The problem can be stated by way of an example: Above we have inferred from linguistic metaphors such as 'I would never *sink that low*' the existence of the conceptual metaphor MORALITY IS VERTICAL ORIENTATION. Similarly, we have also concluded, from phrases such as 'He did *the dirty*', another conceptual metaphor: MORALITY IS CLEANLINESS.

⁷ The Corpus of Contemporary American English shows 'disgusting' appearing in clearly moral contexts in 30 of the first 50 entries (Davies, 2008).

⁸ Although this not clear, as I will discuss in the following sections.

⁹ This is a point to which I will return below.

Hence, it appears that MORALITY is mapped with, among others, these two divergent source domains i.e. VERTICAL ORIENTATION and CLEANLINESS. This means, from the perspective of CMT, that both source domains are necessary for the structuring and understanding of MORALITY; that is, we cannot grasp MORALITY without them.

Now, we could ask the question: When a speaker utters the phrase 'I would never *sink that low*' in a particular instance, are they *only* conceptualizing MORALITY as VERTICAL ORIENTATION (a single source-target mapping) in that instance? If this is the case, then CLEANLINESS, and other source domains, are not necessary for the understanding, or structuring of MORALITY. However, if it is suggested, on the other hand, that the speaker, in this instance, is conceptualizing MORALITY in terms of both VERTICALITY and CLEANLINESS simultaneously – which is what needs to be the case in order to save CMT's claim of necessary conceptual structuring – another troubling question arises: Why, in that instance, did the speaker use a metaphor which only made visible one of the two necessary source domains?

In light of this observation, the CMT theorist is faced with a dilemma: Either it is the case that not all source domains are conceptually necessary to the structure of the target, or it is the case that source domains *are* all conceptually necessary, but are not sufficient for the production of figurative utterances. If the first horn of this dilemma is true, CMT's conceptual claim is defeated, and if the second horn is true, the standard methodology of inferring conceptual metaphors from linguistic metaphors and figurative language (as demonstrated above) is ineffective, as in order to validly infer the existence of a conceptual metaphor from the presence of a linguistic metaphor, the former must be necessary *and* sufficient for the latter¹⁰.

I claim that a possible way out of this dilemma is to hold onto the conceptual claim and try to explain how an isolated figurative utterance, appearing to show conceptual structuring from only one source domain on the surface, in fact, carries the semantic and conceptual content from *all* source domains simultaneously. I posit that

¹⁰ The issue of necessary and sufficient conditions is a serious problem for CMT; one to which I will return in the following section.

this would need to be done by postulating a conceptual relation between all divergent source domains which map onto a given target; such that a single metaphorical utterance carries all the conceptual/semantic weight of all necessary source domains, which I will precede to elucidate. However, I will ultimately conclude that although the proposed theoretical enrichment of CMT avoids the problem of multiple source domains, it does not get away from the second horn of the dilemma: that conceptual metaphors are not sufficient for linguistic metaphors.

4.2.11 Building Conceptual Layers

From the above CMT-style analysis of moral language, the target domain MORALITY appears to be structured by the divergent source domains of VERTICALITY, CLEANLINESS, BRIGHTNESS, FAIR TRANSACTION and EMOTION. In light of this observation, I have claimed that CMT is faced with the problem of explaining the necessity of divergent source domains which are found to be structuring a single target. This is not a new observation and has been discussed before in the literature as the phenomenon of *scope* (Kövecses, 1995). Conceptual domains are said to have varying degrees of scope. The scope of a domain is said to be 'wide' if the domain can be used as a source domain for a multitude of targets. Having a wider scope means, therefore, that a domain is conceptually simple, and it is this simplicity that means it can be used more generally to structure a diverse range of abstract concepts. Domains that are highly abstract are predicted to have a narrower scope and will be found as source domains for relatively few targets.

4.2.12 The Scope of Metaphor

Zoltan Kövecses presents an analysis of the concept FRIENDSHIP in American culture and, through analysis of linguistic metaphors, finds that there are several divergent source domains used to structure the target of FRIENDSHIP. Kövecses claims there are no source domains that exclusively structure FRIENDSHIP as they can act as source domains for other targets. In order to make sense of this, Kövecses first introduces the concept of 'scope' into the vocabulary of Conceptual Metaphor Theory. He makes the observation that some conceptual domains make use of a variety of linguistic metaphors suggesting that they have several divergent source domains and *vice versa*. As Kövecses explains, 'this paper raises the issue of the generality of the application of particular source domains

to particular target domains and suggests a new, theoretically useful notion for this purpose: that of the "scope of metaphor" (Kövecses, 1995). Kövecses shows, using the case of FRIENDSHIP, that source domains come together to form underlying *metaphor systems* and suggests that the metaphorical conceptualization of FRIENDSHIP is determined by the links between all domains involved in the metaphor system.

Understanding that conceptual domains exhibit varying degrees of scope appears intuitive, and indeed provides us with a rough answer to the question of why certain domains can be found to structure multiple target domains.

As useful as this notion is, the problem we are faced with here is the inverse. That is to say, the problem unearthed from the analysis of MORALITY presented above is not how each source domain can be applied to several targets, but how a single target can be mapped with multiple sources. We could, using the notion of scope, argue that the highly abstract domain MORALITY is structured by several source domains of wider scope and that these source domains sit in a hierarchical *multi-layer mapping*, the order of which is determined by the scope of each domain in the mapping. However, I claim that for CMT to be able to adequately account for the phenomenon of multiple source domains, an abstract domain such as MORALITY which is mapped with, and therefore conceptually structured by multiple sources, there must exist a conceptual relationship between each source domain; such that, if a target domain x is mapped with source domains y and z , there must also be a conceptual relationship held between y and z .

If conceptual metaphors arise in virtue of there being a structural or conceptual similarity between source and target, then in being able to map the same target, it is not implausible that the divergent source domains could also be conceptually similar to each other in some non-trivial way. If such a relationship can be shown to exist, it will be possible to construct a hierarchically organized multi-layer mapping which includes all source domains to a single target.

To give a brief example, the concept 'BOOK' is not comprehensible unless we already understand PAPER. It is also a pre-requisite for understanding BOOK that we know

the concept *OBJECT*, which is also, of course, needed to understand *PAPER*. *BOOK* therefore is a concept dependent upon the logically prior concepts *OBJECT* and *PAPER*.

My claim is that if understanding *OBJECT* and *PAPER* are conceptual prerequisites for understanding *BOOK*, that *OBJECT* and *PAPER* must also bear some conceptual relation to each other – and, of course, they do. We must, first of all, understand *OBJECT* in order to understand *PAPER* and grasping *PAPER* is necessary to comprehending *BOOK*. Hence, we have a chain of concepts *OBJECT* → *PAPER* → *BOOK* which are dependent upon each other and sit together in a hierarchical order, with *OBJECT* being the most superordinate and logically prior concept.

With this example in-hand, I will attempt, in the following paragraphs, to show how an analogous conceptual structure could be shown for the abstract target domain *MORALITY*. The theoretical construct of the multi-layer mapping might help CMT to overcome the problem posed by multiple source domains, in showing that all source domains are necessary for the structuring of the target.

If this can be shown, it will allow CMT theorists to plausibly claim that all divergent source domains are necessary to the conceptual structure of the target, insofar as understanding the target logically implies all necessary source domains. If this cannot be shown, it is difficult to see on what basis CMT theorists can claim that all divergent source domains are necessary to the target, which is the cornerstone of the theory.

4.1.13 From the Ground Up

Having already made explicit the divergent source domains which apparently structure *MORALITY*, such as *VERTICAL ORIENTATION* and *CLEANLINESS*, we can now, following my above suggestion, proceed to analyse the source domains in respect to each other so that we might see if a source-target relationship holds between them.

If we do this in the same fashion as above, by looking at conventional metaphors, it would appear that the domain of *CLEANLINESS* is already mapped with *VERTICAL ORIENTATION*. *VERTICAL ORIENTATION* acts as a source domain for *CLEANLINESS*, with *DIRTY* as *DOWN* and *CLEAN* as *UP*. This mapping is shown in phrases such as ‘get *down* and *dirty*’ ‘clean *up*’ and ‘you *scrub up* well’. This specific mapping might also be well-grounded in experience as

the floor is a prominent source of dirt and, consequently, keeping ourselves and our possessions off the ground is routinely done in avoidance of getting dirty which means the conceptual link between VERTICALITY and CLEANLINESS is often reinforced, making the connection a salient one. As Zoltan Kövecses explains 'Our experiences with the physical world serve as a natural and logical foundation for the comprehension of more abstract domains' (Kövecses, 2010). Such claims express a version of the embedded embodiment hypothesis¹¹ (outlined above) which has multiple interpretations throughout cognitive science and indeed within the field of Cognitive Linguistics (Rohrer, 2010). In CMT itself, 'embodiment' has been cashed out in three separate ways, as is noted by Kövecses:

As Lakoff and Johnson observe, we have three ways in which simple, or primary, metaphors are embodied: (1) as we just saw, the correlations are embodied in our neuroanatomy; (2) the source domains arise from the experiences of the human body; and (3) we repeatedly experience in the world situations in which source and target domains are connected. (Kövecses, 2010, p. 244)

In this case, the link between VERTICAL ORIENTATION and CLEANLINESS can be said to be embodied in the third sense mentioned above; there is a repeated experiential link between the two domains as the earth and soil beneath our feet is dirty and dust naturally gathers in physically low places. It is, then, in virtue of there being a constantly experienced link between the two domains that a conceptual link between VERTICAL ORIENTATION and CLEANLINESS is constantly reinforced, making it particularly salient and embodied.

To recapitulate, we have seen that MORALITY is mapped with, and subordinate to, both VERTICAL ORIENTATION *and* CLEANLINESS meaning that if MORALITY is the target domain, both VERTICAL ORIENTATION and CLEANLINESS are necessary source domains. I have also claimed that in order to account for the possibility of a single target being structured by

¹¹ Lawrence Shapiro gives an excellent critical overview of the embodiment hypothesis and how it is used in various areas of cognitive science in *Embodied Cognition* (Shapiro, 2011).

multiple source domains, the CMT theorist would need to hypothesize a link between all source domains of a given target. In order to show how this might be shown within the framework of CMT, I have briefly analyzed VERTICAL ORIENTATION and CLEANLINESS in respect to each other which has made clear a link between these two domains showing how CLEANLINESS is conceived of in terms of VERTICALITY. So, what about the other source domains which we have found to be structuring MORALITY?

4.1.14 A One-Way Relationship

With this initial rough analysis of VERTICAL ORIENTATION, CLEANLINESS and MORALITY, I have started to sketch out a hypothesized conceptualization process which could be called *conceptual layering*. In order to give a more logical and conceptually precise understanding of the conceptual relations that hold between all domains which structure a single target, we can start by introducing some organizational principles.

Firstly, all source domains must be *superordinate* to the target; that is, they are conceptually prior to, and, therefore, give structure to, the target domain. Each superordinate domain brings its unique conceptual material to the target domain. We can state this formally as follows:

The Principle of Superordinacy (PS):

For a domain x to be superordinate to y , x must have y within its scope, such that if x is a superordinate domain to y , it will be possible to understand y in terms of x , but not *vice versa*.

This principle restates fundamental assumptions already made by CMT (directionality) and it should be seen as applying to *all* divergent source domains to MORALITY. That is to say, if a target can be understood in terms of more than one source domain, all those source domains must also be conceptually compatible with each other in the way stated by PS.

It could be proposed, then, that all source domains of a particular target sit together as constituents of a well-formed hierarchical structure, which could be called a *multi-layer mapping*, and that the order of constituent domains in such a multi-layer mapping would be determined by PS. We can explicate this as it relates to our current

example as follows: the abstract source domain *MORALITY* is mapped with (understandable in terms of both) *VERTICAL ORIENTATION* and *CLEANLINESS*, as we have already established. It has further been shown that *CLEANLINESS* is also plausibly mapped with *VERTICAL ORIENTATION* in such a way as to suggest that *VERTICAL ORIENTATION* is a superordinate (conceptually prior) to both *CLEANLINESS* and *MORALITY*. *VERTICAL ORIENTATION*, *CLEANLINESS* and *MORALITY* are, therefore, linked together in a logically structured way forming a multi-layer mapping as follows:

Multi-layer Domain Mapping

VERTICAL ORIENTATION → *CLEANLINESS* → *MORALITY*

Corresponding Concepts Within Domains

UP → CLEAN → RIGHT (Positive)

DOWN → DIRTY → WRONG (Negative)

The most concrete conceptual domains are not mapped onto any lower-level domains; they can be understood purely without the need for conceptual structuring from a further superordinate domain. In my above notation, then, the left-most concept *VERTICAL ORIENTATION* in the mapping denotes a maximally concrete, or maximally superordinate source domain and is the first level of this three-domain mapping. The subsequent levels can be read off from left to right and the final concept represents the target domain. So, in the above multi-level mapping, *VERTICALITY* is the first-level domain, *CLEANLINESS* is the second-level domain and *MORALITY* is the target.

All constituent domains of a given multi-layer mapping will, therefore, be a subordinate of all higher domains and a superordinate of all lower domains in a given mapping. We thus expect to see a well-formed hierarchy of conceptual domains structuring a target at graded levels of abstractness.

Table 1 below, shows this hierarchy in relation to our ongoing example which makes explicit the mappings between domains and gives example concepts from

each domain. Hence, we can understand the multi-layer mapping of MORALITY, thus far, as shown below:

Table 1: Multi-layer Mapping of MORALITY (Three Levels)

Level	Domain	Concept	Metaphorical Mapping
Source	VERT. ORIENT.	UP	(UP)
Second	CLEANLINESS	CLEAN	(CLEAN IS UP)
Target	MORALITY	RIGHT	(RIGHT IS CLEAN+UP)

Notice that in the *RIGHT IS CLEAN* metaphorical mapping, *UP* is implicitly preserved (although it is not explicitly apparent on the surface) insofar as *CLEAN* is already structured by *UP*. There is, then, a *conceptual-semantic permeation* from each constituent domain retained through every one of its subordinate domains which goes through to the target. The target, thus, is necessarily structured by all divergent source domains.

4.1.15 Directionality

As outlined above, the relation between layers in a multi-layer mapping is asymmetrical; concepts in subordinate domains can draw conceptual structure from any superordinate domain in a particular mapping, but not *vice versa*. The hierarchy, therefore, preserves what is referred to as *directionality* in the CMT literature. The phenomenon of directionality was observed by Lakoff and Johnson in their original work on metaphor (1980):

First, we have suggested that there is directionality in metaphor, that is, we understand one concept in terms of another. Specifically, we tend to structure the less concrete and inherently vaguer concepts (like those for emotions) in terms of more concrete concepts which are more clearly delineated in our experience. (Lakoff and Johnson, 1980)

To illustrate this with concepts from our current example, this means that RIGHT can be understood in terms of UP or CLEAN, however CLEAN and UP cannot be understood in terms of RIGHT as it would represent a reverse mapping from an abstract subordinate domain to a more concrete superordinate one. Thus, conceptual mapping is unidirectional and represents an asymmetrical relationship.

We can understand directionality as a governing principle to guide our analysis of abstract concepts. The principle of directionality allows us to (1) analyse the degree of abstraction of various concepts within a certain multi-layered mapping, (2) identify constituent domains of a given multi-layered mapping, and (3) expose the hierarchy of a certain multi-layer mapping. In order to do this, we can consider the following question which tests the relation held between two domains x and y :

Test 1: Is it possible to understand x in terms of y ?

If the answer is affirmative, then x is the subordinate (and more abstract) concept; if the answer is negative, y is the subordinate concept – or they have no relevant relationship.

Applying this to our current example, we can ask: is it possible to understand *MORALITY* in terms of *CLEANLINESS*? – *yes*. Is it possible to understand *MORALITY* in terms of *VERTICAL ORIENTATION*? – *yes*. *MORALITY* must, therefore, be a subordinate concept to both *CLEANLINESS* and *VERTICAL ORIENTATION*. We can double check this inference in reverse: Is it possible to understand *VERTICAL ORIENTATION* in terms of *MORALITY*? – *no*. Is it possible to understand *CLEANLINESS* in terms of *MORALITY*? – *no*. Hence, we can confirm that *MORALITY* is subordinate to both *VERTICAL ORIENTATION* and *CLEANLINESS*. And, if we run the same test with *VERTICAL ORIENTATION* and *CLEANLINESS*, we will find that *CLEANLINESS* is the subordinate concept of *VERTICAL ORIENTATION*, as already noted above. This test can also be used to show a maximally concrete or base level domain, as it will be impossible to think of such a domain in terms of any other. Hence, we have a conceptual hierarchy as shown in Figure 1.

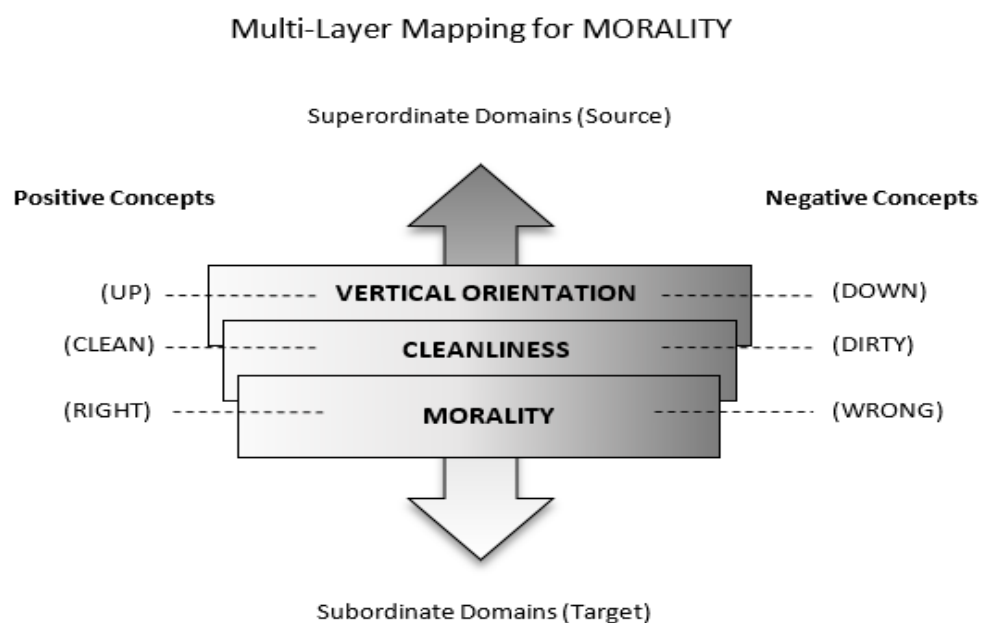


FIGURE 1 Three-level Multi-layer Mapping for MORALITY

4.2.16 Constituents of Multi-Level Mappings

I have shown how the conceptual domain of *MORALITY* is understood in terms of two superordinate mappings of *VERTICAL ORIENTATION* and *CLEANLINESS*. Thus, we have a three-layer mapping of *MORALITY* → *CLEANLINESS* → *VERTICAL ORIENTATION*. But, does this work with all

domains found to be structuring MORALITY? In order for the current proposed analysis to work for CMT, all apparent source domains of MORALITY exposed in the analysis above (plus other which are yet to be found) would need to be proper constituents of the multi-layer mapping. Constituents can be characterized in the following manner:

The principle of Constituency (PC)

A domain x is a constituent of any multi-layer mapping if and only if x bears a relationship to every other constituent domain in the mapping such that it is either superordinate or subordinate of that domain.

In line with the above principles (PS & PC), it will be possible to use Test 1 to look at other putative constituents of the MORALITY mapping and determine their order within the multi-layer mapping.

4.2.17 Doing Good

Before running the proposed test on the source domains already uncovered above, let us look, first, at the domain MERIT from which general terms of approval and disapproval, GOOD and BAD, emanate. This constituent is so obvious as to be trivial, but MERIT is indeed a constituent domain of, and not merely the same as, MORALITY, as MERIT has a much wider application which extends to things outside of the moral domain, such as 'good food' or 'bad idea'.

Following the above principles, an analysis shows MERIT to be a superordinate domain to both MORALITY and CLEANLINESS, but not to VERTICAL ORIENTATION. It is, however, subordinate to, and understandable in terms of, VERTICAL ORIENTATION. MERIT is vertically oriented with GOOD as UP and BAD as DOWN. This metaphor has been known for some time (Lakoff & Johnson, 1980) and has gathered much empirical support in recent years (Gottwald et al., 2015) supporting the hypothesis that the link to also be a well-entrenched and embodied one. Hence, it can be said to be a constituent of the multi-level mapping MORALITY and sits as the second layer in the mapping: VERTICAL ORIENTATION → MERIT → CLEANLINESS → MORALITY. This analysis can be run on all other putative constituent source domains which were found through the prior CMT-based analysis above.

When analysing the next contender, BRIGHTNESS, it becomes apparent that it is indeed a constituent and sits within the hierarchy as follows: VERTICAL ORIENTATION → BRIGHTNESS → MERIT → CLEANLINESS → MORALITY. BRIGHTNESS is a subordinate domain to VERTICAL ORIENTATION as LIGHT and DARK can be understood in terms of UP and DOWN, respectively, and this is another very salient link due to its experiential grounding in the natural world. The sky above is, of course, our main source of natural light and all earthly organisms have evolved in this environment where light comes from above. Our bodies have been moulded by evolution, adapting to deal with such physical phenomena. Our brow shields us from overhead light and does not offer the same protection from low light sources, for example. This link is also observed in expressions such as ‘the darkness *fell*’, ‘we waited for *nightfall*’ and ‘the sky has *brightened up*’.

The next domain in the mapping, MERIT, is a subordinate to BRIGHTNESS: GOOD IS LIGHT and BAD IS DARK (‘a *bright future/dark past*’, ‘*light* at the end of the tunnel’, ‘his reputation was *tarnished*’). BRIGHTNESS is also a superordinate of CLEANLINESS as is evidenced by such phrases as ‘it was clean *and shiny*’ and ‘his feet were *black!*’ such a link is also made salient and reinforced through experience. Thus, we find BRIGHTNESS situated in the mapping as follows: VERTICAL ORIENTATION → BRIGHTNESS → MERIT → CLEANLINESS → MORALITY.

When coming to the domain of FAIR TRANSACTION, it is difficult to see how it fits into the mapping. This may be to do with the fact that this metaphor, as already noted, is formulated at its most schematic and, therefore, includes many varied interactions. In order to make this domain ‘fit’ it could be construed as QUANTITY. This can be observed in the phrases ‘morally *bankrupt*’ and ‘moral *values*’. QUANTITY is understood in terms of VERTICAL ORIENTATION and is a subordinate to BRIGHTNESS, as BRIGHTNESS is understood in terms of abundance of, or lack of, light. After reflection, it becomes clear that each other constituent domain of MORALITY also relates to QUANTITY. We find that QUANTITY is subordinate to VERTICAL ORIENTATION with MORE as UP and LESS as DOWN and superordinate to BRIGHTNESS, MERIT, CLEANLINESS, and MORALITY, as shown in Table 2.

Table 2 Multi-layer Mapping of MORALITY (Six Levels with Positive Concepts)

Level	Domain	Concept	Metaphorical Mapping
Source	VERT. ORIEN.	UP	(UP)
Second	QUANTITY	MORE	(UP IS MORE)
Third	BRIGHTNESS	LIGHT	(LIGHT IS MORE+UP)
Fourth	MERIT	GOOD	(GOOD IS LIGHT+MORE+UP)
Fifth	CLEANLINESS	CLEAN	(CLEAN IS GOOD+ LIGHT+MORE+UP)

Target	MORALITY	RIGHT	(RIGHT IS CLEAN+GOOD+LIGHT+MORE+UP)
--------	----------	-------	--------------------------------------------

As above, we can apply our analysis to the domain of *EMOTION* in order to find its place within the mapping for *MORALITY*. It is, as already noted, superordinate of *MORALITY* as it is clearly evident in linguistic metaphors. *CLEANLINESS*, is understood in terms of *EMOTION* as with ‘the floor was *disgusting*’ making *EMOTION* a superordinate of *CLEANLINESS*. *EMOTION* is a subordinate to *MERIT* and *VERTICAL ORIENTATION* as it can be understood in terms of *MERIT* as is evidenced by ‘I feel *good/bad*’, and *VERTICAL ORIENTATION* as shown by ‘feeling *up/down*’, meaning that *EMOTION* sits between *CLEANLINESS* and *MERIT* in the mapping, as shown in Table 3 and Table 4 which give multi-layer mappings and examples of both positive and negative concepts.

Table 3 Multi-layer Mapping of *MORALITY* (Eight Levels with Positive Concepts)

Level	Domain	Concept	Metaphorical Mapping
Source	VERT.	UP	(UP)
Second	QUANTITY	MORE	(UP IS MORE)
Third	BRIGHTNESS	LIGHT	(LIGHT IS MORE+UP)
Fourth	MERIT	GOOD	(GOOD IS LIGHT+MORE+UP)
Fifth	EMOTION	COMPASSION	(COMPASSION IS GOOD+ LIGHT+MORE+UP)

Sixth	CLEANLINESS	CLEAN	(CLEAN IS COMPASSION+GOOD+LIGHT+MORE+UP)
Target	MORALITY	RIGHT	(RIGHT IS CLEAN+COMPASSION+GOOD+LIGHT+MORE+UP)
Next Level	HOLINESS	HOLY	(HOLY IS RIGHT+CLEAN+COMPASSION+GOOD+LIGHT+MORE+UP)

Table 4 Multi-layer Mapping of MORALITY (Eight Levels with Negative Concepts)

Level	Domain	Concept	Metaphorical Mapping
Source	VERT.	DOWN	(DOWN)
Second	QUANTITY	LESS	(DOWN IS LESS)
Third	BRIGHTNESS	DARK	(DARK IS LESS+DOWN)
Fourth	MERIT	BAD	(BAD IS DARK+LESS+DOWN)
Fifth	EMOTION	HATE	(DISGUST IS BAD+DARK+LESS+DOWN)

Sixth	CLEANLINESS	FILTH	(DIRTY IS DISGUST+BAD+DARK+LESS+DOWN)
Target	MORALITY	WRONG	(WRONG IS DIRTY+DISGUST+BAD+DARK+LESS+DOWN)
Next Level	HOLINESS	EVIL	(EVIL IS WRONG+DIRTY+DISGUST+BAD+DARK+LESS+DOWN)

Tables 3 and 4 show the complete multi-layer mapping for MORALITY with example concepts. The first row of each table shows the maximally concrete concept in the mapping VERTICALITY which is a superordinate to all other domains in the mapping. The domains in the proceeding rows appear in order of abstractness with EVIL being the most abstract. In the last column I have given metaphorical mappings with example concepts from each domain.

Viewed like this, we can clearly observe the unidirectionality of metaphor, and the layering of concepts as domains become more abstract towards the bottom of the table. The concept of RIGHT, then, carries with it the combined conceptual material of all superordinate concepts and, in being so constituted, is a concept with a high level of abstraction – the combined influence of the superordinate domains could be claimed as providing all the senses in which the concept of morally RIGHT can be construed as each domain provides conceptual-semantic material to the next subordinate domain while also preserving the conceptual material from all previous domains, insofar as each concept is necessarily structured by all superordinate concepts.

The reader will also notice that I have included **HOLINESS** as a further domain (tables 3 and 4). This, I believe, is an uncontroversial addition to the mapping and is used to illustrate how **MORALITY** itself could function as a superordinate domain to further concepts. In accordance with the method of analysis proposed here, **HOLINESS** can be tested to show how it is a subordinate domain to all concepts in the mapping, allowing it to be part of the mapping.

It appears from the preceding analysis that **MORALITY** is a conceptual domain with a relatively high level of abstractness inasmuch as it is necessarily understood in terms of several divergent source domains. Therefore, the conceptual structure of **MORALITY** can be understood as **VERTICALITY** → **QUANTITY** → **BRIGHTNESS** → **MERIT** → **EMOTION** → **CLEANLINESS** → **MORALITY** as shown in Figure 2.

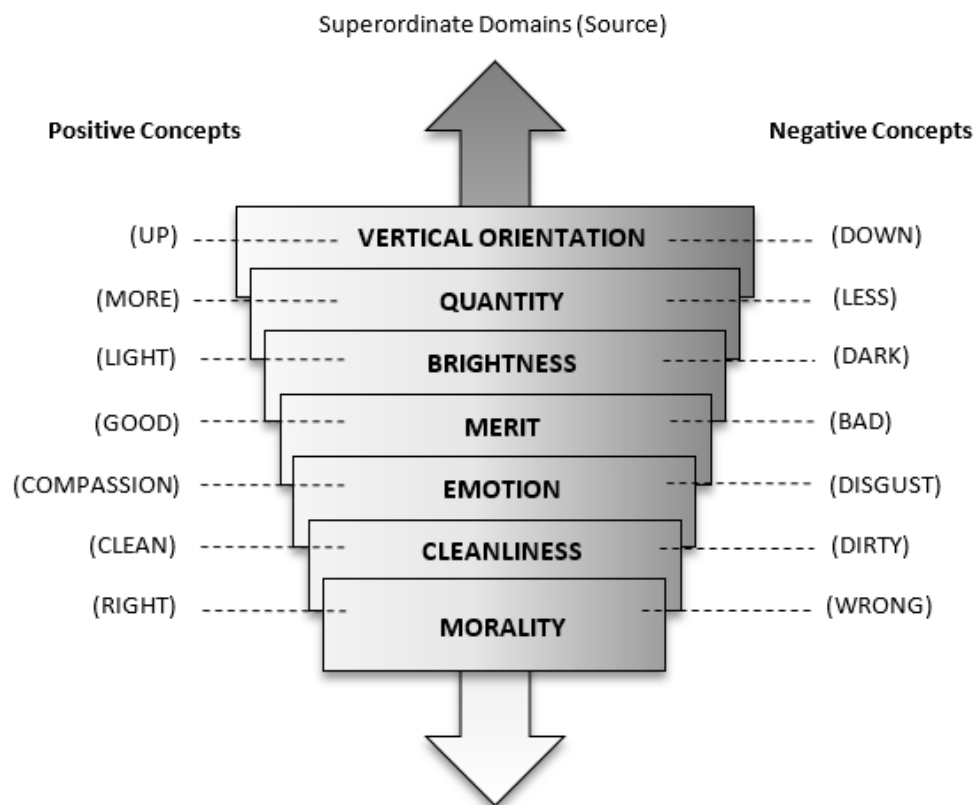


FIGURE 2 Seven-level Multi-layer Mapping for MORALITY

4.2.18 Summary

In the preceding paragraphs, I have proposed and outlined a theory and associated method of analysis which works within the framework of Conceptual Metaphor Theory and aims at explicating why a given target domain might be observed as being structured by multiple, seemingly unrelated, source domains. I have shown that the multiple source domains which structure a single target domain, might all be conceptually linked. I have given an analysis of conceptual domain layering which makes manifest the multi-layer mappings of source domains beneath a given target, thus showing how divergent constituent domains might be arranged hierarchically in multi-layer mappings of the target. Such a structure would need to be logically organized such that, constituent domains satisfied the relevant conditions mentioned above in order to be proper constituents of a given mapping.

4.3 The Form of Moral Concepts

It should be evident from sections 4.1 and 4.2 that there is always a clear and perceptible contrast between *RIGHT* and *WRONG*. That is to say, the concepts *RIGHT* and *WRONG* appear to be binary opposites. This opposition, of course, is of fundamental importance to our understanding of the domain *MORALITY*, from which they emanate, for they act as the compass points by which we navigate the moral landscape. Thus far, however, I have concentrated on the structure of moral concepts which has meant largely treating *RIGHT* and *WRONG* as similar, but in drawing attention to their contrast, we might legitimately ask: What is the nature of this contrasting relationship held between *RIGHT* and *WRONG*?

It is assumed, here, that in describing the same conceptual domain, *RIGHT* and *WRONG* stand in a relation of binary opposition to each other and that the form, or character, of this relation is part of their mental representation. That is to say, part of the mental representation of the concept *RIGHT* is that it stands in a particular relation to the concept *WRONG* and it is the way, or form, in which this relation is mentally represented that is the subject matter of the foregoing investigation. This section, therefore, runs an analysis

which aims to examine how the binary opposite relationship between *RIGHT* and *WRONG* is mentally represented. To this aim, I pose the question:

What is the form of binary opposition held between the moral concepts *RIGHT* and *WRONG*?

I aim to answer this question by running an analysis of moral language which focusses on the use of the antonym pair *right/wrong*. This method is grounded in the assumption stated at the beginning of this thesis: that lexical items encode conceptual form. It follows, if this assumption is true, that the lexical form a given word exhibits will be reflective of the nature of the concept to which the word refers. The result of the analysis shows how *RIGHT* and *WRONG* can be represented in two distinct ways viz. as *mutually exclusive*, or *gradable*.

It is assumed, then, as stated above, that there is a tight connection between concepts and language; such that:

The form in which concepts are mentally represented will determine, to some extent, the lexical and grammatical form of language.

Such a view, as Jones et al. (Steven. Jones et al., 2012) note, is foundational to the field of Cognitive Linguistics. Such frameworks are based on the general assumption that:

[M]eaning is grounded in how we as humans both perceive and understand the world around us. Meanings

of lexical items are dynamic and sensitive to contextual demands, rather than being fixed and stable, and

lexical items evoke meanings rather than have meanings. (Jones, et al, 2012)

A clear example of how this assumption can ground a particular study can be seen in the work of Wisniewski and colleagues (Wisniewski et al., 2003), who, based on this assumption, claim that 'the grammatical distinction between count and mass nouns is systematically related to a conceptual distinction between the referents of count and mass nouns' (Wisniewski et al., 2003). In their investigation into the conceptual basis of count and mass nouns, they support the *cognitive individuation hypothesis* which states

that 'whether a person uses a count or a mass noun to refer to some aspect of reality depends on whether they interpret the referent as an individual or as a non-individuated entity' (Wisniewski et al., 2003; Wisniewska et al., 1996) . Although Wisniewski's work focusses on noun forms, it makes the same guiding assumption that linguistic form is determined by conceptual form and, indeed, there is also empirical support for the conceptual nature of adjectives (Charles et al., 1994; Steven. Jones et al., 2012; Murphy & Andrew, 1993) .

Thus, I claim, in line with the above, that whether a speaker expresses a particular kind of antonymic relation in referring to some aspect of reality depends on whether they mentally represent the referents as mutually exclusive or gradable in nature. The foregoing work, therefore, makes two foundational claims. The general:

- (i) Lexical items, insofar as they semantically encode the meaning of a concept, reflect its conceptual form in their linguistic functioning.

And, the specific:

- (ii) Antonymic form reflects the conceptual form of binary opposite concepts.

These assumptions, underpin the approach taken in the current work and it follows from (i) and (ii) that an analysis of the antonymic relation between moral adjectives such as *right/wrong*, *moral/immoral* and *ethical/unethical* will allow us to probe the conceptual nature of the corresponding moral concepts RIGHT and WRONG, in virtue of the fact that the former is a function of the latter.

Before presenting the analysis of moral adjectives, I will first give a brief overview of previous work on antonymy and binary opposition, which acts as a historical and conceptual background to the following study.

4.3.1 Lexical Relations

Much analysis of antonymy as a lexical relation has hitherto been undertaken, most notably by Lyons (1977) and Cruse (Cruse, 1986) who have

endeavoured to construct taxonomies of antonym types. Antonym types are varied and include 'gradable contrary antonyms', 'complementary antonyms' and 'directional opposition', plus additional subtypes such as 'converse oppositional' and 'independent reversive antonyms'¹².

As well as documenting the various forms of antonymy, another foundational aim of those working in lexical semantics has been to reach a precise definition of what constitutes antonymy. Lynne Murphy (2003) points out that antonymy 'is arguably the archetypical lexical semantic relation' and further notes:

If antonymy is the most robust of the semantic relations, one might argue that it should be the most specifically defined. However, the relevant literature lacks evidence for such specifics. (Murphy, 2003)

Although there is disagreement about how antonymy should be precisely defined, the fruits of these endeavours provide a foundation upon which to build the foregoing study.

Much work on antonymy asserts the importance of the lexico-semantic relations between words or lexical units (Lyons, 1977), however some linguists have stated that the antonymic relation is a purely lexical relation between word forms as opposed to a semantic relation between meanings of words (Gross et al., 1989; Miller et al., 1990).

There are strong and obvious counter examples which threaten the soundness of such a claim. Take, for example, *hot/cold* which are normally considered to be gradable antonyms (they exhibit a scalar relation). However, it is apparent that the pair *hot/cold* can, in fact, be used as complementary antonyms (which exhibit a binary either/or relation) in certain contexts. An instance of this can be observed when referring to hot and cold taps in a sink. In this context, we might ask the question 'is the hot or the cold tap dripping?' where *hot/cold* are clearly represented as complementary antonyms. Such observations seem to point to the context-dependence, and possible conceptually-

¹² See Cruse (1986) and Lyons (1977) for detailed explanations of various antonym types.

determined nature of antonym pairs, as well as the need for a more flexible interpretation of antonymy which accounts for such non-standard uses.

I claim it is possible to understand the flexibility and context-dependence of antonymic form by analysing the underlying concepts to which the antonyms refer. The current work, then, will try to develop upon, and move beyond, previous analyses by exploring and bridging the gap between lexical relations and human cognitive processes. The aim is to understand how various forms of antonymy manifest in language as an *upshot* of more general cognitive processes, and to show how this is evident in the use of moral antonyms *right/wrong*.

4.3.2 The Conceptual Nature of Antonymy

The current work makes the foundational claim: that lexical relations arise as a consequence of human conceptualization processes; such that, conceptual form determines linguistic form. A cognitive view such as this can be adopted in order to describe the differences in lexical categories and the relationships between words such as antonyms, synonyms, hyponyms etc. as being a function of human psychology; that, as has been mentioned above, the meanings and relations between lexical elements are a function of the nature of the conceptual representations which those lexical items connect with.

Carita Paradis' Lexical items as Ontologies and Construals (LOC) framework is a cognitive account of antonymy – or opposition. It describes the cognitive processes involved in the construal of lexical relations. The cognitive process, Paradis holds, starts with *Ontologies* which form the basic conceptual material from which linguistic meaning is made. Ontologies are of two kinds: *Content* and *Configurational structures*. Content structures consist of conceptual dimensions such as MERIT, and properties within those dimensions such as GOOD and BEAUTIFUL, for instance. These content structures are then formatted by the configurational structures such as DEGREE OR FREQUENCY depending on the speaker's communicational intent. These Ontologies are subsequently worked on by *construal processes* which are responsible for generating the final linguistic

form. These construal processes are of four main kinds: *Gestalt*, *Saliency*, *Comparison* and *Perspective*.

Through LOC, antonymy is comprised of a content Ontology where a single dimension e.g. *MERIT* is divided into two opposing antonymic parts – this is part of our natural conceptualization of a domain. This content can be further formatted by a configurational structure of *DEGREE* which can then be processed by the construal of *comparison*, giving rise to gradable antonymy in language. This framework has been applied by other linguists in the analysis of antonyms (Jones et al., 2012) and indeed seems effective in the analysis of adjectives which describe the kinds of physical properties that can be clearly perceived through the senses such as *hot/cold* or *on/off*, as detailed above.

As the LOC model shows, it is possible to reach an understanding of antonymy as arising from, and determined by, the way in which human perceptual systems encode sensory data and build concepts (Acquaviva et al., 2020; Steven. Jones et al., 2012; Murphy & Andrew, 1993; Paradis & Willners, 2011) . So, where does this leave antonyms which describe non-physical or abstract conceptual domains such as *MORALITY*? In order to answer this question, an analysis of the moral antonyms *right/wrong* is subsequently undertaken and it will become evident that *right/wrong* display an antonymic relation which is vague and inconsistent; exhibiting great flexibility throughout moral discourse.

This vagueness, I argue, can be explained, and is, in fact, to be expected from the current cognitive perspective, as moral adjectives refer to abstract conceptual properties which give rise to antonymic vagueness. It appears that, unlike antonyms referring to concrete phenomena such as *on/off* and *hot/cold*, the lexical status of moral antonym pairs is 'vague' or difficult to discern. A variety of contexts show that the antonymous pair *right/wrong* can be used and understood in various and non-uniform ways in moral utterances. The following study will show that they can be observed displaying inconsistent forms of antonymy which depend upon both the context of the utterance and the pre-existing moral convictions of the speaker.

4.3.3 Antonymic Form

In the present work, I focus on two divergent forms of antonymy (i) gradable contrary antonyms, which are assumed to reflect gradable concepts and (ii) complementary antonyms which are assumed to reflect mutually exclusive concepts. I will explicate the nature of these conceptual relations in turn.

4.3.4 Gradable Antonyms and Gradable Concepts

Gradable contrary antonyms can be identified by the logical relation held between a pair of words that are binary opposites. The logical entailments of gradable antonyms can be shown formally as follows:

$$a \Rightarrow \neg b$$

$$\neg a \not\Rightarrow b$$

In gradable contraries such as *hot/cold*, for instance, the assertion of any member of the pair entails the negation of the other member. For example, 'the coffee is hot' entails that the coffee is not cold. However, the negation of one member does not entail the assertion of the other member¹³. The members of a pair of gradable contrary antonyms each represent a point along a variable scale with an indefinite number of intermediate points. Therefore, negating one of the pair rules out the negated term only, leaving many other points on the scale available, including its opposite – the other member of the antonym pair.

The nature of the antonymic relation which holds between a given antonym pair, I claim, is determined by the way in which a concept is mentally represented. The concepts reflected in gradable contrary adjectives are, therefore, assumed to also be gradable in nature. In order to demonstrate this conceptual-lexical relation, let us take the example of the pair *hot/cold*, which exhibit a gradable contrary relation, picking out two opposed ends of a gradable temperature scale. Their status as

¹³ Although it has been argued that this may be pragmatically inferred (Brewer & Lichtenstein, 1975).

gradable contrary antonyms is apt for describing the scalar nature of the TEMPERATURE domain to which they apply. To give an everyday example, when we turn on the shower in the morning, we might first turn on the hot tap and run our hands under the stream of water to test the temperature. We might find, after a minute or two, that the water has become too hot and then proceed to turn on the cold tap too, in order to make the temperature more bearable. We put our hand under the stream of water again, slowly turning on the cold tap until the temperature is just right before stepping into the shower. In the case of temperature – a sensible phenomenon – we are able to gain rich knowledge of the TEMPERATURE domain with our somatosensory system, in the manner just described, which is encoded into the concepts HOT and COLD. Due to the fact that we can sense the graded nature of TEMPERATURE, HOT and COLD are, therefore, mentally represented as gradable in nature. It is congruous, then, that the lexical items which express these concepts in language – in this example the antonymous adjectives *hot/cold* – are construed as gradable contrary antonyms.

4.3.5 Complementary Antonyms and Mutually Exclusive Concepts

In contrast to the gradable contraries mentioned above, the logical relationship between members in a pair of complementary antonyms is that of the exclusive disjunction ($a \oplus b$) and includes such examples as *on/off*, *heads/tails* and *connected/disconnected*. The entailments for complementary antonyms can be shown formally as follows:

$$a \Rightarrow \neg b$$

$$\neg a \Rightarrow b$$

The assertion of one member of the pair entails the negation of its opposite and this is a symmetrical relation. For example, ‘the circuit is *on*’ entails that it is *not off* and ‘the switch is *not on*’ entails that it *is off*. Put simply, there is a binary ‘either-or’ relation between the members of a complementary pair such that there can be no intermediate points or degrees between them. The concepts which complementary antonyms refer to are, therefore, assumed to be mentally represented as mutually exclusive. For instance,

the antonym pair *on/off* mirrors the conceptual nature of the concepts ON and OFF – circuits must be, by conceptual (and physical) necessity, either on or off. In the same manner as the gradable contrary antonyms outlined above, the perceived physical properties of the denotata are encoded into concepts and, insofar as words encode concepts, the conceptual form determines the antonymic form of the lexical items.

4.3.6 Conceptual Representation and Grammar

It is elementary that the form of antonymy which a given pair of adjectives exhibits will determine the morphological form and the grammar of a speaker's utterance. Continuing with the above case, for instance, it is considered ungrammatical to describe a circuit as '*nearly on*', or say that '*circuit X is the on-est*' or '*circuit X is off-er than circuit Y*' as the adjectives *on/off* are complementary antonyms, meaning they cannot be used with the superlative or comparative grammar. Conversely, the superlative or comparative forms *can* be used with *hot/cold* such as '*X is the hottest*' and '*X is colder than Y*'. Such usage, from the current perspective, is not simply the outcome of a learned or conventional grammatical rule, but rather a conceptual necessity. Speakers are prevented from using the superlative or comparative forms with *on/off* due to the mutually exclusive nature of the concepts which they encode. Hence, the kind of antonymy held between a given pair of antonymous adjectives, and therefore also the grammatical form of utterances in which they appear, are determined by the way in which a speaker mentally represents the concepts expressed by the adjectives.¹⁴

4.3.7 Grammatical Analysis

Thus far, I have explained the assumption that antonymous relations between lexical items are determined by the way in which a concept is mentally represented. Based on this assumption, I propose that an analysis of the way in which speakers use moral antonyms, and moral language in general, will be instrumental in uncovering how the concepts RIGHT and WRONG are represented in the minds of agents. I will, then, proceed with an analysis of the antonym pair *right/wrong*¹⁵ in order to uncover

¹⁴ See (Henderson, 2021) for a similar approach with truth predicates.

¹⁵ Although I will limit my analysis to the adjectives *right/wrong*, it is assumed that all moral predicates will function in the same way as they are all assumed to reflect the same concepts RIGHT and WRONG.

the conceptual structure of the corresponding concepts RIGHT and WRONG. The first step in this process is to ascertain which form of antonymy *right/wrong* exhibit by embedding them in various grammatical constructions – some suitable for gradable antonyms, and others suitable for complementary antonyms – in order to see if grammatical and conceptually congruent sentences are generated.

(56) X is right/wrong

A grammatical and conceptually congruent sentence is generated, which would be expected here regardless of the antonymic relation.

(57) #X is right-er/wrong-er than Y.

The comparative form is ungrammatical. A grammatically correct paraphrase of this would be 'X is worse than Y' (or 'X is better than Y') in which 'worse' and 'best' are used to express moral sentiment and are conceptually congruous. This, therefore, points at the comparative grammatical form, which morphologically modifies the adjectives 'right' or 'wrong' with the suffix 'er', being unacceptable, but not the concept of *degrees of* rightness or wrongness. Hence, on the surface, (65) points to *right/wrong* as having an antonymic relation that is complementary, whereas at a conceptual level a gradable nature can be understood – as seen with the paraphrase.

(58) X is not as right/wrong as Y.

This sentence is not strikingly ungrammatical, although it seems awkward in English, and so might suggest complementary antonymy at the sentence level, but this is not obviously the case. Conceptually, (66) is ambiguous.

(59) X is neither right nor wrong.

This sentence is grammatical and so points to gradable antonymy. However, it is unclear if this is congruous at the conceptual level. The same is also true with the following:

(60) 'X is *either* right or wrong'

(68) is grammatical, suggesting complementary antonymy at the sentence level, but is, again, ambiguous at the conceptual level.

(61) #X is *the wrong-est*.

The superlative form generates an ungrammatical sentence which suggests complementarity. Here, as with (65), this sentence could be paraphrased as 'X is *the worst...*' in order to generate a grammatical sentence. Hence, it appears, as with (65), that *right/wrong* are used as complementary antonyms which resist morphological modification. On a conceptual level it seems that a gradable understanding is possible.

(62) X is *completely* wrong.

The totality modifier generates a grammatical sentence which suggests bounded graded antonymy¹⁶, but not unbounded gradable antonymy or complementarity. Intuitively, this form is conceptually congruous.

(63) X is *not completely* wrong.

¹⁶ Some gradable antonyms are bounded, meaning that there is an absolute maximum and/or minimum degree such as *transparent/opaque* and hence can be used in conjunction with totality modifiers like 'completely'. Others are unbounded suggesting that there is no upper limit, such as *hot/cold* and cannot be used with such modifiers (W. Croft & Cruse, 2004; Steven. Jones et al., 2012).

A grammatical sentence is generated when negating the totality modifier 'completely'. This does not work with complementary antonyms, as it suggests a gradable, but bounded, scale. At the conceptual level, this points toward *RIGHT* and *WRONG* as having a gradable representation.

4.3.8 Interpretations

From this brief grammatical analysis, we can start to build a picture of the unclear or inconsistent antonymic relation held between the moral adjectives *right/wrong*. Although we can discern that *right/wrong* resist morphological modification at the sentence level, (65) and (69), they can be used in grammatical frames which suggest gradeability, (67), (70) and (71), as well as complementary antonymy; (64) and (68). At the conceptual level, it is unclear how *RIGHT* and *WRONG* are represented from this analysis, it appears that in some cases we are able to conceive of them as gradable, and in others, as mutually exclusive. It is also important to note that grammaticality and conceptual congruity do not always correlate. This can be seen clearly in (65) where the sentence is ungrammatical, but conceiving of something as being worse than (*wrong-er*) is not incongruous. It can be concluded, therefore, that from the examples (64-71) above, we observe grammatical and conceptual ambiguity with divergences between grammatical and conceptual form.

4.3.9 Context and Conceptual Form

The above analysis is enlightening insofar as it hints at the possibility of *right/wrong* not having a fixed antonymic relation. However, it does not show a clear relation between linguistic and conceptual form. This is indeed a problem for the current theory because if the assumption that conceptual form determines linguistic form is correct, we would expect to see a clear convergence between the two; but we don't. This result, however, can be explained. The methodology of the above analysis is flawed, as the grammatical constructions are decontextualized sentences which make use of the non-descript placeholders X and Y. The use of these abstract placeholders to represent moral actions conceals an important insight: that the kind of antonymic relation held between *right/wrong* is determined by the conceptual nature of the states of affairs to

which the predicates are applied. Hence, it would appear that a sufficiently detailed analysis should use contextualised examples.

4.3.10 Mutually Exclusive Concepts

I will now, in light of the above, proceed with an analysis of contextualized moral judgements which will make it perspicuous that the antonymic relation held between *right/wrong*, is contingent upon the context in which they are employed. Let us start by taking the following moral judgement:

(64) Eating meat is *wrong*. (*Held by a vegetarian*)

In (72), 'wrong' is used in a way which suggests the speaker understands *right/wrong* to be complementary antonyms. Intuitively, we understand that 'wrong', is not being used in a gradable way here. This is not understood by simply looking at the grammatical form of the sentence, but is rather inferred from the conceptual nature of the verb 'eating' to which the predicate 'wrong' is applied.¹⁷ The nature of this action is such that either one is eating, or one is not – eating and not eating are mutually exclusive states. It is understood, then, from the judgement in (72) that the action of eating meat is always wrong and, consequently, that whether one eats a large quantity of meat, or only a small amount, the action is always considered to be morally wrong. It is incongruous, therefore, for a vegetarian who sincerely assents to judgement (72), but who has knowingly and wilfully eaten meat, to justify their action or mitigate its severity by stating that they only ate a small amount – the action of eating meat is judged, in this case, to be simply *wrong*, allowing for no degrees of wrongness, because there are no degrees of eating. This example suggests that the antonymous adjectives 'right' and 'wrong' take on a complementary form when predicated of a non-gradable mutually exclusive state of affairs. Hence, an agent who assents to (72) is assumed to mentally represent *RIGHT* and *WRONG* as mutually exclusive concepts. We can now see that it would be incongruous for

¹⁷ This also suggests a certain level of pragmatic inference from background knowledge, in this case knowledge of vegetarianism. There is mounting evidence to suggest that pragmatic information, such as information in the visual field, is incorporated in the online processing of language (Aparicio, 2018).

such an agent to assert that 'eating a small amount of meat is *not as wrong as* eating a lot', or that 'eating a lot of meat is *wrong-er (worse) than* eating a little'. That is to say, the superlative and comparative forms are grammatically *and* conceptually wrong for those who assent to judgement (72).

There is further evidence for the fact that agents who assent to the moral judgement (72) do indeed understand RIGHT and WRONG as mutually exclusive concepts by looking at real-world examples of moral discourse. In the following excerpts, we can assume that the writer assents to (72), as is evidenced from the content of the text. The extra content included in these excerpts, shows clearly that the writers understand RIGHT and WRONG to be mutually exclusive concepts. Take, for instance, the following text taken from PETA.org:

(65) [T]here is no such thing as "humane meat." Giving animals a few more inches of living space is simply not enough—and even if their quality of life is high, we still don't have the right to take that life for something as trivial as a particular meal. [...] There is no humane or ethical way to eat animals.

(Debate Kit: Is It Ethical to Eat Animals? | PETA, n.d.)

It is clear from the passage from PETA (73), expressed in characteristically emotive language, that, in relation to eating meat, RIGHT and WRONG are conceptualized as having a mutually exclusive relation, even though the predicates *right/wrong* aren't used explicitly. Hence, we can conclude that in assenting to the moral judgement (72), RIGHT and WRONG are mentally represented as mutually exclusive. The judgement in (72) is not the only context where moral concepts are understood as mutually exclusive. There are further examples showing this same conceptual form in the following excerpt from a website which answers children's questions about Christianity:

(66) Are things either right or wrong?

Here's the answer:

Yes, the Bible teaches *some things are either right or wrong*.

Sometimes a decision is a matter of which choice you like better. Should I eat chocolate ice cream or vanilla?

Should I play basketball or volleyball? Other choices are either right or wrong. God gives many commands

in the Bible about the things we should do as Christians and things we should not do.

(*Are Things Either Right or Wrong?*, n.d.) <https://www.gqkidz.org/right-or-wrong.html> - emphasis added)

In this excerpt, the predicates *right/wrong* are clearly being used as complementary antonyms, in order to express the mutually exclusive conceptual nature of RIGHT and WRONG as conceptualized by the writer. This is highlighted by the use of the 'either/or' grammatical construction in 'Yes, the Bible teaches *some things are either right or wrong*'. This case differs slightly from the case in (72) in that it is not the physical nature of a specific state of affairs which determines the mutually exclusive conceptual form, but the metaphysics of Christian morality itself – if what is right and wrong are decided by God, the logic of (74) goes, then God's approval makes the action right and God's prohibition makes the action wrong – there is no middle ground. The following examples (75) and (76) also show that moral concepts are understood in a mutually exclusive manner by Christians:

(67) Every situation and every decision we make boils down to one thing – it is either right or wrong.

Sometimes it is difficult for humanity to accept this and we try to justify what we know to be a wrong

decision by claiming what we face was somehow in a grey area in between both right and wrong.

(*It Is Either Right or Wrong; There's No Grey Area | The Daily Walk*, n.d.) [https://www.thedailywalk.org/it-](https://www.thedailywalk.org/it-is-either-right-or-wrong-theres-no-grey-area/)

[is-either-right-or-wrong-theres-no-grey-area/](https://www.thedailywalk.org/it-is-either-right-or-wrong-theres-no-grey-area/))

(68) Where an issue of right and wrong is concerned, there is never any gray. There is always a right and everything else is a wrong

(*Morality Is Black And White: There Is No Gray – THE ROAD TO CONCORD*, n.d.)

<https://theroadtoconcord.com/natural-law/derevation/rights-bubbles-the-origin-of-universal-morality/morality-is-black-and-white-there-is-no-gray/>

Hence, it is evident from the use of the adjectives *right/wrong*, together with the explicit content of the text, that the writers of (73), (74), (75) and (76) understand RIGHT and WRONG as being mutually exclusive concepts.

4.3.11 Gradable Concepts

From the above, we can see that in certain contexts, people mentally represent RIGHT and WRONG as non-gradable, mutually exclusive concepts. It is not always the case, however, that moral concepts are mentally represented in such a manner. Take moral judgement (77) below, for instance:

(69) Minimising the suffering of farm animals is *right*.

(Expressed by a non-vegetarian animal rights campaigner)

It appears that, in contrast to judgement (72), the adjective 'right' in (77) is understood as gradable. This, also, is not inferred from the grammar of the judgement, but by the nature of the state of affairs to which the predicate 'right' is applied i.e. minimising suffering. It appears, then, that the perceived nature of that to which 'right' or 'wrong' is predicated determines how the concepts RIGHT and WRONG are represented in the mind of the agent. Therefore, if the state of affairs to which a moral predicate is applied is understood as gradable, then RIGHT and WRONG will also be mentally represented in such a manner; that is, abstract moral concepts are mapped onto concrete states of affairs.

To explain this further, we could contemplate an example situation which pertains to the judgement in (77). Let us imagine three farms. The first farm ensured the

lowest levels of animal suffering; the second farm moderate levels of suffering; and the third, very high levels of animal suffering. We might imagine that an agent who assents to moral judgement (77), when comparing the farms, would judge the agricultural practices of the first farm to be morally right, the third farm to be wrong, and the second farm to be neither right nor wrong, but 'right-er' (better) than the third. Here, we see a conceptual moral middle-ground where the moral status of the agricultural practices on the second farm are not adequately covered by either RIGHT or WRONG which is conceivable only if moral concepts are not represented as mutually exclusive. Hence, an individual who assents to (77) sees a moral middle-ground, which is not conceptually accessible to the agent who assents to (72).

This example suggests that *right/wrong* can also be used as gradable antonyms when agents mentally represent RIGHT and WRONG as gradable concepts, but we have also seen, from the initial grammatical analysis above, that *right/wrong* resist being morphologically modified with the comparative and superlative suffixes 'er' and 'est'. It remains to be shown, then, that RIGHT and WRONG can indeed be used in a gradable way. This can be shown, as I will proceed to do, and an interesting observation can be made: the linguistic restriction which prevents the morphological modification of *right/wrong*, forces speakers to use alternative lexical items when expressing moral concepts as gradable. We can see this effect happening in a range of contexts; take the following text from Vox.com for instance:

(70) [N]early 400 companies, including Hyatt and Marriott, committed themselves to *better*

conditions for animals. [...] From one perspective, factory farming of animals is one of the few social

problems in the world today that, rather than getting *better*, gets *worse* each year, as we continue to

breed animals in terrible, even monstrous conditions. (*Animal Rights: 2018's Big Wins — and Big*

Losses — for Animals - Vox, n.d.) [https://www.vox.com/2018/12/24/18148698/2018-year-in-](https://www.vox.com/2018/12/24/18148698/2018-year-in-review-for-animals)

[review-for-animals](https://www.vox.com/2018/12/24/18148698/2018-year-in-review-for-animals) - emphasis added)

As pointed out above, in discussion on animal welfare, it is often assumed that the level or degree of suffering those animals endure is correlated with the moral status of the action, meaning that actions which cause more suffering are seen as ‘wronger’. In (78), the writer is using the words ‘better’ and ‘worse’ to describe the moral status of the conditions in which farm animals are kept and this phenomenon is widespread. Speakers regularly turn to phrases such as ‘morally *better*’, ‘morally *worse*’, ‘morally *superior*’ or ‘morally *preferable*’ when describing the moral status of a state of affairs which are gradable, which can also be seen in the following examples:

(71) Eating chicken is *morally worse* than killing Cecil the lion

(*Eating Chicken Is Morally Worse than Killing Cecil the Lion* - Vox, n.d.)

<https://www.vox.com/2015/7/30/9074547/cecil-lion-chicken-meat> - emphasis added)

(72) [I]t is *morally preferable* that a just warrior be better protected from unnecessary harm, other things being equal. Wouldn't it have been *morally better*, for example, if Allied pilots in World War II could have remotely flown planes to defeat the Nazis rather than risk being shot down?

(*Coming to Terms With How Drones Are Used* - NYTimes.Com, n.d.)

[https://www.nytimes.com/roomfordebate/2012/09/25/do-drone-attacks-do-more-harm-than-](https://www.nytimes.com/roomfordebate/2012/09/25/do-drone-attacks-do-more-harm-than-good/coming-to-terms-with-how-drones-are-used)

[good/coming-to-terms-with-how-drones-are-used](https://www.nytimes.com/roomfordebate/2012/09/25/do-drone-attacks-do-more-harm-than-good/coming-to-terms-with-how-drones-are-used) - emphasis added)

(73) Most People Consider Themselves to Be *Morally Superior*

(*Most People Consider Themselves to Be Morally Superior* - Scientific American, n.d.)

<https://www.scientificamerican.com/article/most-people-consider-themselves-to-be-morally->

[superior](https://www.scientificamerican.com/article/most-people-consider-themselves-to-be-morally-superior) - emphasis added)

(74) Is wearing fur *morally worse than* wearing leather?

(*Is Wearing Fur Morally Worse than Wearing Leather?* | *Ethical and Green Living* | *The Guardian*, n.d.) -

[https://www.theguardian.com/environment/2015/feb/15/is-wearing-leather-less-moral-than-](https://www.theguardian.com/environment/2015/feb/15/is-wearing-leather-less-moral-than-wearing-fur)

[wearing-fur](#) - emphasis added)

The above examples show how speakers are forced to express degrees of rightness and wrongness by using 'better' and 'worse' instead of using the modified forms 'right-er' or 'wrong-er' which makes manifest a tension between the conceptual and linguistic levels. This tension is apparent in the work of Richard Arneson who uses the ungrammatical forms 'right-er' and 'wrong-er' in order to more accurately reflect his nuanced view of morality:

(75) [T]he act consequentialist should downplay the distinction between acts that are right and wrong. Her more important task is to grade acts as "righter" and "wronger" depending on the extent of the shortfall between the act being evaluated and the best that could have been done in the circumstances... We can think of the acts an agent could do on some occasion as ordered in an array of groups of acts that have consequences that range from very close to the consequences of the best act to very close to the very worst one could have done. With this picture in view, we can see that options of a sort have an important role in moral life and moral assessment. Far more important than determining whether one's act on an occasion was right or wrong would be fixing the degree of wrongness if it is not the very best one could have done. (Arneson, 2009)

It can be concluded, in light of the above, that the conceptual nature of RIGHT and WRONG, can be mentally represented as both mutually exclusive and gradable in

nature, and that the form of representation varies with context. The conceptual form, as either gradable or mutually exclusive, is determined by the perceived nature of the states of affairs to which the concepts are applied. It is also apparent that when states of affairs dictate a gradable understanding of *RIGHT* and *WRONG*, agents resort to using typically non-moral evaluative language to express this gradeability – such as ‘better’ and ‘worse’. This is due to the adjectives *right/wrong* being resistant to morphological modification. Explaining this resistance is beyond the scope of the current work, however, I might conjecture that it is plausibly due to linguistic convention based on culturally engrained views about morality. Notwithstanding, it is clear from the above that speakers do use language which reflects a gradable conceptualization of morality. Moreover, an important observation can be made: linguistic conventions appear to limit the expression of certain concepts, suggesting the link between concepts and words is not a straightforward relationship.

4.3.12 Grounding Right and Wrong

As we have seen above, the antonymic status of moral adjectives appears to vary with context and this represents a linguistic manifestation of the way in which the speaker mentally represents moral concepts. Given this, we might wonder why this happens; why the conceptual form of *RIGHT* and *WRONG* need to be grounded in or mapped onto other concepts. This is not the case, of course, for other non-moral concepts which seem to have a fixed or intrinsic conceptual form such as *HOT* and *COLD*, or *ON* and *OFF*, so we might wonder why *RIGHT* and *WRONG* differ. This phenomenon, I claim, can be plausibly explained by alluding to the highly abstract conceptual nature of *RIGHT* and *WRONG*. Moral concepts are abstract in the sense that we are unable to understand and interact with them through our senses on a multi-modal level. In contrast to the case of *HOT* and *COLD* – about which we are able to collect rich sensory data directly; building up detailed conceptual knowledge of the properties – *RIGHT* and *WRONG* have no sensible form. The abstract conceptual form is therefore ‘malleable’ and requires a concrete concept upon which they can be mapped, grounding their conceptual and, therefore, linguistic nature.

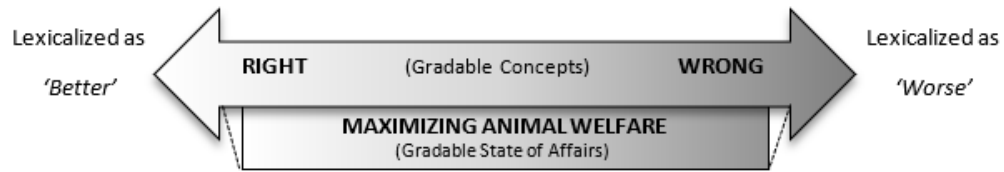


FIGURE 3 *RIGHT & WRONG as Gradable Concepts*



FIGURE 4 *RIGHT & WRONG as Mutually Exclusive Concepts*

4.3.13 Summary

It appears from the analysis above that, unlike the non-moral antonym pairs such as those mentioned in my introductory remarks, there is no fixed form of antonymy observed between *right/wrong* when used as predicate adjectives in English. This observation suggests that the way an agent mentally represents *RIGHT* and *WRONG* is flexible and context-dependent also. It has indeed been shown, from the study of contextualized moral utterances, that the conceptual form of *RIGHT* and *WRONG* is flexible and can be represented as gradable and mutually exclusive. This finding can be explained, and is indeed to be expected, from the current perspective. The conceptual nature of *RIGHT* and *WRONG* as abstract concepts is malleable and, therefore, takes on the nature of the states of affairs to which they are applied.

CHAPTER FIVE

CONCLUSIONS & IMPLICATIONS

5. Conclusions and Implications

As a whole, this work has been a defence of the thesis that:

The moral concepts RIGHT and WRONG are emotionally grounded; such that being able to experience moral emotions is a necessary condition for grasping moral concepts. The conceptual form of RIGHT and WRONG is flexible due to their abstract nature and can, therefore, be mentally represented in divergent forms as either mutually exclusive or gradable opposites.

This thesis was supported by three separate analyses of moral language (§ 4.1, 4.2 & 4.3) which had the following specific aims:

- i. To elucidate the emotional structure of the moral concepts RIGHT and WRONG (§ 4.1).
- ii. To understand whether Conceptual Metaphor Theory is an effective framework through which to study the structure of moral concepts (§ 4.2).
- iii. To show the divergent conceptual forms which RIGHT and WRONG, as binary opposites, assume in the mind (§ 4.3).

The conclusions of each of these studies are summarized in turn in the proceeding paragraphs.

5.1 The Emotional Constitution of Moral Concepts

In the first study (§ 4.1), I argued, by way of linguistic analysis, that moral concepts *RIGHT* and *WRONG* are emotionally constituted, or embodied, such that being disposed to experience moral emotions is a necessary condition for grasping moral concepts.

In supporting this claim, I surveyed the extant psychological and neurological evidence which suggests strong influence of emotional brain structures in moral thought. I then presented an analogical argument drawing a parallel between moral properties and concepts and emotion-dispositional properties and concepts. I showed that moral language functions in an analogous way to emotion-dispositional language, thus giving reason to believe that moral concepts are emotional in nature. In offering further support, I drew attention to the fact that moral judgements are often explicitly expressed in emotion-dispositional terms and also frequently commit the attribution error. I developed a reading of the attribution error which allows us to explain why it happens by alluding to figure and ground construal and showing how emotional stimuli are foregrounded due to their cognitive salience – thus, explaining why the attribution error is commonly made with emotional and moral language.

Hence, in light of the linguistic data and the convergent neurological evidence, it is concluded that the moral concepts *RIGHT* and *WRONG* can be understood as being a particular class of emotion-dispositional concepts and, as such, are emotionally constructed.

5.2 Conceptual Metaphor

It can be concluded from the second study (§ 4.2), which aims at examining the adequacy of Conceptual Metaphor Theory in analyzing the conceptual structure of *MORALITY*, that a clear and robust understanding of *MORALITY* cannot be achieved within this framework.

Firstly, I undertook an analysis of *MORALITY* by compiling examples of figurative language in English which, according to CMT, are the linguistic products of conceptual metaphors and can therefore, they claim, be used as evidence for their existence. It became apparent through using this procedure that several divergent source domains

present themselves as structuring MORALITY, namely: VERTICALITY; QUANTITY; BRIGHTNESS; MERIT; EMOTION and CLEANLINESS. This, I claim, raises the issues of how this is possible; why it is necessary and if it is possible that these divergent domains have some conceptual relation to each other. It is vital that this issue is solved, as it seemingly undermines the foundational claims of CMT: that abstract domains are necessarily structured by source domains and that the conceptual structure on these source domains is therefore necessary to the concept; that the target concept is always understood in terms of the source concept. If this is true, then cases which show the structuring of the same concept in terms of a divergent source domain are counter examples to the claimed conceptual necessity of any other source-target matching, as either it is the case that the other source domain is not conceptually necessary, or it *is* conceptually necessary but has not been sufficient for a figurative utterance and so the CMT theorist is faced with the fact that either the conceptual claim is wrong or the linguistic claim is. I endeavoured to offer a possible way of circumventing this problem by postulating and outlining a possible enrichment to CMT which hypothesised that: if divergent source domains can be mapped onto the same target domain, then all those source domains must bear a conceptual relation to each other. I showed how this could be done by positing some simple principles of organization which are consistent with the main assumptions of CMT and showing how these relations between source domains can be confirmed by applying the standard procedure of figurative language analysis used in CMT.

5.3 Conceptual Form

Based on the foundational assumption that conceptual form determines linguistic form, I have presented an analysis of moral language (§ 4.3) which shows that moral concepts RIGHT and WRONG can be represented in the minds of agents in two distinct ways viz. as mutually exclusive concepts, or as gradable concepts. I have shown that the way in which these concepts are mentally represented is determined by the conceptual nature of the state of affairs alluded to by the agent's moral judgements. I have further claimed that the form in which moral concepts are mentally represented in the mind of an

agent will have a bearing on the judgements that an agent assents to and on the strength of an agent's motivation to act in accordance with their judgements. This is due to the fact that agents who represent moral concepts as mutually exclusive will automatically form corollary judgements by conceptual necessity. It follows that if motivational internalism is true, agents who represent moral concepts in such a way will be motivated toward the same course of action by two separate moral judgements, thus, such an agent will be more strongly motivated to act in accordance with their judgements when moral concepts are represented as mutually exclusive, as opposed to gradable. If such a connection between conceptual representation and motivation holds, we could expect that agents will be more likely to act in accordance with some judgements than others. Indeed, this thesis could be tested empirically and, if correct, provides us with a logical way of explaining why certain moral judgements seem to motivate agents more strongly than others.

5.4 Further Insights & Implications

In this final section I aim to draw out and reconcile some of the further implications and insights gained from the studies presented above.

5.4.1 Kinds of Concept

It should be apparent to the reader that the analyses given above, although related in subject matter (morality), deal with conceptual entities of distinct kinds. The first two analyses deal with the structure of moral concepts, whereas the third analysis, in contrast, looks at their representational form. The difference between the conceptual phenomena studied is akin, therefore, to that between *structure* and *shape*.

What, then, can these separate analyses, when taken together, tell us about human conceptualization? I propose that two observations of paramount importance are to be gained from this work:

- i. That concepts not only differ in size, or amount of content (e.g. concepts, frames, domains), but also in *kind*.

- ii. That comprehensive understanding of divergent conceptual kinds is necessary for a full

explanation of how complex concepts are built.

Apropos of these two points, I will proceed to explain the importance of divergent conceptual kinds in accounting for the possibility of complex conceptual compositionality.

5.4.2 Building A Theory of Conceptual Kinds

Consider an analogy between concepts and language. One *prima facie* and standardly-held observation about language is that it exhibits the property of compositionality. Sometimes called the Principle of Compositionality, this states that the meaning of a complex expression in a language is determined jointly by its structure and the individual meanings of its parts.¹⁸ Hence, the meaning of a phrase or sentence is a function of its syntactical rules and lexical semantics. Varying lexical forms (noun, verb, adjective etc.) are vital to the composition of more complex meanings expressed through larger phrases and sentences. If all words were nouns, for example, it is extremely difficult to see how the construction of complex and meaningful sentences would be possible. Take the string of words '*coffee, desk, office*' for instance. These words have no clearly comprehensible combined meaning – they do not form a meaningful sentence or proposition. This can be explained, in part, by pointing to the lack of qualitatively different lexical forms (i.e. only nouns are present) which would be essential for enabling the composition of a meaningful sentence e.g. '*My coffee is on the desk in the office*'.

It is argued by some that concepts also display compositionality. The most prominent version of such a thesis is the Language of Thought Hypothesis, first posited by Jerry Fodor (Fodor, 1979), although others have explored the phenomenon of conceptual combination (Bose et al., 2018; Costello, 2000; Hampton & Jönsson, 2012; Medin & Shoben, 1988; G. L. Murphy, 1988; Wisniewski, 1997). In addition, there is recent

¹⁸ There is a distinction to be made between ontological compositionality and functional compositionality (Pelletier,

2017). By 'compositionality' I mean the latter functional kind.

data from the neuroscience literature which implicates the brain's default mode network (DMN) in conceptual combination (Frankland & Greene, 2020). Although there is a traditional rejection of the formalist theories which understand concepts as being compositional by cognitive views, the proposal that concepts are compositional is not necessarily one that a cognitivist approach must reject in principle. Conceptual compositionality is, in fact, being studied from a range of theoretical perspectives; some formalist and others cognitivist in nature (Barsalou, 2017; Hampton & Winter, 2017).

In the following, I will argue concepts display compositionality. Compositionality, that is, in the functional sense. Meaning that combinations of concepts produce new emergent concepts which are seemingly more than the sum of their parts. From this assumption, and continuing the analogy with compositionality in language, we might expect that concepts should have this property in virtue of them having divergent forms and functions; that is to say, in order to construct complex and compound concepts the individual conceptual 'atoms' must come in distinct kinds, just as the diversity of lexical form is necessary to compositionality in language¹⁹. Hence, we could claim:

From the observation that systems of language (formal or natural) display compositionality, which is dependent, in part, on the existence of divergent lexical kinds, it is plausible that if concepts also display compositionality, they also exist in heterogeneous forms.

I will proceed to support this claim by showing that:

1. There exist varying kinds of mental representations which are functionally and qualitatively distinct.
2. Distinct conceptual kinds are necessary for the construction of compound concepts.

¹⁹ Whether or not there is a corresponding conceptual 'syntax' is important to investigate, but such an investigation would be out of place here.

5.4.3 Entity Concepts and Predicate Concepts

Assuming that conceptual composition necessitates a distinction between divergent conceptual kinds, I will endeavor, now, to make manifest such a distinction.

Let us start by looking at concepts in the abstract domain of ethics: RIGHT, WRONG and MOTIVE. Although the concepts of RIGHT, WRONG and MOTIVE are somewhat related in content, and all count as mental representations, it appears, under examination, that they are of fundamentally different kinds. The difference between the concepts RIGHT, WRONG and MOTIVE is not simply that they have different conceptual content or 'meaning', but that they are also different *in kind*. That is, MOTIVE represents a qualitatively distinct conceptual entity to RIGHT or WRONG. This difference can be inferred from the observation that they function in distinctive ways.

In order to explicate this functional difference, let's start with the concept RIGHT. RIGHT can be used to modify other concepts, such as MOTIVE. That is, it can be applied to MOTIVE in order to produce the compound concept RIGHT MOTIVE. To put this in logical terms, RIGHT functions as a 'predicate' and, in linguistic terms, RIGHT functions as an 'adjective'. Let us call concepts which function in this manner *predicate concepts*.

In contrast to RIGHT, MOTIVE is distinct insofar as it does not function in this predicative manner; it cannot be used to modify other concepts, but can be *modified by* predicate concepts such as RIGHT in RIGHT MOTIVE, as seen above. This difference in function must be explained, I claim, by a qualitative difference in kind between concepts like MOTIVE on one hand and RIGHT or WRONG on the other. Let's call the concepts which can be modified by predicate concepts, such as MOTIVE, *entity concepts*.

In making this distinction we recognize an intrinsic functional difference in nature between predicate concepts and entity concepts; such that, RIGHT, as a predicate concept, can be used to modify an entity concept, like MOTIVE, in order to produce the compound RIGHT MOTIVE, but that MOTIVE, as an entity concept, cannot be used to modify a predicate concept such as RIGHT.

We can make this distinction between predicate concepts and entity concepts clearer. I invite the reader to take part in an imaginative thought experiment, so that we may examine our *analysandum* directly. Picture, if you will, a ball. You should have a representation in mind. Now, imagine a black ball. You have just modified the existing representation. The ball has a new property – it's black. Next, imagine that the ball is pink. You have, again, performed a mental operation whereby you updated the base entity concept (BALL) with the predicate concept (PINK) and, as a result, you have a different representation, or image, in your mind's eye. This can be repeated with a large number of predicate concepts. You could, for instance, continue to modify the entity concept BALL with other predicate concepts such as WHITE, SMALL, BIG, SPOTTED, SOFT and so on. This imaginative exercise is only possible, I claim, if the concept BALL is of a different nature, or kind, to that of BLACK, WHITE, SMALL etc. It is part of the nature of BALL that it can be modified in this way, thus producing a compound concept, just as an adjective modifies a noun.

Furthermore, there seem to be limits, or constraints, to such imaginative tasks. Clear your mind of BALL and try, now, to imagine SOFT, HARD, WET OR DRY without a base entity concept; try, similarly, to imagine BLACK, or STRIPEY, or SPOTTED in isolation. This cannot be done. We cannot imagine BLACK apart from the entity that is black; or SOFT without the entity that is soft. When imagining BLACK, you are likely to picture an expanse of black, but this expanse implicitly assumes an entity concept, for the black in your mind has size and shape, even if the shape is not clearly defined. Hence, it appears that predicate concepts cannot function as stand-alone concepts; they need an entity to which they are applied, either explicitly or implicitly.

We can ascertain more about the unique behavior of predicate concepts. Now, with your shapeless black expanse in-mind, try to modify BLACK with PINK. Can you picture PINK BLACK? This, also, cannot be imagined due to the nature of predicate concepts; that is, combinations of predicate concepts alone cannot produce conceptual compounds. It may be argued that this is simply due to the colours I have chosen in this example; that combinations of other predicate concepts, such as BLUE and GREEN, can be imagined. It might be claimed, for instance, that we can modify BLUE with GREEN, and that

BLUE GREEN can easily be imagined – we *can* picture blueish green. This is true. However, I propose that the mental operation here is subtly different. When we imagine blueish green, we are not modifying the base concept BLUE with GREEN, but are imagining a whole new predicate concept– call it TEAL. Imagining TEAL is not akin to imagining BLUE GREEN, hence we are not combining BLUE and GREEN, but picturing a separate predicate concept TEAL.

The same failure of composition can also be seen between two entity concepts. Take for example COFFEE TABLE. This construction is normally understood as a being a noun-noun combination, but it is not the case that in the compound the constituent concepts (COFFEE and TABLE) are both entity concepts, although when understood in isolation, they may be. Try, now, to imagine TABLE as a base concept. Once you have the picture in mind, attempt to modify this with the entity concept COFFEE. This cannot be imagined. The same is true for other supposed constructions of two entity concepts such as HOUSE CAT OR BALL PEN. We do indeed understand these constructions; they are clearly comprehended compound concepts, but they are not the result of a cognitive operation which modifies a base entity concept with another entity concept. They are, instead, a case of a subtly different mental process, which I will make manifest in the following section.

The above observations, I claim, suggest the need for a distinction between concepts of (at least) two distinct kinds which relate roughly to the difference in physical kind between objects (entity concepts) and properties (predicate concepts), or, in language, to nouns and adjectives, respectively. Hence, I propose a preliminary characterization of these conceptual kinds in respect to their functionality:

Entity Concept: A mental representation which can be modified by predicate concept, but cannot be used to modify such a concept.

Predicate concept: A mental representation which presupposes an entity concept and can be used to modify, but cannot be modified by, such a concept.

It is probable that more conceptual kinds than this are needed to account for a range of concepts normally expressed through lexical items such as verbs and pronouns or closed-class lexical items which include morphological features such as ‘er’

and 'ing'²⁰ and also for logical connectives such as 'and', 'or' and 'if'. Exploring such cases is an important task, but is one which will take me beyond the scope of the current work.

5.4.4 Compositionality and Emergence

As we are beginning to see, concepts display distinct imaginative functions that determine the ways in which they can be combined. Hence, an understanding of these divergent conceptual forms should shed light on the process of conceptual composition. In order to see this, let's return to the observation mentioned above from Fodor and Lepore (Fodor & Lepore, 1996). Recall that the compound concept $PET\ FISH$ was described as having an emergent conceptual structure, or prototypicality profile, which appears to be more than the sum of its parts. We can make sense of this case by recognizing that PET and $FISH$ are both entity concepts and, as such, cannot be combined. As we have started to see above, conceptual diversity is required for compositionality, hence, the apparent compound concept $PET\ FISH$ cannot be a construction of two entity concepts. The concept PET in the compound $PET\ FISH$ must, therefore, be a predicate concept – meaning something like 'DOMESTIC' – which modifies, the entity concept $FISH$. Hence, the compound $PET\ FISH$ is, in actuality, a compound of the concepts $DOMESTIC\ (PET)$ (predicate concept) and $FISH$ (entity concept). This is made clearer if we understand that 'PET', when comprehended as an entity concept, is really $DOMESTIC-ANIMAL$ and 'PET' as a predicate concept is simply $DOMESTIC$. Given this, if $PET\ FISH$ were a compound of two entity concepts the resulting concept would be $DOMESTIC-ANIMAL, FISH$ and not $DOMESTIC\ FISH$, which is how we naturally understand $PET\ FISH$. Thus, from PET to $PET\ FISH$, a subtle, and almost imperceptible, shift has happened from PET^{Entity} ($DOMESTIC\ ANIMAL$) to $PET^{Predicate}$ ($DOMESTIC$). This is difficult to detect due to language misleading us. The wordform 'PET' by itself and the wordform 'PET' in 'PET FISH' are the same, masking the conceptual shift. The same phenomenon is evident in linguistic constructions where the wordform 'pet' changes from noun to adjective, as shown below.

²⁰ Evans, 2007 has noticed this and calls such concepts *lexical concepts* which he assumes are represented by open-class lexical items. I will return to Evans' distinction below.

(1) 'pet'^{noun} (e.g. 'I have a pet') = [DOMESTIC ANIMAL]^{Ent} (entity concept)

(2) 'pet'^{adj} (e.g. 'A pet fish') = [DOMESTIC]^{Pred} (predicate concept)

Such phrases as 'pet fish' are often referred to in the concept literature as noun-noun constructions, but this is mistaken. We can see that the noun form of 'pet' expresses an entity concept (DOMESTIC ANIMAL) (1) whereas the adjective form of 'pet' expresses a predicate concept (DOMESTIC) (2). With this understanding and our conceptual distinction in-hand we can see how the compound concept PET FISH [DOMESTIC FISH] can only be formed with the coupling of an entity concept and a predicate concept (3), and not between two entity concepts (4):

(3) 'pet fish'^{adj+noun} (e.g. I have a pet fish) = [DOMESTIC FISH]^{Pred+Ent}

(4) 'pet; fish'^{noun+noun} (e.g. a pet, a fish...) = [DOMESTIC-ANIMAL; FISH]^{Ent+Ent}

Hence, 'PET' in PET FISH functions as a predicate concept. This is necessarily the case, as if the two were predicate concepts, they would not compound, but simply refer to a collection, or list, of separate objects. This can be observed in the difference between 'pet fish' and 'pet, fish'. The former being an emergent meaning from the composition of a noun and an adjective and the latter being simply a list of two nouns and, hence, displaying no compositionally emergent meaning.

So, the mystery of why PET FISH has a different typicality profile to PET and FISH combined is understood. PET FISH is a compound concept and, therefore, must be comprised of divergent conceptual forms (one base entity concept and one modifying predicate concept) which allow for functional compositionality and the emergent conceptual profile.

5.4.5 Empirical Evidence for Divergent Conceptual Kinds

The need to acknowledge distinctions between various conceptual kinds has been emphasized by others who give both theoretical and empirical support for such

a view (Gentner & Beranek, 1981; Gentner & France, 2013; Medin et al., 2000). Waxman and Gelman (Waxman & Gelman, 2010) in their chapter entitled *Different kinds of concepts and different kinds of words: What words do for human cognition* who state the following:

Languages include many different kinds of words (e.g. nouns, adjectives, verbs), each of which is recruited to convey a different kind of concept (e.g. categories of objects, properties, and events, respectively). Yet even in the current literature, most of the developmental research on words and concepts has focused on only one kind of word: nouns. Although important insights have been gained, this focus is not without costs. Chief among them is that the principles underlying the acquisition of nouns and the conceptual consequences of their use differ importantly from the principles underlying the acquisition of other grammatical forms, including adjectives and verbs, and the conceptual consequences of using these forms. (Waxman & Gelman, 2010)

Research in developmental psychology shows that, from infancy, children are able to understand distinctions which relate to various conceptual kinds such as I have posited here. For instance, it has been shown that even pre-verbal infants can understand the differences between nouns, adjectives and verbs through grammatical and situational context, and, based on this understanding, are able to make a conceptual distinction between objects, properties and actions accordingly. Children appear to interpret newly encountered nouns as referring to objects or individuals (Gelman & Taylor, 1984; Macnamara, 1984; Markman, 1989; Waxman, 1990), adjectives as picking out properties (Klibanoff & Waxman, 2000; Mintz & Gleitman, 2002; Waxman & Markow, 1998) and verbs as applying to actions and events (Fisher, 2002; Waxman et al., 2009). Such evidence supports the claim that different kinds of concept exist and shows that an understanding of these heterogeneous conceptual kinds allows children to grasp new words. Indeed, if there were no difference in conceptual kind, it is difficult to understand how children would be able to learn language at all, as knowing a particular word form is a noun, and that nouns refer to entities, for instance, is what allows children to know that a particular word applies to an object, rather than a property or an action, in a particular

scene. Thus, the qualitative difference in conceptual kind is seen as vital to language acquisition and must logically and developmentally precede it.

Another notable observation is that certain kinds of concept appear to emerge earlier than others. Waxman and Gelman note that ‘infants appear to tease apart first the grammatical form noun and map this form specifically to object categories. This noun-category link sets the stage for the evolution of more specific expectations linking adjectives and verbs to their respective meanings’ (2010). That infants would need to be able to understand nouns – and, therefore, entity concepts – first is logically consistent with the view outlined here; that predicate concepts – which relate to properties and, therefore, adjectives – need, or pre-suppose entity concepts for their application. Hence, the thesis that entity concepts are functionally distinct, and are logically and conceptually prior to predicate concepts enjoys empirical support (Waxman & Gelman, 2010).

5.4.6 The Flexibility of Words and Concepts

There is clearly a parallel between conceptual kind and lexical kind – as is standardly assumed in psychological investigations such as those referred to in the previous sections – and, thus, an apparent relation between language and thought – an assumption shared in the present study. An important question to ask, however, is: Does conceptual kind determine lexical kind? Traditional views of language would have answered negatively to this question. However, there is a commonly held view in psychology and cognitive linguistics which sees the distinction between nouns, verbs, adjectives and other lexical forms as conceptually grounded. I will proceed to show that the answer to this question is affirmative, but also that the link between language and thought is not straightforward.

To elaborate, ‘black’ which is normally considered an adjective can in fact be used as a noun in specific contexts, such as in a game of pool. I might, for instance, refer to the black ball as simply ‘black’ when I say ‘I’m on *black*’ or ‘the black’ when I say ‘I potted *the black*’. In this case, however, it is important to remember that although there appears to be linguistic flexibility brought about through use and context (there has been a lexical transformation from adjective to noun) this does not necessarily equate to, or

reflect, a corresponding conceptual transformation. To illustrate, the word 'black' when used as a noun, is a case of using the same *wordform* in a non-standard way to refer to a different concept. Under normal circumstances, 'black' might be used as an adjective to refer to the predicate concept BLACK, but, in this specific instance (when playing pool), it is being used as a noun to refer to the entity concept BALL. Hence, a single wordform can be used flexibly to refer to various concepts which will determine the lexical status of 'black' as either an adjective or noun:

(5) 'black'^{adj} (e.g. 'The black ball') = [BLACK] (predicate concept)

(6) 'black'^{noun} (e.g. 'I potted the black') = [BALL] (entity concept)

As should be evident from the above, when the *lexical form* is morphed from adjective to noun, this does not mean that the *concept* is being morphed in the same manner. When 'black' is being used as a noun, one's understanding of the mentally stored concept BLACK has not thereby changed from a predicate concept to an entity concept; it is simply that the lexical item has been applied to a different concept altogether (BALL). We should exercise caution, therefore, when making inferences from lexical form to conceptual form, for if we were to conclude that in using 'black' as a noun, the agent therefore understands BLACK as an entity concept, we would be mistaken.

We can see in (6) that the noun form of 'black' does relate to an entity concept, but the concept is not BLACK, it's BALL. Hence, it appears we would be justified in concluding that lexical kind correlates with conceptual kind; such that nouns refer to entity concepts and adjectives refer to predicate concepts. We would be wrong, however, to conclude that when a particular wordform is morphed from adjective to noun in a specific context, the predicate concept morphs into an entity concept – this is not the case. The concept itself has not changed, but the wordform has merely been applied to a concept of a divergent form, thus inheriting a new lexical form. Hence, it appears that even though lexical form is flexible, conceptual form is stable.

In sum, it can be shown that there is indeed a link between lexical form and conceptual form; such that nouns refer to entity concepts and adjectives refer to

predicate concepts. The analysis above, however, suggests that identical wordforms undergo a lexical transformation when connected to a different concept of a divergent conceptual kind and not due to a morphing of conceptual kind. A distinction in conceptual kind, then, allows for the clearer understanding of how lexical form is determined.

With this understanding of conceptual forms and how they construe lexical form, I will now proceed to compare this view with a notable theory of conceptual kinds from the field of cognitive linguistics in order to orient it on the theoretical landscape.

5.4.7 Other Conceptual Kinds

The need to recognize concepts of qualitatively distinct kinds has been noted by psychologists, as mentioned above, but has had comparatively less attention in the field of cognitive linguistics. This has also been noted by Vyvyan Evans, who has done much in recent years to deepen the understanding of conceptual forms in the area (Evans, 2006, 2009b, 2015).

Evans claims that a distinction is needed to separate two kinds of concept which have traditionally been conflated in the research viz. *lexical concepts* and *cognitive models*. In his framework of meaning construction, Lexical Concepts and Cognitive Models (LCCM), Evans emphasizes the need to recognize the distinction between the rich conceptual content (cognitive models) that is conveyed through open-class lexical items such as nouns, verbs and adjectives, and the schematic content (lexical concepts) expressed by closed-class lexical items like prepositions, copulas and morphemes, such as plural markers (Evans, 2006). Evans claims that the rich content stored in the conceptual system and the schematic semantic content stored in the linguistic system represent qualitatively different kinds of concept, and that both are needed in the construction of meaning through language. The linguistic system, he claims, acts as an 'executive control function' enabling the expression of rich knowledge stored in the conceptual system (Evans, 2015). As Evans explains:

I suggest that the linguistic system evolved, in part, by facilitating more effective control of the extant representations in the conceptual system. That is, linguistic

representations are specialised for providing a 'scaffolding' to structure conceptual representations, thereby facilitating their use in communication. (Evans, 2009a)

Since its initial formulation, Evans has elaborated on the LCCM framework by introducing the distinction between *analogue concepts* and *parametric concepts* (Evans, 2015). Analogue concepts are integrated into the existing LCCM model by being associated with open-class words and the rich content which they are assumed to afford access to, and, accordingly, parametric concepts are associated with closed-class lexical items which are, conversely, assumed to represent highly schematic content. As he explains:

the distinction between the conceptual and linguistic systems relates to concepts that are analog in nature, on the one hand, and those that are parametric in nature, on the other. In so doing, I argue against received accounts of embodied cognition that fail to recognize such a distinction. (Evans, 2015)

Evans' LCCM model is a theory of *access semantics*. On this view, meaning is said to be derived in virtue of the fact that language provides access to specific portions of an agent's full encyclopedic knowledge background. This view is integrated with the theoretical constructs of analog concepts and parametric concepts just mentioned as follows:

parametric concepts facilitate access to analog concepts in the process of meaning construction. Although both types of concept are derived from embodied, or as I shall prefer, grounded experience, they are qualitatively distinct. Parametric concepts are schematic, while analog concepts are richer, more closely constituting analogs of the experience types they are grounded in. (Evans, 2015)

Evans' distinction between analog and parametric kinds differs to the one I make here between entity concepts and predicate concepts. Both entity and predicate concepts are

forms of what Evans calls analogue concepts, as they are built of rich embodied encyclopedic knowledge and inhere in the conceptual system. The distinction proposed herein, however, could be reconciled with Evans' model. If we were to integrate the LCCM model with the current view, then, we would have a distinction between two kinds of analog concepts (entity and predicate) plus the parametric concepts proposed by Evans. Hence, entity and predicate concepts are subsets of analog concepts. Such an integration seems plausible as Evans appears to make the same claim about the distinction between entities and predicates as I do, and even underscores the importance of demarcating these conceptual kinds when he states that 'complex thoughts, actions, and so on require that our concepts can be combined compositionally in order to form complex ideas' (Evans, 2015).

This being said, however, Evans claims that it is the parametric lexical concepts which are responsible for the distinction between the two forms, and that the difference between entity and predicate is, therefore, not manifest in pure analog knowledge. Instead, the divergence, he suggests, is created when analogue content is accessed and construed through the parametric schematic content. So, for Evans, the difference between object (entity) and property (predicate) is created by the linguistic system and is not therefore a distinction which exists in the conceptual system proper. Evans explains this using the example of 'red' and 'redness':

[T]he grammatical distinction between the adjective (red) and noun (redness) appears to relate to a semantic distinction between the notion of property versus thing. The words red and redness, while indexing the same (or similar) perceptual state, also encode schematic concepts: PROPERTY versus THING. (Evans, 2015)

Although LCCM theory has many merits, I find it difficult to see how the difference in entity and predicate concepts – or 'PROPERTY' and 'THING' – only arises through the interplay of analog concepts with lexical concepts in the linguistic system, as, if this is the case, non-linguistic creatures or pre-verbal infants would not be able to distinguish the difference in

properties of objects and the objects themselves. As mentioned above, this conceptual distinction does indeed seem to be present in pre-verbal infants and appears to be a necessary pre-requisite for the acquisition of language.

There is much merit to the LCCM theory and Evans presents a convincing analysis. This is supported by experimental evidence suggesting the existence of highly schematic parametric concepts and how they are encoded into the linguistic system. However, I would argue that there needs to be a further distinction in kind between various analog concepts that distinguishes – at least – the difference in function between entity concepts and predicate concepts.

5.4.8 Summary

I have argued, here, for a clear distinction between functionally and qualitatively distinct conceptual kinds. Such a distinction, I claim, is necessary in order to account for the complexity and compositional nature of conceptualization. I have presented a phenomenological analysis in support of this claim which reveals the divergent imaginative functionality between *predicate concepts* – which represent properties and function in such a way as to modify entity concepts – and *entity concepts* – which cannot modify other entity concepts or predicate concepts. I have further supported the existence of these conceptual kinds by making reference to empirical data which points to entity concepts being logically, and developmentally prior to predicate concepts. This difference is manifest in language through the lexical distinctions between noun (entity concepts) and adjective (predicate concepts). I have also endeavored to show that, the concept-word link is not straightforward; such that it cannot be inferred from the flexibility of lexical forms that concepts are also malleable in the same manner, but that conceptual form does determine lexical form.

Describing the nature of conceptual representation is a Herculean task – without even considering the link between concepts and language. The distinct nature of conceptual kinds must be recognized in order to fully understand the process of conceptual compositionality and demarcating entity concepts and predicate concepts is seen as a necessary, albeit incomplete, start to this task.

5.5 In Conclusion

The current work highlights an important theoretical point: current frameworks within Cognitive Linguistics have not taken enough care in clearly delineating the different kinds of mental phenomena which we can broadly call 'concepts'. That is to say, even the simplest concepts must be extremely complex entities – let alone complexes of such concepts such as domains and frames – and the theoretical treatment of such mental representations, therefore, deserves an extremely detailed elucidation which should start, I suggest, with reaching an understanding of the divergent conceptual kinds, and explicating the most simple concrete concepts before expanding such an account to include and encompass the more abstract concepts and more complex construction built from them i.e. domains, frames and domain matrices.

The task of describing the nature of concepts is, in itself, a monumental task – without even considering the link between these concepts and language. Much work in Cognitive Linguistics is unsatisfactory in this regard, as it posits and analyzes the more complex mental representations (domains, frames, domain matrices) without first reaching a detailed understanding of the nature of the most basic conceptual kinds, the analysis of which, as I have endeavored to show in the previous section, will inform and lead the study of more complex conceptual structures.

We can now, in bringing the insights from this work together, sketch out a multi-dimensional understanding of moral concepts, as shown in Figure 3.

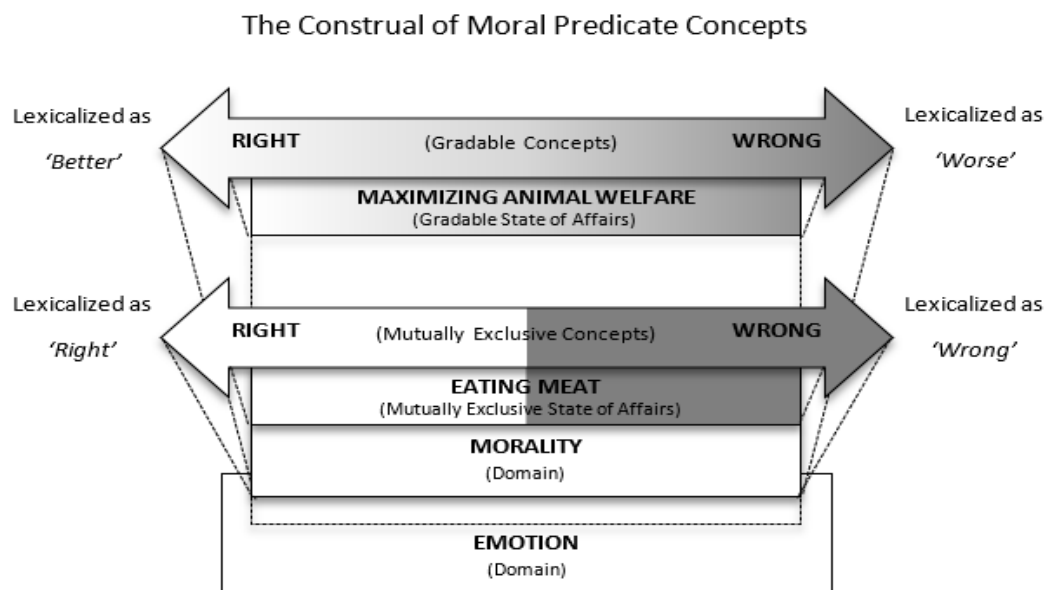


FIGURE 5 Integrated Schematic Diagram Showing Structure and Form of Moral Concepts

The findings of the above can be stated in the following propositions:

- I. Moral concepts are embodied; grounded by emotional states.
- II. Being able to experience certain emotional states is a necessary condition for grasping the conceptual domain EMOTION.
- III. Understanding the conceptual domain EMOTION is a necessary condition for grasping the domain MORALITY and its constituent concepts RIGHT and WRONG.
- IV. Moral concepts RIGHT and WRONG are abstract concepts and as such can be represented as either mutually exclusive or gradable in nature.

- V. Moral concepts RIGHT and WRONG are predicate concepts due to their conceptual functioning.

- VI. Moral concepts RIGHT and WRONG, in being predicate concepts, are applied to, and therefore modify, entity concepts.

- VII. Moral concepts RIGHT and WRONG, as binary opposite, abstract predicate concepts, can be mentally represented in either mutually exclusive or gradable form, which is determined by the nature of the entity concept to which they are applied.

5.5.1 Final Thoughts

The fundamental goal of this work has been to reach a comprehensive understanding of moral concepts. I hope I have been successful in enabling a deeper insight into how moral concepts are represented and structured in the human conceptual system.

During the process of conducting this study, it has become clear that, although there is undoubtedly a close link between thought and language, we theorists working on such problems and adopting varied frameworks need to be extremely cautious when drawing inferences and conclusions about the structure and representation of particular mental phenomena from the analysis of language alone, as, although language is used to express and communicate conceptual content, linguistic data is of an inherently different and more schematic kind to that of analogue conceptual data which is rich, encyclopaedic and, in some cases, multimodal, or embodied. It is inevitable, given this understanding, that from the analysis of language alone, one can only hope to gain a very simplified and impoverished view of the mind.

Having said this, as shown above, it is possible to make use of linguistic analysis in studying the conceptual system, but such work ideally needs to be supplemented with non-linguistic experimental evidence and coupled with rigorous logical and conceptual analysis.

Although, I am confident that novel insights into the phenomenon of morality have been gained through this study, it has, of course, also spawned many further questions which, unfortunately, it leaves unanswered; such is the nature of philosophical and scientific investigation. I will give just three examples of such questions here, each directed at a specific discipline:

A question for psychologists: How, in principle, are conceptual associations to be identified as distinct from conceptually necessary cross-domain mappings in the brain?

A question for philosophers: What are the logical restrictions of studying the human conceptual system which are imposed by the limitations of the conceptual system itself?

A question for linguists: If conceptual data is inherently richer and more complex than linguistic data, how far can linguistic analysis take us in the study of the human conceptual system?

In closing, I wish to express my hope that the knowledge gained from the current work not only adds to our understanding of that profound and fundamental human concept MORALITY, but also that the insights – as well as any inconsistencies – here will act as a foundation upon which a fuller and more comprehensive understanding of morality, mind and meaning can eventually be built.

Justin J. Bartlett

Srinakharinwirot University, 2022

REFERENCES

- Acquaviva, P., Lenci, A., Paradis, C., & Raffaelli, I. (2020). Models of lexical meaning. In *Word Knowledge and Word Usage* (pp. 353–404). De Gruyter Mouton.
<https://doi.org/10.1515/9783110440577-010>
- Adamaszek, M., D'Agata, F., Ferrucci, R., Habas, C., Keulen, S., Kirkby, K. C., Leggio, M., Mariën, P., Molinari, M., Moulton, E., Orsi, L., van Overwalle, F., Papadelis, C., Priori, A., Sacchetti, B., Schutter, D. J., Styliadis, C., & Verhoeven, J. (2016). Consensus Paper: Cerebellum and Emotion. *The Cerebellum* 2016 16:2, 16(2), 552–576. <https://doi.org/10.1007/S12311-016-0815-8>
- Acquaviva, P., Lenci, A., Paradis, C., & Raffaelli, I. (2020). Models of lexical meaning. In *Word Knowledge and Word Usage* (pp. 353–404). De Gruyter Mouton.
<https://doi.org/10.1515/9783110440577-010>
- Adamaszek, M., D'Agata, F., Ferrucci, R., Habas, C., Keulen, S., Kirkby, K. C., Leggio, M., Mariën, P., Molinari, M., Moulton, E., Orsi, L., van Overwalle, F., Papadelis, C., Priori, A., Sacchetti, B., Schutter, D. J., Styliadis, C., & Verhoeven, J. (2016). Consensus Paper: Cerebellum and Emotion. *The Cerebellum* 2016 16:2, 16(2), 552–576. <https://doi.org/10.1007/S12311-016-0815-8>
- Allport, G. W. (1954). *The nature of prejudice*. Addison-Wesley.
- Acquaviva, P., Lenci, A., Paradis, C., & Raffaelli, I. (2020). Models of lexical meaning. In *Word Knowledge and Word Usage* (pp. 353–404). De Gruyter Mouton.
<https://doi.org/10.1515/9783110440577-010>
- Adamaszek, M., D'Agata, F., Ferrucci, R., Habas, C., Keulen, S., Kirkby, K. C., Leggio, M., Mariën, P., Molinari, M., Moulton, E., Orsi, L., van Overwalle, F., Papadelis, C., Priori, A., Sacchetti, B., Schutter, D. J., Styliadis, C., & Verhoeven, J. (2016). Consensus Paper: Cerebellum and Emotion. *The Cerebellum* 2016 16:2, 16(2), 552–576. <https://doi.org/10.1007/S12311-016-0815-8>
- Allport, G. W. (1954). *The nature of prejudice*. Addison-Wesley.
- Anderson, S. R., & Lightfoot, D. W. (2002). The Language Organ. In *The Language*

- Organ*. Cambridge University Press. <https://doi.org/10.1017/cbo9780511613869>
- Animal rights: 2018's big wins — and big losses — for animals* - Vox. (n.d.). Retrieved February 4, 2021, from <https://www.vox.com/2018/12/24/18148698/2018-year-in-review-for-animals>
- antisocial personality disorder* – *APA Dictionary of Psychology*. (n.d.). Retrieved September 19, 2020, from <https://dictionary.apa.org/antisocial-personality-disorder>
- Aparicio, H. (2018). *PROCESSING CONTEXT-SENSITIVE EXPRESSIONS: THE CASE OF GRADABLE ADJECTIVES AND NUMERALS A DISSERTATION SUBMITTED TO THE FACULTY OF THE DIVISION OF THE HUMANITIES IN CANDIDACY FOR THE DEGREE OF DOCTOR OF PHILOSOPHY DEPARTMENT OF LINGUISTICS*.
- Are things either right or wrong?* (n.d.). Retrieved February 4, 2021, from <https://www.gqkidz.org/right-or-wrong.html>
- Aristotle. (1963). *Categories, and De interpretatione*, (J. L. Ackrill, Ed.). Oxford Calrendon Press.
- Aristotle., & Heath, Malcolm. (1996). *Poetics*. Penguin Books.
- Arneson, R. J. (2009). *What Do We Owe to Distant Needy Strangers*.
- Ayto, J. (2003). *The Oxford dictionary of rhyming slang*. Oxford University Press.
- Barsalou, L. W. (2007). Grounded Cognition.
[Http://Dx.Doi.Org/10.1146/Annurev.Psych.59.103006.093639](http://Dx.Doi.Org/10.1146/Annurev.Psych.59.103006.093639), 59, 617–645.
<https://doi.org/10.1146/ANNUREV.PSYCH.59.103006.093639>
- Barsalou, L. W. (2017). Cognitively Plausible Theories of Concept Composition.
Language, Cognition, and Mind, 3, 9–30. https://doi.org/10.1007/978-3-319-45977-6_2
- Bartlett, J. J. (2020). An Expected Error: An Essay in Defence of Moral Emotionism.
Axiomathes, 1–19. <https://doi.org/10.1007/s10516-020-09524-5>
- Bartlett, J. J. (2021). Motivating Emotions: Emotionism and the Internalist Connection.
Axiomathes, 1–21. <https://doi.org/10.1007/s10516-021-09550-x>
- Beaney, M. (2021). *Analysis: Conceptions of Analysis in Analytic Philosophy* (Stanford

- Encyclopedia of Philosophy*). <https://plato.stanford.edu/entries/analysis/s6.html>
- Benczes, R., & Ságvári, B. (2018a). Where metaphors really come from: Social factors as contextual influence in Hungarian teenagers' metaphorical conceptualizations of life. *Cognitive Linguistics*, 29(1), 121–154. <https://doi.org/10.1515/COG-2016-0139>
- Benczes, R., & Ságvári, B. (2018b). Where metaphors really come from: Social factors as contextual influence in Hungarian teenagers' metaphorical conceptualizations of life. *Cognitive Linguistics*, 29(1), 121–154. <https://doi.org/10.1515/COG-2016-0139/PDF>
- Bermúdez, J. L. (2014). *Cognitive Science: An Introduction to the Science of the Mind* (2nd Edition). Cambridge University Press.
- Bernhard, R. M., Chaponis, J., Siburian, R., Gallagher, P., Ransohoff, K., Wikler, D., Perlis, R. H., & Greene, J. D. (2016). Variation in the oxytocin receptor gene (OXTR) is associated with differences in moral judgment. *Social Cognitive and Affective Neuroscience*, 11(12), 1872. <https://doi.org/10.1093/SCAN/NSW103>
- Blair, R. J. R. (2008). The amygdala and ventromedial prefrontal cortex: functional contributions and dysfunction in psychopathy. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 363(1503), 2557. <https://doi.org/10.1098/RSTB.2008.0027>
- Blair, R. J. R., & Cipolotti, L. (2000). Impaired social response reversal. A case of “acquired sociopathy.” *Brain*, 123(6), 1122–1141. <https://doi.org/10.1093/brain/123.6.1122>
- Blair, R. J. R., Hwang, S., White, S. F., & Meffert, H. (2016). Emotional Learning, Psychopathy, and Norm Development. In *Moral Brains* (pp. 185–202). Oxford University Press. <https://doi.org/10.1093/acprof:oso/9780199357666.003.0009>
- Bose, A., Alonso, J., & Meteyard, L. (2018). Compound production in aphasia: Contributions of semantic transparency and imageability. *Frontiers in Human Neuroscience*, 12. https://doi.org/10.3389/FNHUM.2018.228.00028/EVENT_ABSTRACT
- Brewer, W. F., & Lichtenstein, E. H. (1975). Recall of logical and pragmatic implications

- in sentences with dichotomous and continuous antonyms. *Memory & Cognition*, 3(3), 315–318. <https://doi.org/10.3758/BF03212917>
- Carey, S. (2009). The Origin of Concepts. In *The Origin of Concepts*. Oxford University Press. <https://doi.org/10.1093/acprof:oso/9780195367638.001.0001>
- Casasanto, D. (2009). *When is a linguistic metaphor conceptual metaphor?* (pp. 127–145). <https://doi.org/10.1075/hcp.24.11cas>
- Castroviejo, E., McNally, L., & Sassoon, G. W. (2018). *Gradability, Vagueness, and Scale Structure: From the Armchair to the Lab* (pp. 1–24). https://doi.org/10.1007/978-3-319-77791-7_1
- Charles, W. G., Reed, M. A., & Derryberry, D. (1994). Conceptual and associative processing in antonymy and synonymy. *Applied Psycholinguistics*, 15(3), 329–354. <https://doi.org/10.1017/S0142716400065929>
- Churchland, P. Smith. (2011). *Braintrust: What Neuroscience Tells Us About Morality*. Princeton University Press. <https://philpapers.org/rec/CHUBWN>
- Clark, A. (1999). An embodied cognitive science? *Trends in Cognitive Sciences*, 3(9), 345–351. [https://doi.org/10.1016/S1364-6613\(99\)01361-3](https://doi.org/10.1016/S1364-6613(99)01361-3)
- Clark, A., & Chalmers, D. (1998). The extended mind. In *Analysis* (Vol. 58, Issue 1, pp. 7–19). Oxford University Press. <https://doi.org/10.1093/analysis/58.1.7>
- Coming to Terms With How Drones Are Used - NYTimes.com*. (n.d.). Retrieved February 4, 2021, from <https://www.nytimes.com/roomfordebate/2012/09/25/do-drone-attacks-do-more-harm-than-good/coming-to-terms-with-how-drones-are-used>
- Costello, F. (2000). Efficient creativity: constraint-guided conceptual combination. *Cognitive Science*, 24(2), 299–349. [https://doi.org/10.1016/S0364-0213\(00\)00020-3](https://doi.org/10.1016/S0364-0213(00)00020-3)
- Crawford, L. E. (2009). Conceptual Metaphors of Affect. *Emotion Review*, 1(2), 129–139. <https://doi.org/10.1177/1754073908100438>
- Crockett, M. J. (2016). Morphing Morals. In *Moral Brains* (pp. 237–245). Oxford

- University Press. <https://doi.org/10.1093/acprof:oso/9780199357666.003.0011>
- Croft, W. A., & Wood, E. J. (2000). Construal operations in linguistics and artificial intelligence. *Undefined*, 51–78. <https://doi.org/10.1075/CELCR.2.04CRO>
- Croft, W., & Cruse, D. A. (2004). *Cognitive Linguistics*.
<https://doi.org/10.1017/CBO9780511803864>
- Cruse, D. A. (1986). *No Title Lexical Semantics*. Cambridge University Press.
- Cruse, D. A., & Togia, P. (1996). Towards a Cognitive Model of Antonymy. *Lexicology*.
- D'Arms, J., & Jacobson, D. (2007). *Sensibility Theory and Projectivism*.
<https://doi.org/10.1093/OXFORDHB/9780195325911.003.0008>
- Darwall, S., Gibbard, A., & Railton, P. (1992). Toward Fin de siecle Ethics: Some Trends. *The Philosophical Review*, 101(1), 115. <https://doi.org/10.2307/2185045>
- Debate Kit: Is It Ethical to Eat Animals? | PETA*. (n.d.). Retrieved February 4, 2021, from <https://www.peta.org/teachkind/lesson-plans-activities/eating-animals-ethical-debate-kit/>
- Dennett, D. (1987). *The Intentional Stance*. The MIT Press.
- Devinsky, O., Morrell, M. J., & Vogt, B. A. (1995). Contributions of anterior cingulate cortex to behaviour. *Brain : A Journal of Neurology*, 118 (Pt 1)(1), 279–306.
<https://doi.org/10.1093/BRAIN/118.1.279>
- Donald, M. (1991). *Origins of the Modern Mind*. Harvard University Press.
- Dove, G. (2011). On the need for embodied and dis-embodied cognition. *Frontiers in Psychology*, 1(JAN), 242. <https://doi.org/10.3389/FPSYG.2010.00242/BIBTEX>
- Dove, G. (2015). Three symbol ungrounding problems: Abstract concepts and the future of embodied cognition. *Psychonomic Bulletin & Review* 2015 23:4, 23(4), 1109–1121. <https://doi.org/10.3758/S13423-015-0825-4>
- Eating chicken is morally worse than killing Cecil the lion - Vox*. (n.d.). Retrieved February 4, 2021, from <https://www.vox.com/2015/7/30/9074547/cecil-lion-chicken-meat>

- Evans, V. (2004). *The Structure of Time: Language, Meaning and Temporal Cognition*. John Benjamins.
- Evans, V. (2006). Lexical concepts, cognitive models and meaning-construction. *Cognitive Linguistics*, 17(4), 491–534. <https://doi.org/10.1515/COG.2006.016>
- Evans, V. (2007). *A Glossary of Cognitive Linguistics*. Edinburgh University Press.
- Evans, V. (2009a). *Semantic representation in LCCM Theory* (pp. 27–55). <https://doi.org/10.1075/hcp.24.06eva>
- Evans, V. (2009b). How words mean: Lexical concepts, cognitive models, and meaning construction. In *How Words Mean: Lexical Concepts, Cognitive Models, and Meaning Construction*. Oxford University Press. <https://doi.org/10.1093/acprof:oso/9780199234660.001.0001>
- Evans, V. (2015). What's in a concept? Analog versus parametric concepts in LCCM Theory. In E. Margolis & S. Laurence (Eds.), *The Conceptual Mind: New Directions in the Study of Concepts* (pp. 251–290). MIT Press.
- Federmeier, K. D., & Kutas, M. (1999). A Rose by Any Other Name: Long-Term Memory Structure and Sentence Processing. *Journal of Memory and Language*, 41(4), 469–495. <https://doi.org/10.1006/jmla.1999.2660>
- Fillmore, C. J. (1968). The Case for Case. In E. Bach & R. Harms (Eds.), *Universals in Linguistic Theory*. New York: Holt, Rinehart, and Winston.
- Fillmore, C. J. (1976). FRAME SEMANTICS AND THE NATURE OF LANGUAGE. *Annals of the New York Academy of Sciences*, 280(1 Origins and E), 20–32. <https://doi.org/10.1111/j.1749-6632.1976.tb25467.x>
- Fisher, C. (2002). Structural limits on verb mapping: The role of abstract structure in 2.5-year-olds' interpretations of novel verbs. *Developmental Science*, 5(1), 55–64. <https://doi.org/10.1111/1467-7687.00209>
- Fodor, J. A. (1979). *The language of thought*. Harvard University Press.
- Fodor, J., & Lepore, E. (1996). The red herring and the pet fish: Why concepts still can't be prototypes. *Cognition*, 58(2), 253–270. [https://doi.org/10.1016/0010-0277\(95\)00694-X](https://doi.org/10.1016/0010-0277(95)00694-X)

Fogelin, R. J. (2011). *Figuratively Speaking: Revised Edition. Figuratively Speaking: Revised Edition*, 1–144.

<https://doi.org/10.1093/ACPROF:OSO/9780199739998.001.0001>

Frankland, S. M., & Greene, J. D. (2020). Concepts and Compositionality: In Search of the Brain's Language of Thought. <https://doi.org/10.1146/Annurev-Psych-122216-011829>, 71, 273–303. <https://doi.org/10.1146/ANNUREV-PSYCH-122216-011829>

Frege, G. (1879). *Begriffsschrift: Eine der arithmetischen nachgebildete Formelsprache des reinen Denkens*. <https://philpapers.org/rec/FREBED>

Fumagalli, M., & Priori, A. (2012). Functional and clinical neuroanatomy of morality. *Brain*, 135(7), 2006–2021. <https://doi.org/10.1093/BRAIN/AWR334>

Gallese, V., Fadiga, L., Fogassi, L., & Rizzolatti, G. (1996). Action recognition in the premotor cortex. *Brain : A Journal of Neurology*, 119 (Pt 2)(2), 593–609. <https://doi.org/10.1093/BRAIN/119.2.593>

Gallese, V., & Lakoff, G. (2005). The Brain's concepts: the role of the Sensory-motor system in conceptual knowledge. *Cognitive Neuropsychology*, 22(3), 455–479. <https://doi.org/10.1080/02643290442000310>

Geeraerts, D., & Cuyckens, H. (2007). *The Oxford handbook of cognitive linguistics*. Oxford University Press.

Gelman, S. A. (2007). The Essential Child: Origins of Essentialism in Everyday Thought. In *The Essential Child: Origins of Essentialism in Everyday Thought*. Oxford University Press. <https://doi.org/10.1093/acprof:oso/9780195154061.001.0001>

Gelman, S. A., & Taylor, M. (1984). How Two-Year-Old Children Interpret Proper and Common Names for Unfamiliar Objects. *Child Development*, 55(4), 1535. <https://doi.org/10.2307/1130023>

Gentner, D., & Beranek, B. (1981). SOME INTERESTING DIFFERENCES BETWEEN VERBS AND NOUNS. *Cognition and Brain Theory*.

Gentner, D., & France, I. M. (2013). The verb mutability effect: Studies of the combinatorial semantics of nouns and verbs. *Lexical Ambiguity Resolution: Perspective from*

- Psycholinguistics, Neuropsychology and Artificial Intelligence*, 343–382.
<https://doi.org/10.1016/B978-0-08-051013-2.50018-5>
- Gibbard, Allan. (1990). *Wise choices, apt feelings : a theory of normative judgment*.
 Harvard University Press.
- Gibbard, Allan. (2006). Moral Feelings and Moral Concepts. In R. Shafer-Landau (Ed.),
Oxford Studies in Metaethics (Vol. 1). Oxford University Press.
- Gibbs, J., & Colston, H. L. (2012). Interpreting Figurative Meaning. In *Interpreting
 Figurative Meaning*. Cambridge University Press.
<https://doi.org/10.1017/cbo9781139168779>
- Gibbs Jr, R. W. (2017). Metaphor Wars. In *Metaphor Wars*. Cambridge University Press.
<https://doi.org/10.1017/9781107762350>
- Gibbs, R. (2008). The Cambridge Handbook of Metaphor and Thought. In *The Cambridge
 Handbook of Metaphor and Thought*. Cambridge University Press.
<https://doi.org/10.1017/cbo9780511816802>
- Gibbs, R. W. (2014). Why Do Some People Dislike Conceptual Metaphor Theory?
Cognitive Semiotics, 5(1–2), 14–36. <https://doi.org/10.1515/cogsem.2013.5.12.14>
- Gibbs; Raymond. (1994). *The Poetics of Mind: figurative thought language and
 understanding*. Cambridge University Press.
- Gibson, J. (1966). *The senses considered as perceptual systems*. Houghton Mifflin.
- Goldberg; Adele E. (1995). *Constructions: A Construction Grammar Approach to
 Argument Structure*. University of Chicago Press.
- Goldinger, S. D., Papesh, M. H., Barnhart, A. S., Hansen, W. A., & Hout, M. C. (2016). The
 poverty of embodied cognition. *Psychonomic Bulletin & Review* 23:4, 23(4),
 959–978. <https://doi.org/10.3758/S13423-015-0860-1>
- Gottwald, J. M., Elsner, B., & Pollatos, O. (2015). Good is up-spatial metaphors in action
 observation. *Frontiers in Psychology*, 6(OCT).
<https://doi.org/10.3389/fpsyg.2015.01605>
- Greene, J. D., Nystrom, L. E., Engell, A. D., Darley, J. M., & Cohen, J. D. (2004). The

- Neural Bases of Cognitive Conflict and Control in Moral Judgment. *Neuron*, 44(2), 389–400. <https://doi.org/10.1016/J.NEURON.2004.09.027>
- Greene, J. D., Sommerville, R. B., Nystrom, L. E., Darley, J. M., & Cohen, J. D. (2001). An fMRI investigation of emotional engagement in moral judgment. *Science*, 293(5537), 2105–2108. <https://doi.org/10.1126/science.1062872>
- Greenwald, A. G., & Banaji, M. R. (1995). Implicit Social Cognition: Attitudes, Self-Esteem, and Stereotypes. *Psychological Review*, 102(1), 4–27. <https://doi.org/10.1037/0033-295X.102.1.4>
- Gross, D., Fischer, U., & Miller, G. A. (1989). The organization of adjectival meanings. *Journal of Memory and Language*, 28(1), 92–106. [https://doi.org/10.1016/0749-596X\(89\)90030-2](https://doi.org/10.1016/0749-596X(89)90030-2)
- Group, P. (2007). MIP: A method for identifying metaphorically used words in discourse. *Metaphor and Symbol*, 22(1), 1–39. <https://doi.org/10.1080/10926480709336752>
- Hampton, J. A., & Jönsson, M. L. (2012). *Typicality and Composition a Lily: the Logic of Combining Vague Concepts*. <https://doi.org/10.1093/OXFORDHB/9780199541072.013.0018>
- Hampton, J. A., & Moss, H. E. (2003). Concepts and meaning: Introduction to the special issue on conceptual representation. *Language and Cognitive Processes*, 18(5–6), 505–512. <https://doi.org/10.1080/01690960344000161>
- Hampton, J. A., & Winter, Y. (2017). *Compositionality and Concepts in Linguistics and Psychology* (J. A. Hampton & Y. Winter, Eds.; Vol. 3). Springer International Publishing. <https://doi.org/10.1007/978-3-319-45977-6>
- Harenski, C. L., Antonenko, O., Shane, M. S., & Kiehl, K. A. (2008). Gender differences in neural mechanisms underlying moral sensitivity. *Social Cognitive and Affective Neuroscience*, 3(4), 313–321. <https://doi.org/10.1093/SCAN/NSN026>
- Harenski, C. L., Edwards, B. G., Harenski, K. A., & Kiehl, K. A. (2014). Neural correlates of moral and non-moral emotion in female psychopathy. *Frontiers in Human Neuroscience*, 8(SEP), 1–10. <https://doi.org/10.3389/FNHUM.2014.00741/ABSTRACT>

Harenski, C. L., & Hamann, S. (2006). Neural correlates of regulating negative emotions related to moral violations. *NeuroImage*, *30*(1), 313–324.

<https://doi.org/10.1016/J.NEUROIMAGE.2005.09.034>

Haser, V. (2005). Metaphor, Metonymy, and Experientialist Philosophy. In *Metaphor, Metonymy, and Experientialist Philosophy*. DE GRUYTER.

<https://doi.org/10.1515/9783110918243>

Henderson, J. (2021). Truth and Gradability. *Journal of Philosophical Logic*, 1–25.

<https://doi.org/10.1007/s10992-020-09584-3>

Hills, D. (n.d.). *Metaphor*. Retrieved from <https://plato.stanford.edu/entries/metaphor/>

Horstmann, G., & Ansorge, U. (2011). Compatibility Between Tones, Head Movements, and Facial Expressions. *Emotion*, *11*(4), 975–980.

<https://doi.org/10.1037/a0023468>

Huangfu, G., Li, L., Zhang, Z., & Sheng, C. (2021). Moral metaphorical effect of cleanliness on immoral workplace behaviors: Environmental cleanliness or self-cleanliness? <https://doi.org/10.1177/18344909211034257>, 15.

Humphreys, G. W., Price, C. J., & Riddoch, M. J. (1999). From objects to names: A cognitive neuroscience approach. *Psychological Research*, *62*(2–3), 118–130.

<https://doi.org/10.1007/s004260050046>

Hutchins, E. (1995). *Cognition in the Wild*. The MIT Press.

Inhelder, B., & Piaget, J. (1964). *The Early Growth of Logic in the Child: Classification and Seriation*. Routledge and Kegan Paul.

Is wearing fur morally worse than wearing leather? | Ethical and green living | The Guardian. (n.d.). Retrieved February 4, 2021, from

<https://www.theguardian.com/environment/2015/feb/15/is-wearing-leather-less-moral-than-wearing-fur>

Isen, A. M., & Levin, P. F. (1972). Effect of feeling good on helping: Cookies and kindness. *Journal of Personality and Social Psychology*, *21*(3), 384–388.

<https://doi.org/10.1037/h0032317>

It is either right or wrong; there's no grey area | *The Daily Walk*. (n.d.). Retrieved February 4, 2021, from <https://www.thedailywalk.org/it-is-either-right-or-wrong-theres-no-grey-area/>

Johnson, M. (1993). *Moral imagination : implications of cognitive science for ethics*. University of Chicago Press.

Jones, J. W. (2019). *Living religion : embodiment, theology, and the possibility of a spiritual sense*. Oxford University Press.

Jones, Steven., Murphy, M. Lynne., Paradis, Carita., & Willners, Caroline. (2012). *Antonyms in English : Construals, Constructions and Canonicity*. Cambridge University Press.

Keysar, B., Shen, Y., Glucksberg, S., & Horton, W. S. (2000). Conventional Language: How Metaphorical Is It? *Journal of Memory and Language*, 43(4), 576–593. <https://doi.org/10.1006/jmla.2000.2711>

Klibanoff, R. S., & Waxman, S. R. (2000). Basic level object categories support the acquisition of novel adjectives: Evidence from preschool-aged children. *Child Development*, 71(3), 649–659. <https://doi.org/10.1111/1467-8624.00173>

Kosfeld, M., Heinrichs, M., Zak, P. J., Fischbacher, U., & Fehr, E. (2005). Oxytocin increases trust in humans. *Nature*, 435(7042), 673–676. <https://doi.org/10.1038/NATURE03701>

Kövecses, Z. (1995). American friendship and the scope of metaphor. *Cognitive Linguistics*, 6(4), 315–346. <https://doi.org/10.1515/cogl.1995.6.4.315>

Kövecses, Z. (2000). *Metaphor and emotion: Language, culture, and body in human feeling*. Cambridge University Press.

Kövecses, Z. (2010). *Metaphor: A Practical Introduction* (Second). Oxford University Press.

Kövecses, Z. (2015). Metaphor and Culture. In *Where Metaphors Come From* (pp. 73–96). Oxford University Press. <https://doi.org/10.1093/acprof:oso/9780190224868.003.0005>

- Kövecses, Z. (2020). Extended Conceptual Metaphor Theory. In *Extended Conceptual Metaphor Theory*. Cambridge University Press.
<https://doi.org/10.1017/9781108859127>
- Lakoff, G. (1990). The Invariance Hypothesis: Is abstract reason based on image-schemas? *Cognitive Linguistics*, 1(1), 39–74.
<https://doi.org/10.1515/cogl.1990.1.1.39>
- Lakoff, G. (1996b). *Moral Politics: What Conservatives Know That Liberals Don't*. University of Chicago Press.
- Lakoff, G. (2014). Mapping the brain's metaphor circuitry: metaphorical thought in everyday reason. *Frontiers in Human Neuroscience*, 8(1), 958.
<https://doi.org/10.3389/fnhum.2014.00958>
- Lakoff, G., & Johnson, M. (1980). *Metaphors We Live By*. University of Chicago Press.
- Lakoff, G., & Johnson, M. (1999). *Philosophy in the flesh : the embodied mind and its challenge to Western thought*. Basic Books.
- Langacker, R. W. (1987). Foundations of cognitive grammar. Theoretical prerequisites. In *Language Sciences* (Vol. 1, Issue 2). Stanford university press.
[https://doi.org/10.1016/0388-0001\(89\)90007-7](https://doi.org/10.1016/0388-0001(89)90007-7)
- Leezenberg, M. (2014). From Cognitive Linguistics to Social Science: Thirty Years after *Metaphors We Live By*. *Cognitive Semiotics*, 5(1–2), 140–152.
<https://doi.org/10.1515/cogsem.2013.5.12.140>
- Liao, S. M. (2016). *Moral Brains: The Neuroscience of Morality - Google Books*. Oxford University Press.
- Lin, E. L., Murphy, G. L., & Shoben, E. J. (1997). The Effects of Prior Processing Episodes on Basic level Superiority. *The Quarterly Journal of Experimental Psychology A*, 50(1), 25–48. <https://doi.org/10.1080/027249897392215>
- Lizardo, O. (2012). The conceptual bases of metaphors of dirt and cleanliness in moral and non-moral reasoning. *Cognitive Linguistics*, 23(2), 367–393.
<https://doi.org/10.1515/COG-2012-0011/MACHINEREADABLECITATION/RIS>

- Luo, Q., Nakic, M., Wheatley, T., Richell, R., Martin, A., & Blair, R. J. R. (2006). The neural basis of implicit moral attitude--an IAT study using event-related fMRI. *NeuroImage*, 30(4), 1449–1457.
<https://doi.org/10.1016/J.NEUROIMAGE.2005.11.005>
- Lycan, W. G. (2008). *Philosophy of language: A contemporary introduction*. Routledge.
- Lynott, D., & Coventry, K. (2014). On the ups and downs of emotion: Testing between conceptual-metaphor and polarity accounts of emotional valence-spatial location interactions. *Psychonomic Bulletin and Review*, 21(1), 218–226.
<https://doi.org/10.3758/s13423-013-0481-5>
- Lyons, J. (1977). Semantics. In *Semantics*. Cambridge University Press.
<https://doi.org/10.1017/cbo9780511620614>
- Macnamara, J. (1984). *Names for Things: A Study of Human Learning*. The MIT Press.
- Margolis, E., & Laurence, S. (1999). *Concepts: Core Readings*.
<https://philpapers.org/rec/MARCCR>
- Margolis, E., & Laurence, S. (2015). *The Conceptual Mind* (E. Margolis & S. Laurence, Eds.). The MIT Press.
- Markman, E. (1989). *Categorization and Naming in Children: Problems of Induction*. The MIT Press.
- McDowell, J. (1985). *Values and secondary qualities*. In Ted Honderich (ed.), *Morality and Objectivity*. London: Routledge. pp. 110-129.
- McGlone, M. S. (2007). What is the explanatory value of a conceptual metaphor? *Language and Communication*, 27(2), 109–126.
<https://doi.org/10.1016/j.langcom.2006.02.016>
- McGlone, M. S. (2011). Hyperbole, Homunculi, and Hindsight Bias: An Alternative Evaluation of Conceptual Metaphor Theory. *Discourse Processes*, 48(8), 563–574.
<https://doi.org/10.1080/0163853X.2011.606104>
- McNeil, D. (1992). *Hand and mind: What gestures reveal about thought*. University of Chicago Press.

- Medin, D. L., Lynch, E. B., & Solomon, K. O. (2000). Are there kinds of concepts? *Annual Review of Psychology, 51*, 121–147.
<https://doi.org/10.1146/ANNUREV.PSYCH.51.1.121>
- Medin, D. L., & Schaffer, M. M. (1978). Context theory of classification learning. *Psychological Review, 85*(3), 207–238. <https://doi.org/10.1037/0033-295X.85.3.207>
- Medin, D. L., & Shoben, E. J. (1988). Context and structure in conceptual combination. *Cognitive Psychology, 20*(2), 158–190. [https://doi.org/10.1016/0010-0285\(88\)90018-7](https://doi.org/10.1016/0010-0285(88)90018-7)
- Meier, B. P., Hauser, D. J., Robinson, M. D., Friesen, C. K., & Schjeldahl, K. (2007). What's "Up" With God? Vertical Space as a Representation of the Divine. *Journal of Personality and Social Psychology, 93*(5), 699–710. <https://doi.org/10.1037/0022-3514.93.5.699>
- Meier, B. P., & Robinson, M. D. (2006). Does "feeling down" mean seeing down? Depressive symptoms and vertical selective attention. *Journal of Research in Personality, 40*(4), 451–461. <https://doi.org/10.1016/j.jrp.2005.03.001>
- Meier, B. P., Sellbom, M., & Wygant, D. B. (2007). Failing to take the moral high ground: Psychopathy and the vertical representation of morality. *Personality and Individual Differences, 43*(4), 757–767. <https://doi.org/10.1016/j.paid.2007.02.001>
- Meneses, A., & Liy-Salmeron, G. (2012). Serotonin and emotion, learning and memory. *Reviews in the Neurosciences, 23*(5–6), 543–553.
<https://doi.org/10.1515/REVNEURO-2012-0060>
- Miller, G. A., Beckwith, R., Fellbaum, C., Gross, D., & Miller, K. J. (1990). Introduction to WordNet: An On-line Lexical Database*. *International Journal of Lexicography, 3*(4), 235–244. <https://doi.org/10.1093/IJL/3.4.235>
- Miller, K. A., Raney, G. E., & Demos, A. P. (2020). Time to Throw in the Towel? No Evidence for Automatic Conceptual Metaphor Access in Idiom Processing. *Journal of Psycholinguistic Research, 49*(5), 885–913. <https://doi.org/10.1007/s10936-020-09728-1>
- Mintz, T. H., & Gleitman, L. R. (2002). Adjectives really do modify nouns: The

- incremental and restricted nature of early adjective acquisition. *Cognition*, 84(3), 267–293. [https://doi.org/10.1016/S0010-0277\(02\)00047-1](https://doi.org/10.1016/S0010-0277(02)00047-1)
- Moll, J., de Oliveira-Souza, R., Bramati, I. E., & Grafman, J. (2001). *Functional Networks in Emotional Moral and Nonmoral Social Judgments*. <https://doi.org/10.1006/nimg.2002.1118>
- Moll, J., de Oliveira-Souza, R., Eslinger, P. J., Bramati, I. E., Mourão-Miranda, J., Andreiuolo, P. A., & Pessoa, L. (2002). The Neural Correlates of Moral Sensitivity: A Functional Magnetic Resonance Imaging Investigation of Basic and Moral Emotions. *Journal of Neuroscience*, 22(7), 2730–2736. <https://doi.org/10.1523/jneurosci.22-07-02730.2002>
- Morality Is Black And White: There Is No Gray – THE ROAD TO CONCORD*. (n.d.). Retrieved February 4, 2021, from <https://theroadtoconcord.com/natural-law/derevation/rights-bubbles-the-origin-of-universal-morality/morality-is-black-and-white-there-is-no-gray/>
- Morton, A. (1980). *Frames of Mind: Constraints on the Common-Sense Conception of the Mental*. <https://philpapers.org/rec/MORFOM-2>
- Most People Consider Themselves to Be Morally Superior - Scientific American*. (n.d.). Retrieved February 4, 2021, from <https://www.scientificamerican.com/article/most-people-consider-themselves-to-be-morally-superior/>
- Murphy, G. (2002). *The Big Book of Concepts*. The MIT Press.
- Murphy, G. L. (1988). Comprehending Complex Concepts. *Cognitive Science*, 12(4), 529–562. https://doi.org/10.1207/S15516709COG1204_2
- Murphy, G. L. (2016). Is there an exemplar theory of concepts? *Psychonomic Bulletin and Review*, 23(4), 1035–1042. <https://doi.org/10.3758/s13423-015-0834-3>
- Murphy, G. L., & Allopenna, P. D. (1994). The Locus of Knowledge Effects in Concept Learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 20(4), 904–919. <https://doi.org/10.1037/0278-7393.20.4.904>
- Murphy, G. L., & Andrew, J. M. (1993). The Conceptual Basis of Antonymy and

- Synonymy in Adjectives. *Journal of Memory and Language*, 32(3), 301–319.
<https://doi.org/10.1006/jmla.1993.1016>
- Murphy, G. L., & Medin, D. L. (1985). The Role of Theories in Conceptual Coherence. *Psychological Review*, 92(3), 289–316. <https://doi.org/10.1037/0033-295X.92.3.289>
- Murphy, M. L. (2003). Semantic Relations and the Lexicon. In *Semantic Relations and the Lexicon*. Cambridge University Press.
<https://doi.org/10.1017/cbo9780511486494>
- Nosofsky, R. M., & Alfonso-Reese, L. A. (1999). Effects of similarity and practice on speeded classification response times and accuracies: Further tests of an exemplar-retrieval model. *Memory and Cognition*, 27(1), 78–93.
<https://doi.org/10.3758/BF03201215>
- Nosofsky, R. M., & Palmeri, T. J. (1997a). An Exemplar-Based Random Walk Model of Speeded Classification. *Psychological Review*, 104(2), 266–300.
<https://doi.org/10.1037/0033-295X.104.2.266>
- Nosofsky, R. M., & Palmeri, T. J. (1997b). Comparing exemplar-retrieval and decision-bound models of speeded perceptual classification. *Perception and Psychophysics*, 59(7), 1027–1048. <https://doi.org/10.3758/BF03205518>
- Oms, S., & Zardini, E. (2019). The Sorites Paradox. In S. Oms & E. Zardini (Eds.), *The Sorites Paradox*. Cambridge University Press.
<https://doi.org/10.1017/9781316683064>
- Osherson, D. N., & Smith, E. E. (1981). On the adequacy of prototype theory as a theory of concepts. *Cognition*, 9(1), 35–58. [https://doi.org/10.1016/0010-0277\(81\)90013-5](https://doi.org/10.1016/0010-0277(81)90013-5)
- Paradis, C., & Willners, C. (2011). Antonymy. *Review of Cognitive Linguistics*, 9(2), 367–391. <https://doi.org/10.1075/rcl.9.2.02par>
- Pelletier, F. J. (2017). Compositionality and Concepts—A Perspective from Formal Semantics and Philosophy of Language. *Language, Cognition, and Mind*, 3, 31–94. https://doi.org/10.1007/978-3-319-45977-6_3
- Prinz, J. (2006). The emotional basis of moral judgments. *Philosophical Explorations*,

- 9(1), 29–43. <https://doi.org/10.1080/13869790500492466>
- Prinz, J. J. (2007). *The Emotional Construction of Morals*. Oxford University Press.
- Rakova, M. (2002). The philosophy of embodied realism: A high price to pay? *Cognitive Linguistics*, 13(3), 215–244. <https://doi.org/10.1515/cogl.2002.015>
- Rips, L. J., Shoben, E. J., & Smith, E. E. (1973). Semantic distance and the verification of semantic relations. *Journal of Verbal Learning and Verbal Behavior*, 12(1), 1–20. [https://doi.org/10.1016/S0022-5371\(73\)80056-8](https://doi.org/10.1016/S0022-5371(73)80056-8)
- Rohrer, T. (2010). *Embodiment and Experientialism*. <https://doi.org/10.1093/OXFORDHB/9780199738632.013.0002>
- Rosch, E., & Mervis, C. B. (1975). Family resemblances: Studies in the internal structure of categories. *Cognitive Psychology*, 7(4), 573–605. [https://doi.org/10.1016/0010-0285\(75\)90024-9](https://doi.org/10.1016/0010-0285(75)90024-9)
- Rosch, E., Mervis, C. B., Gray, W. D., Johnson, D. M., & Boyes-Braem, P. (1976). Basic objects in natural categories. *Cognitive Psychology*, 8(3), 382–439. [https://doi.org/10.1016/0010-0285\(76\)90013-X](https://doi.org/10.1016/0010-0285(76)90013-X)
- Schmahmann, J. D., & Caplan, D. (2006). Cognition, emotion and the cerebellum. In *Brain : a journal of neurology* (Vol. 129, Issue Pt 2, pp. 290–292). Oxford Academic. <https://doi.org/10.1093/brain/awh729>
- Schnall, S., Haidt, J., Clore, G. L., & Jordan, A. H. (2008). Disgust as embodied moral judgment. *Personality and Social Psychology Bulletin*, 34(8), 1096–1109. <https://doi.org/10.1177/0146167208317771>
- Schubert, T. W. (2005). Your highness: Vertical positions as perceptual symbols of power. In *Journal of Personality and Social Psychology* (Vol. 89, Issue 1, pp. 1–21). <https://doi.org/10.1037/0022-3514.89.1.1>
- Shapiro, L. A. (2011). *Embodied cognition*. Routledge.
- Siegel, J. Z., & Crockett, M. J. (2013). How serotonin shapes moral judgment and behavior. *Annals of the New York Academy of Sciences*, 1299(1), 42. <https://doi.org/10.1111/NYAS.12229>
- Smoke, K. L. (1932). An objective study of concept formation. *Psychological*

- Monographs*, 42(4), i–46. <https://doi.org/10.1037/h0093297>
- Spalding, T. L., & Murphy, G. L. (1996). Effects of background knowledge on category construction. *Journal of Experimental Psychology: Learning Memory and Cognition*, 22(2), 525–538. <https://doi.org/10.1037/0278-7393.22.2.525>
- Steen, G. J., Dorst, A. G., Herrmann, J. B., Kaal, A., Krennmayr, T., & Pasma, T. (2010). A Method for Linguistic Metaphor Identification. In *Converging Evidence in Language and Communication Research*. John Benjamins. <https://doi.org/10.1075/celcr.14>
- Suchman, L. (1987). *Plans and situated actions: The problem of human-machine communication*. Cambridge University Press.
- Talmy, L. (1975). Figure and Ground in Complex Sentences. *Annual Meeting of the Berkeley Linguistics Society*, 1, 419–430. <https://doi.org/10.3765/bls.v1i0.2322>
- Talmy, L. (2000). *Toward a Cognitive Semantics, Volume 1*. The MIT Press.
- Vervaeke, J., & Kennedy, J. M. (1996). Metaphors in Language and Thought: Falsification and Multiple Meanings. *Metaphor and Symbolic Activity*, 11(4), 273–284. https://doi.org/10.1207/s15327868ms1104_3
- Vigliocco, G., & Vinson, D. P. (2012). Semantic representation. In *The Oxford Handbook of Psycholinguistics*. Oxford University Press. <https://doi.org/10.1093/oxfordhb/9780198568971.013.0012>
- Vigliocco, G., Warren, J., Siri, S., Arciuli, J., Scott, S., & Wise, R. (2006). The role of semantics and grammatical class in the neural representation of words. *Cerebral Cortex*, 16(12), 1790–1796. <https://doi.org/10.1093/cercor/bhj115>
- Waismann, F. (1977). What is Logical Analysis? In *Philosophical Papers* (pp. 81–103). Springer Netherlands. https://doi.org/10.1007/978-94-010-1144-0_8
- Wattenmaker, W. D., Dewey, G. I., Murphy, T. D., & Medin, D. L. (1986). Linear separability and concept learning: Context, relational properties, and concept naturalness. *Cognitive Psychology*, 18(2), 158–194. [https://doi.org/10.1016/0010-0285\(86\)90011-3](https://doi.org/10.1016/0010-0285(86)90011-3)
- Waxman, S. R. (1990). Linguistic biases and the establishment of conceptual

- hierarchies: Evidence from preschool children. *Cognitive Development*, 5(2), 123–150. [https://doi.org/10.1016/0885-2014\(90\)90023-M](https://doi.org/10.1016/0885-2014(90)90023-M)
- Waxman, S. R., & Gelman, S. A. (2010). Different kinds of concepts and different kinds of words: What words do for human cognition. In *The Making of Human Concepts*. Oxford University Press. <https://doi.org/10.1093/acprof:oso/9780199549221.003.06>
- Waxman, S. R., Lidz, J. L., Braun, I. E., & Lavin, T. (2009). Twenty four-month-old infants' interpretations of novel verbs and nouns in dynamic scenes. *Cognitive Psychology*, 59(1), 67–95. <https://doi.org/10.1016/j.cogpsych.2009.02.001>
- Waxman, S. R., & Markow, D. B. (1998). Object Properties and Object Kind: Twenty-One-Month-Old Infants' Extension of Novel Adjectives. *Child Development*, 69(5), 1313–1329. <https://doi.org/10.1111/j.1467-8624.1998.tb06214.x>
- Weitz, M., & Russell, B. (1961). My Philosophical Development. *The Philosophical Review*, 70(1), 112. <https://doi.org/10.2307/2183411>
- Wertheimer, M. (1923). Untersuchungen zur Lehre von der Gestalt. II. *Psychologische Forschung* 1923 4:1, 4(1), 301–350. <https://doi.org/10.1007/BF00410640>
- Wheatley, T., & Haidt, J. (2005). Hypnotic disgust makes moral judgments more severe. *Psychological Science*, 16(10), 780–784. <https://doi.org/10.1111/j.1467-9280.2005.01614.x>
- Wierzbicka, A. (1986). Metaphors linguists live by. *Paper in Linguistics*, 19(2), 287–313. <https://doi.org/10.1080/08351818609389260>
- Wilkinson, P. R. (2002). *Thesaurus of traditional English metaphors (Second Edition)*. Routledge.
- Wisniewski, E. J. (1997). When concepts combine. *Psychonomic Bulletin and Review*, 4(2), 167–183. <https://doi.org/10.3758/BF03209392>
- Wisniewski, E. J., Lamb, C. A., & Middleton, E. L. (2003). On the conceptual basis for the count and mass noun distinction. In *Language and Cognitive Processes* (Vol. 18, Issues 5–6, pp. 583–624). <https://doi.org/10.1080/01690960344000044>
- Wisniewski, E. J., & Murphy, G. L. (1989). Superordinate and Basic Category Names in

Discourse: A Textual Analysis. *Discourse Processes*, 12(2), 245–261.

<https://doi.org/10.1080/01638538909544728>

Wisniewska, E. J., Imai, M., & Casey, L. (1996). On the equivalence of superordinate concepts. *Cognition*, 60(3), 269–298. [https://doi.org/10.1016/0010-0277\(96\)00707-X](https://doi.org/10.1016/0010-0277(96)00707-X)

Wittgenstein, L. (1922). *Tractatus Logico-Philosophicus*. Routledge.

Wittgenstein, L. (2009). *Philosophical Investigations* (P. M. S. Hacker & J. Schulte, Eds.; 4th Edition). Wiley.

Yu, N. (2015). Metaphorical Character of Moral Cognition: A Comparative and Decompositional Analysis. <Http://Dx.DoI.Org/10.1080/10926488.2015.1049500>, 30(3), 163–183. <https://doi.org/10.1080/10926488.2015.1049500>

Yu, N. (2016). Spatial Metaphors for Morality: A Perspective from Chinese. *Metaphor and Symbol*, 31(2), 108–125. <https://doi.org/10.1080/10926488.2016.1150763>

Yu, N., Wang, T., & He, Y. (2016). Spatial Subsystem of Moral Metaphors: A Cognitive Semantic Study. <Http://Dx.DoI.Org/10.1080/10926488.2016.1223470>, 31(4), 195–211. <https://doi.org/10.1080/10926488.2016.1223470>



APPENDIX



APPENDIX A

Received: 1 July 2020 / Accepted: 12 October 2020

Springer Nature B.V. 2020



An Expected Error: An Essay in Defence of Moral Emotionism

Justin J. Bartlett¹

Received: 1 July 2020 / Accepted: 12 October 2020
© Springer Nature B.V. 2020

Abstract

This work draws an analogical defence of *strong emotionism*—the metaethical claim that moral properties and concepts consist in the propensity of actions to elicit emotional responses from divergent *emotional perspectives*. I offer a theory that is in line with that of Prinz (The emotional construction of morals. Oxford University Press, Oxford, 2007). I build an analogy between moral properties and what I call *emotion-dispositional properties*. These properties are picked out by predicates such as ‘annoying’, ‘frightening’ or ‘deplorable’ and appear to be uncontroversial and frequent cases of *attribution error*—the attributing of subjective emotional states as mind-independent properties. I present a linguistic analysis supporting the claim that moral properties and their related concepts are reducible to a subset of emotion-dispositional properties and concepts. This is grounded in the observation that utterances featuring moral predicates function linguistically and conceptually in analogous ways to emotion-dispositional predicates. It follows from this view that asserted moral utterances are truth-apt relative to ethical communities, but that speakers misconceive the extensions of predicates. I show how the framework of Cognitive Linguistics allows us to explain this error. Further analysis of moral and non-moral utterances exposes the deeper conceptual schemas structuring language through cognitive construal processes. An understanding of these processes, coupled with an emotionist elucidation of moral properties and concepts, makes the attribution error an expected upshot of the emotionist thesis, rather than an uncomfortable consequence.

Keywords Metaethics · Moral psychology · Cognitive linguistics · Emotion · Moral concepts

✉ Justin J. Bartlett
justinbartlettswu@gmail.com

¹ Srinakarinwirot University, Bangkok, Thailand

1 A Difficult Relationship

Much has been said about the relationship between emotion and morality. Aside from the academic work of philosophers, psychologists and sociologists, the relationship between the two is even a well-established notion in folk intuition. Quotidian experience tells us that the transgressions of ethical rules and moral disagreements are often followed with emotionally charged reactions and outbursts, thus making the connection between the moral and the emotional particularly salient. But is the relationship between morality and emotion simply a reactionary one? Do moral properties give rise to emotional responses or do emotional responses give rise to moral properties? This question is not simply a case of chiasmus, a cheap rhetorical device. Its implications have great philosophical weight. Exactly *how* our emotions interact with, or constitute, morality is still a topic of fierce debate among philosophers and psychologists alike; a wide range of metaethical and psychological theories attempt to explain the relationship between moral judgement and emotion. The theoretical melee rages, and discerning the stronger opponent is not straightforward. One observation, however, that is fairly easily made, is how rarely empirical findings from the field of moral psychology, and its related disciplines, enter into the metaethical arena. This has been noted previously by other authors such as Prinz and Nichols who write that:

[E]ven where moral philosophers have invoked emotions, they seldom attend carefully to the psychological characteristics of the emotions to which they appeal. Indeed, it would be hard to exaggerate the extent to which philosophers, even self-described sentimentalists, have neglected psychological research on the moral emotions (Prinz and Nichols 2010).

Our question, here, is an empirical one. This is not to say, however, that a priori deduction or phenomenological analysis are not effective methods of investigation, but rather to point out that answering such a question fully and accurately will inevitably need empirical support as well as philosophical insight. It is my aim in this paper, therefore, to answer the question of whether moral properties give rise to emotional responses, or vice versa by offering an analysis of emotion's necessary role in the formation of moral properties, concepts and judgements. This work is interdisciplinary in nature and makes use of work from psychology and Cognitive Linguistics in supporting the philosophical claim that emotional properties are necessarily conceptually prior to, and constitutive of, moral properties. This is the so-called *emotionist*¹ claim, or rather, a version of it. In the following paragraphs, I assume a strong emotionist stance towards morality which makes the following claim and related assumptions:

Moral judgements are those judgements referring to any conduct ϕ perceived as having the propensity to elicit certain emotional responses in themselves and observers, in certain situations. Such that, a judgement α is a moral

¹ I follow Jesse Prinz's terminology. See Prinz (2007).

judgement *iff* the conduct referred to by α is perceived as having the propensity to elicit certain emotional states in agent x and observer y in S .

This elucidation of moral judgements makes two implicit assumptions:

- i. Metaphysical Claim: Moral properties can be understood as consisting in the propensity of certain actions to elicit moral emotional reactions in agents and observers.
- ii. Epistemic Claim: Grasping moral concepts consist in having the predisposition to experience the emotional reactions alluded to in the metaphysical assertion.

I will proceed to build a defence for the strong emotionist elucidation of moral properties, concepts and judgements as outlined above by presenting an analogical argument highlighting the similarities between moral language and more general emotional language which, based on the assumption that language structure mirrors conceptual structure, supports the conclusion that moral properties can be understood as a subset of emotional properties.

Firstly, I will lay the philosophical foundations of such a theory (Sect. 2) and offer an overview of the psychological evidence which supports a necessary link between emotion and morality (Sect. 3). I will then give further support for this reading of moral judgements by offering an analysis of language which draws strong parallels between moral language and non-moral emotional language. It will be made clear from an analysis of everyday utterances about emotional dispositions, that humans frequently make systematic *attribution errors*² when predicating *emotion-dispositional*³ properties of extramental entities. I will endeavour to show how widespread and pervasive this phenomenon is in ordinary emotional language, and, furthermore, that this is not merely a linguistic mistake but is also, at its foundation, a conceptual one. An analogy will be made such that there can be no *in-principle* objections to systematic attribution error when predicating emotion-dispositional properties of extramental entities. This emotional language use will be compared to moral language and an argument by analogy will be employed supporting the conclusion that moral properties and concepts can be reduced to emotional properties and concepts (Sects. 4 and 5). Lastly, I will briefly offer a plausible explanation as to why we are prone to making this kind of error by employing the conceptual framework of Cognitive Linguistics (Sect. 6). The view I offer here shows that the root of this linguistic and conceptual error is easily explained by making reference to the influence of *construal phenomena*, which have been proposed and explicated by linguists such as Langacker (1987), Talmy (1978)

² The term 'fundamental attribution error' is normally used to refer to the tendency of observers to attribute an agent's wrong behavior as resulting from their fundamental character traits. Here, I re-purpose this term for the present argument. However, as will be made clear, it still retains much of its original character.

³ Here I will use the term 'emotion-dispositional' in order for the concept to stand out against other properties, such as red and bitter which are considered to be 'response-dependent' or secondary properties. See (D'Arms and Jacobson (2007)). Emotion-dispositional properties are response-dependent but have the propensity to elicit emotional reactions whereas standard response-dependent properties, such as 'red' typically do not.

and Croft and Wood (2000). It will be shown that an understanding of construal phenomena and how they affect language is instrumental in accounting for the prevalence of attribution error in emotional and moral language.

2 Philosophical Foundations

The view of moral judgments defended here can be referred to as a strong emotionist account of moral judgements. A strong emotionist theory is one which makes both the metaphysical claim that ‘moral properties are essentially related to emotions’ and the epistemic claim that ‘moral concepts are essentially related to emotions’ (Prinz 2007; p 14–16). Claims of the former metaphysical kind can be seen in theories which liken moral properties to response-dependent properties such as colour (McDowell 1985). The latter epistemic kind can be seen in Allan Gibbard’s work (Gibbard 1990, 2006) where he defines moral concepts in emotional terms. Understanding such concepts, therefore, consists in recognizing the potential emotional responses caused by certain acts. Gibbard (2006) states that moral concepts of ‘wrong’, for example, can be defined as those which lead to the elicitation of such emotions as guilt on the part of the transgressing agent, and which will be seen as *reprehensible* on the part of the observer; that for an action to be considered wrong, it must be one for which, at least, a guilty response is *warranted* by the agent. Hence, we have an account of morality which holds that moral concepts can be defined in terms of the actual and potential emotional dispositions elicited:

I say that the basic narrowly moral concept is being blameworthy or reprehensible. That an action is reprehensible just means that reprehension over it on the part of others and guilt over it on one’s own part are warranted. (Gibbard 2006).

A strong emotionist theory is one which combines both metaphysical and epistemic claims outlined in (i) and (ii) above. The work of Jesse Prinz (2007; 2006) provides an example of a strong emotionist theory, but other similar theories are to be found in sensibility theories (Darwall et al. 1992; McDowell et al. 1985). Sensibility theories, make the claim that moral properties are ‘response-dependant’. An analogy is often made with colour, the perception of which is supposed to be dependent upon the agent experiencing the property (D’Arms and Jacobson 2007). In McDowell’s words, moral properties are to be understood as ‘qualities not adequately conceivable except in terms of certain subjective states’ (Mc Dowell 1985, p.136). Hence, we can see moral properties as defined by, their propensity to elicit certain moral emotional reactions such as guilt, shame, remorse or compassion. Moral concepts, therefore, consist in the predisposition to experience these emotional reactions, and it follows from this that moral judgements can be defined as judgements about the propensity of an action to elicit these emotional responses.

Consequently, I claim, moral utterances of the form ‘*x* is *P*’ express belief-like states and are truth-apt propositions. We avoid an error theory as moral predicates,

thus understood, have as their extension certain forms of conduct. We could say that conduct perceived as having a certain ‘character’—a particular character recognized as having the propensity to elicit certain emotional responses in agents and observers—is the extension of moral predicates and, accordingly, propositions containing such predicates, do refer properly to states of affairs. The truth conditions of such judgements, however, are relative to a particular moral community—a group that shares the same values. The moral community is the set of all people who share a particular moral judgement.⁴ The truth of a moral proposition will, therefore, be relative. There may, however, be some judgements which are universally accepted if shared by all humans, but this would be a contingent fact. Such a theory, therefore, can be said to be *minimally* realist (Van Roojen 2015, p 116). However, although we avoid a semantic error theory, we do not avoid a conceptual one, as I claim that speakers mistakenly perceive moral properties as objective mind-independent and non-reducible properties.

Emotionist or sensibility theories have, of course, been challenged and one such objection comes from Harman in response to Gibbard (2006) who, as seen above, proposes a guilt-focussed account of moral concepts. Harman (2009) warns against positing a definitional explication of moral concepts, especially in the case of guilt, on the grounds that it is possible for an agent to commit a morally reprehensible action and yet not feel guilty. Harman argues that a necessary link cannot be true a priori and that moral concepts cannot be defined in terms of guilt feelings due to the fact that it is possible to be motivated to act morally by other feelings such as empathy or compassion. Harman’s objection works against guilt-focussed accounts of moral concepts but not with emotionist theories in general and I will claim, due to these considerations, that an emotionist account of morality must encompass all *moral* emotional states.

Prinz (2006, 2007) has defended a strong emotionist thesis by pointing to the wealth of empirical data from psychology and neuroscience as support, and it is my view that, in light of the empirical data, such an account should indeed be taken seriously.

3 Psychological Evidence

It appears, intuitively, that a connection between emotion and morality holds, at least as a causal relation. Generally, emotions can be seen as a psychological and physiological reaction to a personally-significant stimulus:

[Emotion is] a complex reaction pattern, involving experiential, behavioral, and physiological elements, by which an individual attempts to deal with a personally significant matter or event. The specific quality of the emotion (e.g., fear, shame) is determined by the specific significance of the event. (*Emotion—APA Dictionary of Psychology*).

⁴ This definition needs to be formalised but this rough characterisation will suffice for our current purpose.

We can see that such a causal relationship between stimulus and reaction holds, *prima facie*, for moral emotions too. For instance, if I have acted in a way which you and I both judge to be wrong, then I expect you to respond with anger, and, furthermore, I expect to feel guilty for my transgression (if not for the action in itself, then at least for upsetting you.) We do, then, for the most part, appear to have predictable emotional responses to the transgression of moral rules from both agent and observer. Armchair speculation about responses to what we might broadly call *moral stimuli* do not yet, however, constitute the grounds for a compelling emotionist argument, but they do provide us with a jump-off point; one which is strongly supported by empirical evidence.

The link between emotion and morality is widely accepted in the psychological literature, and we can even see anger, which is typically understood as being a non-moral emotion, profiled in vaguely moral terms; as an emotion which arises when, injustice or transgression is detected. As Prinz and Nichols (2010) point out:

In the recent moral psychological literature on anger, the familiar characterization of the profile of anger is that it's caused by a judgment of transgression and it generates an inclination for aggression and retributive punishment [...] Much of the evidence for this profile comes from work in social psychology over the last two decades. (Prinz and Nichols 2010).

Apart from emotional profiling and categorization, the psychological literature seems to give strong support for a link between morality and emotion which shows emotion to be more than simply an output or reaction. There are two prominent ways in which emotion has been connected to morality in the psychological and neuroscientific literature: (i) in interfering with, or guiding, moral decision making, and (ii) in being a constituent part of the moral judgement-making process. I will offer an overview of the work which documents these phenomena in turn.

An important and oft-cited example of how emotional centres of the brain affect moral decision making can be seen in the work of Greene and colleagues (2001) who observed neural activity in two distinct decision-making systems when subjects contemplated moral dilemmas. Greene et al. used fMRI scans to identify the areas of the brain used when making decisions in moral dilemmas and found clear activity from emotional centres of the brain when participants contemplated the scenarios. This was coupled with reduced activity in the pre-frontal lobe and structures associated with rational and logical thought, when participants were asked to contemplate more 'personal' moral dilemmas (Green 2001). This work is widely interpreted as showing that moral decision-making processes are, at least in part, influenced by emotional stimuli,⁵ and seems to provide evidence for a connection between morality and emotion which is more than simply reactional.

In addition, there is evidence which suggests that emotional states actually guide or antecedently influence people's moral judgements. Isen and Levin (1972), for instance, have shown that inducing positive emotional states in subjects, makes

⁵ Greene et al.(2001) even suggests that this evidence show that classic consequentialist and utilitarian ethics require rational thought processes whereas deontological thinking is driven by the non-rational emotional structures of the brain.

them more likely to help others, suggesting that emotional priming encourages prosocial behaviour. In another study, where people were asked to evaluate the wrongness of certain actions, participants who were sitting at a dirty desk gave consistently higher ‘wrongness ratings’ as compared with their counterparts who were sitting at clean desks (Schnall et al. 2008). Schnall et al.’s observations show the corollary of Isen and Levin’s; that the moral decision-making process can also be affected negatively when emotions such as disgust are present. This has also been corroborated by similar observations made by Wheatley and Haidt who induced disgust in subjects hypnotically (Wheatley and Haidt 2005). Of course, these findings do not provide conclusive evidence but do give compelling empirical support for a link between emotion and morality at the level of decision-making, such that emotion is shown to be more than just a reaction to moral stimuli, but as also being actively involved in the moral decision-making process antecedently influencing moral judgement.

In addition to the work mentioned above, there are further empirical studies which give more direct support to the current argument about moral concepts, and these studies examine moral judgement in subjects without prior emotional priming. Before I move on, it is important to note that the examination of moral judgement, as opposed to moral *decision-making processes*, at the neurophysiological level, has a more direct bearing on the emotionist claim—that moral concepts can be reduced to emotional concepts. This, of course, is due to it being imperative to the making of a judgement that one understands the concepts employed by the judgement; that to be able to make a moral judgement, one must first grasp moral concepts. It can be concluded, then, that if making moral judgements requires the use of emotional structures in the brain, then moral concepts are likely to be emotional in nature. Evidence for this could come in two forms: evidence showing the involvement of emotional brain structures in the process of moral judgement making, and, conversely, evidence showing that individuals with emotional impairment fail to make moral judgements or grasp moral concepts. Both types of data are, in fact, available and I will briefly mention some examples here.

Studies by Moll et al. (2001, 2002) found that when participants made judgements about moral statements such as ‘we break the law when necessary’, distinct regions of the brain were activated including the frontopolar cortex, medial frontal gyrus, right anterior temporal cortex and the cerebellum, which is normally associated with emotional regulation (Schmahmann and Caplan 2006), that were not active when processing non-moral statements such as ‘stones are made of water’.

Even more promising, perhaps, is the data gathered from investigations into psychopathy, and individuals with antisocial personality disorder, which has shed considerable light on moral judgement-making. The results of such studies are indeed pertinent to the current argument. Antisocial personality disorder is characterized by the APA as ‘the presence of a chronic and pervasive disposition to disregard and violate the rights of others. Manifestations include repeated violations of the law, exploitation of others, deceitfulness, impulsivity, aggressiveness, reckless disregard for the safety of self and others, and irresponsibility, accompanied by lack of guilt, remorse, and empathy’ (*Antisocial Personality Disorder – APA Dictionary of Psychology*). As can be seen from this delineation, a

lack of prosocial moral emotions such as guilt, remorse and empathy are outstanding features of psychopathy and so we might expect such individuals, following our claims about the emotional nature of moral concepts, to lack competence with moral concepts; that their emotional impairment would mean moral impairment. Indeed, it has been shown, in line with this thought, that individuals with antisocial personality disorder fail to grasp the so-called moral/conventional distinction, meaning that they are not able to distinguish moral transgressions, such as stealing, from conventional non-moral transgressions such as talking in class or not waiting in line (Blair et al. 2016; Blair and Cipolotti 2000; Liao 2016). Interestingly, further research into the neurotransmitter serotonin, which has been found to be lacking in individuals who display characteristics typical of psychopathy such as callous-unemotional (CU) traits (Dolan and Anderson 2003; Soderstrom et al. 2001, 2003), has shown that serotonin function covaries positively with prosocial behaviours and negatively with antisocial behaviours, suggesting that emotional neurochemistry might play an intrinsic role in moral appraisal and action.

In sum, there is a significant body of empirical evidence which shows the involvement of emotional regions of the brain in various moral thinking processes and, more importantly for the current argument, that emotional states are not simply reactions in response to moral transgressions but also aid and influence moral judgements antecedently. Furthermore, in the case of psychopathy, emotional impairment correlates with an inability to make the moral/conventional distinction which is indicative of the inability to grasp certain moral concepts. Such evidence, although not conclusive, I believe licenses much optimism for a strong emotionist reading of morality.

4 Moral Emotions and Perspectives

I have, so far, been using the term ‘emotion’ in a rather broad sense. Before building my defence of emotionism, it is necessary to give a more detailed explication of the emotional import that the current theory proposes is present in moral judgments. In order to do this, we need to make reference to three distinct groups of *moral* emotions. Firstly, there are *pro-social* emotions such as sympathy, empathy and compassion, which are assumed to promote positive moral behaviour. Secondly, there are *self-blame* emotions such as guilt and shame which are directed inwardly and suffered by the agent upon transgression. And, thirdly, *other-blame* emotions which include contempt, anger and disgust which are experienced by the observer and are directed outwardly at the transgressing agent. Understanding these distinct emotions and the divergent *perspectives* from which they arise is important for the foregoing argument. The three categories of pro-social, self-blame and other-blame form a triad of possible perspectives from which a moral action can be appraised: observer (agent-focussed), agent (self-focused) and observer (victim-focussed). Pro-social emotions emerge from an observer’s victim-focussed appraisal and are assumed to motivate morally good actions. Self-blame emotions represent negative self-focussed appraisal of the agent from the perspective of the agent themselves. These are assumed to motivate the avoidance of, and are occurrent from, actions

perceived by the agent as being morally wrong. And, other-blame emotions represent the perspective of the agent-focussed observer and are assumed to motivate retaliation, retribution and blame of the agent. With the distinct moral emotions, we have three divergent perspectives from which moral emotion can be elicited and these, in turn, provide separate emotional reactions.⁶

With an understanding of the moral emotions and their divergent perspectives in-hand, we can now define moral judgements in virtue of their propensity to elicit these specific emotional responses. We can, therefore, define moral judgements as:

Moral judgements are those judgements referring to any conduct ϕ perceived as having the propensity to elicit moral emotional responses in themselves and observers, in certain situations. Such that, a judgement α is a moral judgement *iff* the conduct referred to by α is perceived as having the propensity to elicit moral emotional states in agent x and observer y in S .

This formulation of moral judgements follows, as shown above, from the metaphysical claim that moral properties consist in the propensity of an action to elicit moral emotional states, and the epistemic claim that the disposition to experience such moral emotions is to grasp moral concepts. Actions, therefore, will be perceived as having 'moral' properties (rightness/wrongness) if they are understood as having the propensity to elicit the appropriate moral emotional responses in agents and observers. Hence, sincerely assenting to a moral judgement consists in a recognition that a particular act has the propensity to elicit moral emotions.

In the following section, I will argue that moral properties and concepts, understood thusly, are analogous to myriad other dispositional properties and their corresponding concepts which we ascribe in error, and that moral properties appear to be a subset of such properties.

5 Emotion-Dispositional Properties

The wealth of empirical evidence garnered from the work of moral psychologists does indeed improve the soundness of our claims, but although it suggests heavy influence from emotional structures on higher-level cognitive functioning, we cannot use it to directly support either the metaphysical or the epistemic claims posited here. It is, then, in addition to the empirical data, necessary to provide observations which will bolster our claims about moral properties and concepts more directly.

As briefly alluded to above, the emotionist takes a view of moral properties to be broadly in line with sensibility theories (D'Arms and Jacobson 2007). Sensibility theories come in many flavours but have their origins in the work of David Hume.⁷

⁶ It should also be noted that moral transgression might produce collateral non-moral emotions such as fear of repercussions and sadness in the agent, thus providing further motivation to avoid transgressive behaviour.

⁷ I invoke Hume simply as a historical reference point and am not claiming that the current understanding of dispositional properties is 'Humean' at all.

Such readings often make use of an analogy between moral properties and colour properties. Colour, it is said, is not a mind-independent objective property but is rather a perceptual one—merely the subjective experience of the way human perceptual systems interpret light in the optic array. Redness, then, is a ‘dispositional’ property, and not a mind-independent one, like mass for instance. It is held that colour concepts, therefore, consist in being able to experience these dispositions. This, it is often suggested, is how we should understand moral properties too—as being dispositional in the way that colours are. We do indeed have a helpful analogy here with colour, and it has clear pedagogical merit. But this, I fear, is the full extent of its efficacy. This is because colour properties are in fact very distinct from the moral properties of rightness and wrongness, in some important respects. Colour is visually perceived; moral properties are not. Colour is emotionally inert; moral properties are not. The analogy with colour, then, is a useful tool for grasping the concept of dispositional properties but has little argumentative force when explaining moral values.⁸ There are, however, alternative analogies to be drawn from other groups of dispositional properties that mimic more closely our understanding of moral properties. These, I claim, provide not only a more effective analogy but can also exert some explanatory and argumentative leverage on the phenomenon currently under analysis. One such example can be seen in the case of ‘annoyingness’. As with colour, we routinely ascribe annoyingness to extramental objects—most commonly people—and states of affairs. We are, in fact, very prone to saying things such as ‘he is annoying’, ‘this song is annoying’ or ‘the situation is annoying’ and in doing so we attribute the property of ‘annoyingness’ to people, objects and situations as if it were an objective and intrinsic part of their physical existence; as if it were ‘out there’. However, this is incorrect. Being annoyed, of course, is an emotional disposition; it consists solely in our being annoyed. Hence, being annoyed is a subjective emotional state which we misinterpret and wrongly attribute as being a property of extramental entities when uttering such things as ‘Smith is annoying’, for instance. I will call this an *attribution error*. Smith’s actions or mannerisms *do* have a particular character which induces an annoyed state in us but this character is not objective and is dependent upon the reaction this character elicits in the observer; it is a response-dependent property. In fact, what we really describe when we utter ‘Smith is annoying’ is ‘Smith makes me feel annoyed’ which is felicitous and, despite its uncommon and awkward form, correctly assigns ‘annoyingness’ to the speaker as a subjective emotional response. This makes our utterance of ‘Smith is annoying’ a mistake because being annoying is not something that anything can be, in and of itself. This mistake, however, is not simply a linguistic one, but also a sincere conceptual one and it shows through in the surface grammar of our utterance.⁹

⁸ See Railton (2003) for a further objection to the analogy from colour.

⁹ In *Inventing Right and Wrong* (1977: p 42) Mackie puts forward a similar example with the property of being disgusting. Mackie, unfortunately, spends little time on this line of thought and his example of disgust is also subtly different from annoyingness as being disgusted by the sight of vomit, the smell of rotting flesh or the taste of a fungus may well be an automatic or hard-wired reflex.

As per the emotionist view, we understand such utterances not as statements describing the properties of extramental objects, but as statements about one's own emotional state or disposition in response to those entities or actions. However, although it takes just a moment's reflection to realize that annoyingness is in the eye of the beholder, I claim that, under normal circumstances, people *do* in fact believe what their utterances pretend to be; that Smith *is* annoying; Smith *has* that property. If this were not the case, why would we not more commonly say 'I am annoyed' at witnessing Smith's actions or 'Smith's actions annoy me?' I believe that we make this mistake sincerely and that 'Smith is annoying' is not simply a more convenient grammatical form. Such utterances should be taken as indicative of the way people perceive the world. Evidence in support of this claim can be found when looking at normal conversation patterns. Upon declaring that someone is annoying, we normally expect our friends to agree. If they do not, we attempt to convince them 'you don't think so? But he *really is!* He is annoying because...'. But, no amount of justifying will convince you that he is annoying, if, that is, you don't already think so, because no amount of explaining is likely to change your feelings and dispositions. Such examples, uncover a systematic attribution error that is both linguistic and conceptual in nature.

It is indeed plausible that moral predicates function in the same way and I propose that this is how they are best understood. When comparing the conceptual features and linguistic functioning of moral discourse with emotion-dispositional discourse we find some striking similarities. For example, when one says 'that is wrong' the surface grammar of the utterance suggests that one intends to ascribe a mind-independent property to an extramental entity which one believes to be true—as is the case with annoyingness. And, furthermore, we are surprised if others do not agree with our moral judgement and are likely to make an attempt at convincing them that our view is correct—as with annoyingness. It is also the case that—as with annoyingness—no amount of convincing will sway people if they are not already similarly predisposed.

If my analysis is right and moral properties do turn out to be emotion-dispositional properties like annoyingness, this all amounts to a systematic error in our language and with our understanding of these properties. This is the sort of conclusion that many want to avoid. But, however displeasing this conclusion might be, the case of annoyingness gives us reason to believe that it is, at least, possible for us to be mistaken in such a way. Moreover, due to the clear links between morals and emotions, moral properties can plausibly be understood as being a special case of the general emotion-dispositional properties just described. To put this characterization more succinctly, in the case of annoyingness, I think it fairly uncontroversial to claim that:

1. Annoyingness is an emotion-dispositional property and is thus constituted solely by the relevant emotional response to a stimulus.
2. Having the disposition to be annoyed is a necessary condition for grasping the concept of annoyingness.

3. Annoyingness judgements of the form ‘ x is annoying’ are those judgements referring to entities perceived as having the propensity to elicit the relevant emotional response of annoyance in observers.

The analogy of annoyingness might only be slightly better than the classic colour analogy if it were not for the fact that this is only one example in a wide range of emotion-dispositional properties which we routinely attribute in error. Taking a moment to reflect, it becomes clear that such cases are frequent in English.¹⁰ When we use participle adjectives to say that something, or someone, is annoying, frustrating, boring, depressing, scary, confusing, sexy, disgusting, deplorable, disappointing, stressful, relaxing, nerve-wracking, frightening, awesome, inspiring, shocking or infuriating¹¹ we make the same mistake; we are falsely, but sincerely, attributing a subjective emotional disposition to an extramental entity. In light of this, I claim, there can be no in-principle objection to systematic attribution error in the case of moral judgements and, moreover, that the principle of charity starts to seem far too generous. The cases mentioned above are not simply perceptual errors in the way that colour is supposed to be, but they are also all cases of ascribing an emotion-dispositional property as if it were a mind-independent one.

So far, I have shown that various emotion-dispositional properties exist which are not purely extramental features of the world. Such properties consist in emotional reactions to relevant stimuli, and being predisposed to experience these emotional reactions is a necessary condition for grasping emotion-dispositional concepts. I have further claimed that moral properties and concepts mimic the conceptual and linguistic behavior of emotion-dispositional properties and concepts, giving us reason for optimism about an emotionist reading of morality, and supplying us with a structure on which to model moral properties, concepts and judgements. If moral properties are emotion-dispositional properties then from claims 1, 2 and 3 above, 1*, 2* and 3* follow

- 1* As emotion-dispositional properties, ‘rightness’ and ‘wrongness’ are constituted by the propensity of an act to elicit moral emotional responses in agent and observer.
- 2* Being predisposed to experience the relevant emotional responses is a necessary condition for grasping the concepts of rightness and wrongness.
- 3* Moral judgements of the form φ ‘is right/wrong’ are those judgements referring to conduct φ perceived as having the propensity to elicit moral emotional responses in agents and observers, in certain situations.

I have claimed that moral properties, concepts and judgements can be understood as being a subset of the wider set of emotion-dispositional properties and can be labelled as *moral emotion-dispositional* properties. There is, however, an immediate issue with our theory of moral emotion-dispositional properties. Given that this

¹⁰ Other languages need to be considered. It appears that in Thai language emotion-dispositional properties are attributed in error in a similar manner.

¹¹ This is not an exhaustive list but I think it makes my point.

reading is true, we are convicting humans of routinely making systematic conceptual and linguistic errors. It remains to be explained just *why* we do this.

6 Through the Lens of Cognitive Linguistics

We are obliged, *ex hypothesi*, to tackle the now looming question of ‘*why?*’. Given that the above account is true, why is it that we make these sorts of systematic attribution errors? Why do we act *as if* emotion-dispositional properties were objective, when they are not? A more instructive way of phrasing this question would be “What is the difference between uttering ‘Smith is annoying’ and ‘I feel annoyed with Smith?’”. Being able to give a plausible response to this question will be needed to bolster our theory, and this is the task I will now undertake. In order to do this, it will be necessary to take a short but enlightening diversion into the realm of Cognitive Linguistics.

One of the foundational assumptions underlying the cognitive linguistic framework is that linguistic structures are the upshots of, or mirror, more fundamental conceptual structures, as Croft and Cruse (2004) put it:

[L]anguage is not an autonomous cognitive faculty. The basic corollaries of this hypothesis are that the representation of linguistic knowledge is essentially the same as the representation of other conceptual structures, and that the processes in which that knowledge is used are not fundamentally different from cognitive abilities that human beings use outside the domain of language. (Croft and Cruse 2004).

It is asserted that language is emergent; it is a function of more general cognitive processes and conceptual structures. Such an assertion is grounded in the fact that no specific structure in the brain has been found to be responsible for language (Anderson and Lightfoot 2002). Cognitive linguists argue that general cognitive and conceptualization processes underlie not *only* semantic representation but that syntax, morphology, and phonology are also generated by and grounded in general cognitive functioning.

Some theories in Cognitive Linguistics have taken inspiration from, and made use of, theories from other areas of cognitive psychology. One such example is borrowed from Gestalt psychology. The Gestalt theory of perception was first concerned with identifying and understanding the various ways in which the brain orders visual input (Wertheimer 1923). This theorizing led the Gestalt psychologists to posit various ‘principles’ of perception. One prominent Gestalt principle is that of *figure and ground*. It is apparent that visual perception is structured by underlying and unconscious cognitive processes. Figure and ground organization is one such process. This is the process whereby attention can be selectively focused on an object (the figure) allowing it to exist in the perceptual foreground and stand out, or appear separated from, its surroundings (the ground). Although first elucidated in the context of visual perception, linguists have found similar principles to be evident in *construal processes* which appear to structure language in an analogous way (Talmy 1975). From visual organization to grammatical organization, the same

construal principles seem to be at work. For instance, we can see clear cases of figure and ground construal in the passive and active voices. Take the following examples:

- a. 'Jones cleaned the windows'
- b. 'The windows were cleaned by Jones'

Both represent the same state of affairs but there is a difference at a psychological level which can be explained by reference to a figure and ground construal process. In *a* 'Jones' is the figure, or the focus of attention, whereas in *b* 'the windows' become the figure against the background of Jones cleaning, and this represents a change of focus. The figure is the more salient part of the scene being described and so the grammatical form chosen by the speaker is indicative of the way they perceive the situation. We can also recognize the figure and ground construal at work in the case of annoyingness mentioned above. Take our now familiar phrases:

- c. Smith is annoying.
- d. I am annoyed with Smith.

In *c* the figure is 'Smith'—'Smith' appears as the focus of the utterance. In *b* we see the opposite arrangement where 'I', the speaker, is profiled as the figure and 'Smith' as the ground. When analyzing the difference between the two phrasings in terms of figure and ground construal, we see how the focus of each utterance is shifted, but also note that the attribution error disappears from *c* to *d*. The fact that *c* – the phrase which makes the attribution error – is a common phrasing is telling and provides us with further reason for optimism. It is possible to show that phrases of the form in *c* are more frequently found in English than the paraphrase *d* with corpus data. Several English web-based corpora show that the phrase 'am annoyed' appears far less frequently than the phrase 'is annoying' in a range of contexts. For instance, the I-Web corpus showed 'am annoyed' at a frequency of 572 as compared with 'is annoying' at 11,676. And, in the Corpus of Contemporary American English 'am annoyed' appears 52 times versus 'is annoying' which appears 390 times (Davies 2008, 2018). Such results are indeed to be expected if my present analysis is correct. The fact that the form used in *c* appears more frequently is significant for two important reasons. Firstly, it vindicates our claim that attribution error occurs and is widespread. Secondly, it gives us further insight into the way in which humans generally perceive the world, as per the Cognitive linguistic framework. In uttering *c*, the speaker places 'Smith' as the figure—the more salient concept. I claim that this happens naturally as 'Smith' is perceived as the stimulus which gives rise to my annoyed state—recall how emotions are reactions to stimuli. So, although Smith is not intrinsically annoying—as the surface grammar suggests—Smith *did* cause me to be annoyed. Smith is therefore the stimulus and, hence, appears as more cognitively salient. This salience is reflected in the surface grammar of the utterance by construing the stimulus (in this case 'Smith') as the figure of the utterance. This fix of attention is shown in the common phrasing 'x is P' and gives a plausible explanation as to why we naturally and intuitively utter sentences of this form more frequently than alternatives. From an anthropological

perspective, it is possible that humans could have evolved a propensity to focus on and attend to the triggers of emotional reactions before introspecting. The foregrounding of the emotional stimulus provides possible support for this assertion and gives us an explanation as to why we more commonly utter phrases of the form ‘Smith is annoying’, although it is in error. It is easy to see how a propensity to exhibit a cognitive bias of this sort would have been a potentially effective survival tool.

Cognitive construal operations such as figure and ground structure human thought and language. Sentence grammar is therefore indicative of the way we conceptualize states of affairs. The same analysis from above can be run on moral utterances. In moral utterances of the form ‘ ϕ is wrong/right’, ϕ is foregrounded as the more salient figure of the utterance. This suggests that it is the focus of the speaker’s attention and, therefore, is seen as an emotional stimulus. Hence, the figure and ground construal gives us reason to believe that ϕ -like moral acts are identified (possibly subconsciously) as emotional stimuli and are, therefore, deserving of our attention. When viewed in such a light, attribution error is not an anomaly that needs to be explained away, but rather an expected outcome grounded in underlying human conceptualization processes.

7 Emotional Import

As defined here, moral properties have the propensity to elicit certain moral emotions in the agent and the observer. This makes them a special case of emotion-dispositional properties as they elicit emotions from multiple perspectives. Hence, we need to explain, now, the differences between emotion-dispositional properties and moral emotion-dispositional properties (rightness/wrongness).

Recall that with emotion-dispositional properties, there is an available paraphrase which represents an alternative construal and that has the upshot of not committing the attribution error as seen below:

- e. Smith is annoying. (Common phrasing with attribution error)
- f. I am annoyed with Smith. (Less common paraphrase without attribution error)

In the case of moral judgements, there seems not to be a similar paraphrase available:

- g. ϕ is wrong
- h. #I feel wrong about ϕ

Our analogy seems as if it might break down here. However, lack of a paraphrase is to be expected if our current definition of moral emotion-dispositional properties is correct. Standard emotion-dispositional predicates contain only one emotional perspective and so can be used in the first person, as seen in *f* above. Remember, however, that moral concepts such as ‘rightness’ and ‘wrongness’ include, as part of their definitions, more than one emotional perspective and so cannot be used in the first person (*h.*). Hence, judgements of the form ‘ x is wrong’ can be used easily as

‘blanket’ or universal judgements as they allude to the potential emotional responses from divergent perspectives (observer (agent-focused), observer (victim-focused) and agent (self-focused)). In contrast ‘Smith is annoying’ describes one single emotion-dispositional property, from the one perspective of the observer, and so can be felicitously construed in the first person as a subjective judgement about one’s own emotional state. It would seem, then, that given this dynamic, and given our definition of moral emotion-dispositional properties, that a first-person error-free construal of moral utterances might not be possible. However, moral utterances *can* be construed in the first person, thereby avoiding the attribution error, but this construal will be limited to representing only one perspective, and each perspective will make clear the emotional import contained within moral judgements. The judgement, ‘ ϕ is wrong’ can be reconstrued in a number of ways, each way representing a unique emotional perspective:

- i. I am disgusted with ϕ (Observer’s perspective—agent-focused)
- j. I feel sorry for p because of ϕ (Observer’s perspective—victim-focused)
- k. I feel guilty about ϕ , (Agent’s perspective—self-focused)

We can see from *i*, *j* and *k* that moral judgements containing moral emotion-dispositional properties can be reduced to emotion-dispositional properties, as seen in the above examples, and such utterances do not commit the attribution error or use moral predicates. The fact that this is possible gives further support for the thesis that moral properties are reducible to emotion-dispositional properties. We can even see this more explicitly in common parlance where moral opinions are often couched in explicitly emotional terms; we can, and do, talk about moral acts without using moral predicates. For example, it is often said of immoral actions that that they are heinous, deplorable, disgusting, contemptible, loathsome, hateful or detestable.¹² The phrase ‘ ϕ is deplorable’, for example, has a clearly moral tone, despite the use of an emotional predicate, as opposed to a moral one. It is easily inferred pragmatically that this expresses a moral judgement. When analyzing these examples, we find an even tighter analogy with emotion-dispositional properties where a non-moral emotional predicate is being used to convey moral indignation. Hence, in such cases, we have no need to draw a mere *analogy* between moral properties and emotion-dispositional properties, as the former *is reducible to* the latter. In using adjectives like deplorable, heinous or detestable to describe morally wrong acts, we clearly paint morality with an emotional pallet.

¹² Such terms also appear to be used more frequently in sentences of the form ‘ x is P ’ which commit the attribution error. For instance, in the I-Web online corpus, ‘is contemptible’ appeared 180 times vs ‘feel contempt’ at 115; ‘is disgusting’ 4832 vs ‘am disgusted’ at 1124 and ‘is deplorable’ appeared 801 times as compared with ‘I deplore’ at only 410. The relative frequency of such forms shows the stimulus as psychologically salient.

8 In Conclusion

Above, I have argued for a strong emotionist reading of moral properties, concepts and judgements. I, firstly, presented psychological evidence which shows a clear connection between moral judgement making and emotion, implying that a strong link between emotion and morality is evident. Secondly, I claimed that such evidence is best explained by understanding moral properties, and therefore moral concepts and judgments, as emotionally constituted. I, then, offered an analogical argument of non-moral emotion-dispositional properties drawing clear parallels between these two kinds of properties in showing how they function in language and discourse. Moral emotion-dispositional properties (rightness/wrongness) are a subset of the wider set of emotion-dispositional properties which contain properties such as ‘annoyingness’. Accordingly, moral judgements of the form ‘ ϕ is wrong/right’ are defined as:

those judgements referring to any conduct ϕ perceived as having the propensity to elicit moral emotional responses in themselves and observers, in certain situations. Such that, a judgement α is a moral judgement *iff* the conduct referred to by α is perceived as having the propensity to elicit moral emotional states in agent x and observer y in S .

I claim that moral judgements, like emotion-dispositional judgements, are best understood as sincerely intended truth-apt propositions which express belief-like states. Thus understood, their referents are certain forms of conduct perceived as having a moral ‘character’ having the propensity to elicit moral emotional responses in agents and observers. Judgements, therefore, are relative to a particular moral community and have as part of their make-up subjective emotional import. Such judgements, however, commit an attribution error. They admit of a perceptual mistake—the mistake of attributing a response-dependent or emotion-dispositional property as an objective mind-independent one.

Finally, I claimed that our natural propensity to focus attention on emotional stimuli possibly leads humans to make such errors; to wrongly attribute such dispositional properties to, emotional stimuli. In support of this claim I offered an analysis of language from the perspective of Cognitive Linguistics. Analyzing construal operations apparent in moral utterances shows that the most common grammatical form when using emotion-dispositional predicates, x is P , foregrounds x —the emotional stimulus. This suggests that the emotional stimulus it is more psychologically salient. This salience is conceptually significant and is apparent in the grammar of the utterance through the figure and ground construal process. Moral and emotional judgements of the form ‘ x is P ’ are relatively frequent in comparison to passive grammatical forms which represent the emotion (as opposed to the stimulus) as the salient ‘figure’ of the utterance, and which do not commit the attribution error. This suggests a natural predisposition to focus attention on the stimulus, labelling it with an emotional tag, rather than to immediately focus on one’s emotional state.

When viewed this way, our utterances, although mistaken, can be seen as reflective of the way we perceive the world. Underlying conceptual structures augment and determine linguistic structures and do so with the aim of communicating socially significant and useful concepts and not merely at expressing true propositions and, thus, in this process we find an expected error.

Acknowledgements I am grateful for the comments made by anonymous reviewers which have helped to improve the arguments presented here.

References

- Anderson SR, Lightfoot DW (2002) *The language organ*. Cambridge University Press, Cambridge.
- Antisocial personality disorder – APA Dictionary of Psychology. (n.d.). Retrieved September 19, 2020, from <https://dictionary.apa.org/antisocial-personality-disorder>
- Blair RJR, Cipolotti L (2000) Impaired social response reversal. A case of “acquired sociopathy.” *Brain* 123(6):1122–1141. <https://doi.org/10.1093/brain/123.6.1122>
- Blair RJR, Hwang S, White SF, Meffert H (2016) Emotional learning, psychopathy, and norm development. In: *Moral brains*. Oxford University Press, Oxford, pp 185–202. <https://doi.org/10.1093/acprof:oso/9780199357666.003.0009>
- Croft W, Cruse DA (2004) *Cognitive linguistics*. Cambridge University Press. <https://doi.org/10.1017/CBO9780511803864>
- Croft W, Wood E (2000) Construal operations in linguistics and artificial intelligence. In: Albertazzi L (ed) *Meaning and cognition: a multidisciplinary approach*. John Benjamins, Amsterdam, pp 51–78
- D’Arms J, Jacobson D (2007). Sensibility Theory Projectivism. <https://doi.org/10.1093/OXFORDHB/9780195325911.003.0008>
- Darwall S, Gibbard A, Railton P (1992) Toward Fin de siècle ethics: some trends. *Philos Rev* 101(1):115. <https://doi.org/10.2307/2185045>
- Davies M (2008) The corpus of contemporary American English (COCA): million words, 1990–present. <https://www.english-corpora.org/coca/>
- Davies M (2018) The 14 Billion word iWeb Corpus. <https://www.english-corpora.org/iWeb/>
- Dolan MC, Anderson IM (2003) The relationship between serotonergic function and the psychopathy checklist: screening version. *J Psychopharmacol* 17(2):216–222. <https://doi.org/10.1177/0269881103017002011>
- Emotion – APA Dictionary of Psychology. (n.d.). Retrieved September 19, 2020, from <https://dictionary.apa.org/emotion>
- Gibbard A (1990) *Wise choices, apt feelings: a theory of normative judgment*. Harvard University Press, Harvard.
- Gibbard A (2006) Moral feelings and moral concepts. In: Shafer-Landau R (ed) *Oxford studies in metaethics*, vol 1. Oxford University Press, Oxford.
- Greene JD, Sommerville RB, Nystrom LE, Darley JM, Cohen JD (2001) An fMRI investigation of emotional engagement in moral judgment. *Science* 293(5537):2105–2108. <https://doi.org/10.1126/science.1062872>
- Harman G (2009) Guilt-free Morality. *Oxford Stud Metaethics* 4:203–214
- Isen AM, Levin PF (1972) Effect of feeling good on helping: cookies and kindness. *J Pers Soc Psychol* 21(3):384–388. <https://doi.org/10.1037/h0032317>
- Langacker R (1987) *Foundations of cognitive grammar*, vol. 1: theoretical prerequisites. Stanford University Press, Palo Alto.
- Liao SM (2016) *Moral brains: the neuroscience of morality—Google books*. Oxford University Press, Oxford.
- Lyons J (1977) *Semantics*. In: *Semantics*. Cambridge University Press, Cambridge. <https://doi.org/10.1017/cbo9780511620614>
- Mackie JL (1977) *Ethics: inventing right and wrong*. Penguin Books, London.
- McDowell J (1985) Values and secondary qualities. In: Honderich T (ed) *Morality and objectivity*. Routledge & Kegan Paul, London.

- Moll J, De Oliveira-Souza R, Bramati IE, Grafman J (2001) Functional networks in emotional moral and nonmoral social judgments. *Neuroimage*. <https://doi.org/10.1006/nimg.2002.1118>
- Moll J, De Oliveira-Souza R, Eslinger PJ, Bramati IE, Mourão-Miranda J, Andreiuolo PA, Pessoa L (2002) The neural correlates of moral sensitivity: a functional magnetic resonance imaging investigation of basic and moral emotions. *J Neurosci* 22(7):2730–2736. <https://doi.org/10.1523/jneurosci.22-07-02730.2002>
- Prinz JJ (2007) *The emotional construction of morals*. Oxford University Press, Oxford.
- Prinz J, Nichols S (2010) Moral emotions. In: Doris JM (ed) *The moral psychology handbook*. Oxford University Press, Oxford, pp 111–146.
- Prinz J (2006) The emotional basis of moral judgments. *Philos Explorations* 9(1):29–43. <https://doi.org/10.1080/13869790500492466>
- Railton P (2003) Red, Bitter, Good (1998). In: *Facts, values, and norms: essays toward a morality of consequence* (Cambridge Studies in Philosophy, pp. 131–148). Cambridge University Press, Cambridge. <https://doi.org/10.1017/CBO9780511613982.006>
- Schmahmann JD, Caplan D (2006) Cognition, emotion and the cerebellum. In: *Brain: a journal of neurology* (vol 129, Issue Pt 2, pp. 290–292). Oxford Academic, Oxford. <https://doi.org/10.1093/brain/awh729>
- Schnall S, Haidt J, Clore GL, Jordan AH (2008) Disgust as embodied moral judgment. *Pers Soc Psychol Bull* 34(8):1096–1109. <https://doi.org/10.1177/0146167208317771>
- Soderstrom H, Blennow K, Manhem A, Forsman A (2001) CSF studies in violent offenders: I. 5-HIAA as a negative and HVA as a positive predictor of psychopathy. *J Neural Trans* 108(7):869–878. <https://doi.org/10.1007/s007020170036>
- Soderstrom H, Blennow K, Sjodin AK, Forsman A (2003) New evidence for an association between the CSF HVA:5-HIAA ratio and psychopathic traits. *J Neurol Neurosurg Psychiatry* 74(7):918–921. <https://doi.org/10.1136/jnnp.74.7.918>
- Talmy L (1975) Figure and Ground in Complex Sentences. *Ann Meeting Berkeley Linguistics Soc* 1:419–430. <https://doi.org/10.3765/bls.v1i0.2322>
- Talmy L (1978) The relation of grammar to cognition: a synopsis. In *proceedings of the 1978 workshop on theoretical issues in natural language processing (TINLAP '78)*. Association for computational linguistics, USA, 14–24. <https://doi.org/10.3115/980262.980266>
- Van Roojen M (2015) *Metaethics: A contemporary introduction*. a Contemporary Introduction. In: *Metaethics*. Taylor and Francis Inc, Milton Park
- Wertheimer M (1923) [1950] *Laws of organization in perceptual forms* (abridged and translated by Willis D. Ellis). In: Willis D. Ellis (ed) *A source book of Gestalt psychology*. Humanities Press, New York, pp 71–88.
- Wheatley T, Haidt J (2005) Hypnotic disgust makes moral judgments more severe. *Psychol Sci* 16(10):780–784



APPENDIX B

Received: 22 September 2021 / Accepted: 6 December 2021

© The Author(s), under exclusive licence to Springer Nature B.V. 2021



A War of Words: Dissecting the Foundational Claims of CMT

Justin J. Bartlett¹ · Sugunya Ruangjaroon¹

Received: 22 September 2021 / Accepted: 6 December 2021
 © The Author(s), under exclusive licence to Springer Nature B.V. 2021

Abstract

This work presents two theoretical challenges to Conceptual Metaphor Theory (CMT). The first argument shows CMT's foundational *Conceptual Claim*—that abstract concepts are necessarily structured by concrete concepts—entails the blurring of the literal–figurative distinction, which calls into question the legitimacy of standard methods of metaphor identification used in CMT. The second argument aims at the *Linguistic Claim*—that conceptual metaphors are necessary for metaphorical language—by showing that conceptual metaphors are neither necessary nor sufficient for linguistic metaphors and that, therefore, the existence of conceptual metaphors cannot be validly inferred from the presence of their linguistic counterparts. In light of the arguments put forward, the CMT theorist is forced to accept one of four options: (A) hold on to both the Conceptual Claim and Linguistic Claim, by adequately addressing problems presented here, (B) discard the Conceptual Claim and give up the theory, (C) discard both claims and give up the theory, or (D) accept the Conceptual Claim but reject the Linguistic Claim and abandon the methods of discovering conceptual metaphors through analysis of figurative language. I argue that the only tenable option is D.

Keywords Conceptual Metaphor Theory · Cognitive linguistics · Cognitive psychology · Abstract concepts · Conceptual structure

1 Preparing for Battle

The phenomenon of metaphor has been examined by philosophers and rhetoricians for millennia. One of the earliest accounts of metaphor comes from Aristotle who notes the profundity of metaphor in his *Poetics*, stating that:

The greatest thing by far is to be a master of metaphor; it is the one thing that cannot be learnt from others; and it is also a sign of genius, since a good meta-

✉ Justin J. Bartlett
justinbartlettswu@gmail.com

¹ Faculty of Humanities, Srinakharinwirot University, Bangkok, Thailand

phor implies an intuitive perception of the similarity in the dissimilar. (Aristotle and Heath 1996)

Metaphor does indeed have a beautiful and powerful poetic effect, but how metaphors manifest and carry their intended meaning has long been the subject of fierce debate in the philosophy of language and linguistics. Amidst the semantic, pragmatic and naïve simile accounts of metaphor (Lycan 2008), recent research in the field of Cognitive Linguistics has led to a radically different understanding of the phenomenon as more than just a quirk of language or mere literary flourish. This ‘cognitive’ approach holds that metaphor, far from being a rare ‘sign of genius’, is a natural manifestation of innate cognitive processes; not simply a device to aid conceptualization, but, rather, a function, or upshot, of the conceptualization process itself. George Lakoff and Mark Johnson first proposed this view in their work *Metaphors We Live By* (1980) by offering a novel explication of the conceptual structuring processes believed to be responsible for the manifestation of metaphor in language. Lakoff and Johnson’s Conceptual Metaphor Theory (CMT) states that metaphor, contrary to received wisdom, is extremely prevalent in language and, far from simply being a rhetorical device used only by the eloquent, is virulent and pervasive; underpinning and structuring vast swathes of our discourse. The presence of widespread linguistic metaphor is considered, from CMT’s perspective, to be evidence pointing to the existence of so-called ‘conceptual metaphors’ in the mind. Hence, the major insight of this theory is that our concepts and thoughts are metaphorically structured—not merely our language.

In the foregoing work, I challenge Conceptual Metaphor Theory by presenting analyses which expose theoretical faults in its foundational assumptions. In light of these arguments, CMT theorists are forced to reconsider their claims and associated methodologies.

I first explicate the main tenets of CMT and outline previous objections to the theory (Sect. 2), which acts as a background against which I argue that the *Conceptual Claim* of CMT, that abstract target domains are necessarily structured by more concrete source domains, engenders the blurring of the literal-figurative distinction which needs to hold if metaphor identifying techniques are to have any legitimacy (Sect. 3).

In Sect. 4, I offer a second argument which emphasizes the implausibility of CMT’s *Linguistic Claim* by undertaking an analysis that shows conceptual metaphors to be neither necessary nor sufficient for the production of linguistic metaphors. I present cases for the non-sufficient and the non-necessary in turn before showing how, in light of this, it follows that it cannot be reliably inferred from the presence of linguistic metaphors that there is any corresponding conceptual metaphor existing in the mind of the speaker, as such an inference would be invalid. Thus, it is contended that the Linguistic Claim, along with common methodologies employed by conceptual metaphor theorists, whereby linguistic metaphors are analysed in an effort to discover conceptual metaphors, are flawed and should be discarded.

Lastly, (Sect. 5) I pre-emptively acknowledge a possible counterargument to my challenge: that I am deliberately construing an unreasonably ‘strong’ version of

CMT for the purposes of developing my argument. I address this by showing that a strong CMT is indeed the version held by those working in the field and, furthermore, that it is the only psychologically and philosophically interesting version of the thesis. Attempting, I claim, to avoid the challenges presented herein by construing it as a weak thesis comes at the cost of reducing the theory to triviality.

2 Behind Enemy Lines

Conceptual Metaphor Theory has been extremely influential within Cognitive Linguistics and has been adopted by many who believe that CMT provides a way of studying the conceptual system via the analysis of figurative language. Since its conception, CMT has seen much development and elaboration (Gibbs 2017; Kövecses 2020; Lakoff 1996), but the foundational assumptions remain the same. In Lakoff and Johnson's *Metaphors We Live By* (1980), they claim that our conceptual system is 'metaphorical' in nature:

Our ordinary conceptual system, in terms of which we both think and act, is fundamentally metaphorical in nature. (Lakoff and Johnson 1980: p 3)

They go on to state the foundational tenet of their theory more concisely by saying that 'The essence of metaphor is understanding and experiencing one kind of thing in terms of another' (Lakoff and Johnson 1980). This cornerstone of CMT is still espoused in the contemporary literature (Kövecses 2010, 2020). CMT claims that the conceptual system uses a process that 'metaphorically' structures abstract conceptual domains (target domains) with more concrete domains (source domains); that in order for humans to grasp target concepts, our minds must structure them with simpler source concepts.¹ The mental constructs born from this cognitive process are dubbed 'conceptual metaphors'.

A now classic case of a conceptual metaphor is ARGUMENT IS WAR. It is claimed that the abstract domain of ARGUMENT is structured by, or 'understood in terms of', the more concrete domain of WAR. That is to say, the conceptual structure of WAR is used by our conceptual system as a foundation upon which our understanding of ARGUMENT is built in what is termed a 'cross-domain mapping'; thus, forming the conceptual metaphor ARGUMENT IS WAR. Hence, we understand arguments *in terms of* war. This psychological hypothesis is made based on linguistic evidence. Everyday utterances, it is claimed, appear to make manifest these underlying conceptual metaphors. For instance, the conceptual metaphor ARGUMENT IS WAR is made explicit, Lakoff and Johnson claim, when we say such things as 'Your claims are *indefensible*', 'He *attacked every weak point in* my argument' and 'I *demolished* his argument' (1980: p 4).

¹ Some theorists resist talking of domains in terms of abstractness or concreteness (Grady, 1997) and prefer to use simplicity and complexity instead. I will use the terms 'abstract' and 'concrete' here for convenience, but I am not making any substantial claims by using these terms and the reader may choose to substitute them for 'simple' and 'complex', respectively.

CMT, like all theories, has its critics (Bundgaard 2019; Evans 2004; Haser 2005; Leezenberg 2014; Rakova 2002; Vervaeke and Kennedy 1996; Wierzbicka 1986). One oft-mentioned criticism is that of circular reasoning (Gibbs 2017). This claim is indeed a fair criticism when aimed at Lakoff and Johnson's original work (1980). In response, later work on conceptual metaphor has sought to employ non-linguistic experimental methods of revealing cross-domain mappings which tries to circumvent the problem of *petitio principii*. There is now, as Raymond Gibbs explains:

[A]n abundance of experimental evidence to support the claim that conceptual metaphors are an essential part of verbal metaphor use. (Gibbs 2017)

A number of notable experimental studies help to bolster the claims of CMT independently (Crawford 2009; Gibbs and Colston 2012; Gibbs 1994; Horstmann and Ansorge 2011; Meier, et al. 2007a, b; Meier and Robinson 2006; Schubert 2005). Such studies do, (at least in specific cases) appear to show the psychological reality of cross-domain mappings. However, there is also experimental data to the contrary, so this point is certainly not settled (Gibbs 2014; Keysar et al. 2000; McGlone 2007, 2011; Miller et al. 2020).

In light of the empirical data, however, we should ask: If specific cross-domain conceptual mappings were to be conclusively proven as cognitively real, would theorists thereby be justified in taking *linguistic* metaphors as evidence for *conceptual* metaphors? I claim not, as the empirical data only (possibly) justifies the claim that (certain) conceptual metaphors exist as psychologically real (which I do not wish to dispute here), but does not at all justify, or render effective, the methods of linguistic analysis commonly employed. Against this background, I will proceed to show that there are further reasons to be skeptical about CMT's foundational claims and associated methodologies.

3 The Chink in the Armour

Although it is possible for the charge of circular reasoning to be avoided, as the presence of cross-domain conceptual mappings enjoys empirical support, there are still theoretical and methodological flaws present in contemporary work on conceptual metaphor. These issues, I aim to show, come from the foundational assertions made by CMT and are plausibly rooted in the fact that one of the most important metaphors in CMT has been consistently overlooked. Ironically, I am referring to the metaphor of 'metaphor' itself.

It is apparent, going back once again to Lakoff and Johnson's work, that the 'metaphor' in the moniker 'Conceptual Metaphor Theory' is itself a metaphor; that is, it is a figurative use of a term which would normally be applied to linguistic phenomena, but which is used metaphorically to refer to cognitive phenomena i.e. a cross-domain mapping or conceptual metaphor. However, I claim the conflation of the two senses of 'metaphor' has plausibly led to the further conflation of the two separate foundational postulates of CMT: *the Conceptual Claim* and *the Linguistic Claim*. These two claims can be separated and understood as follows:

Conceptual Claim Abstract target domains are necessarily structured by more concrete source domains; thus, conceptual metaphors are formed.

Linguistic Claim Conceptual metaphors are necessary for linguistic metaphors.

With these separate claims clearly delineated, we can start to analyse each of them in turn. I will begin with the Conceptual Claim before moving on to the Linguistic Claim.

It appears, upon inspection, that the Conceptual Claim of CMT has a consequence which threatens the standard methodology by which conceptual metaphors are studied: it appears to call into question the distinction between literal and metaphorical, or figurative, language. To illustrate this point, take the following passage from Lakoff and Johnson where they, again, make reference to the ARGUMENT IS WAR conceptual metaphor:

The metaphor is not merely in the words we use *it is in our very concept of an argument*. The language of argument is not poetic, fanciful, or rhetorical; *it is literal*. We talk about arguments that way because we conceive of them that way—and we act according to the way we conceive of things. (Lakoff and Johnson 1980; emphasis added)

This passage highlights a particularly important and problematic consequence of CMT's Conceptual Claim, which is encompassed by the phrase 'it is literal'. Lakoff and Johnson are justified in concluding that (according to CMT) metaphorical language is 'not poetic, fanciful, or rhetorical' but that 'it is literal' (1980), as this conclusion seems to follow from CMT's Conceptual Claim. However, although was pointed out at CMT's conception, by its founders, it seems not to have been considered problematic.

As standardly defined, a linguistic metaphor is an expression that compares or likens two distinctly different subject matters (Hills n.d.). Figurative language in general is commonly understood as "speech where speakers mean something other than what they literally say" (Gibbs and Colston 2012: p 1). Metaphor, in being a figurative device is figurative in the sense that, as Robert Fogelin puts it 'the speaker is trying to induce in the respondent a (mutually recognized) adjustment or replacement of what the speaker actually said' (Fogelin 1986/2011: p 87). So, at a rough—and overly crude—approximation, on this standard view, if we take it that a literal utterance has its intended meaning, at least in part, by activating an understanding, or concept, in the hearer's mind, figurative utterances would be used to elicit a concept distinct from the one standardly elicited by the utterance used. And, by contrast, literal utterances could be defined as expressions used to elicit a concept standardly evoked by that utterance.

On this traditional reading, therefore, in saying 'He attacked every weak point in my argument' (Lakoff and Johnson 1980), the speaker, it is assumed, is eliciting a concept other than the one standardly meant by the words i.e. they talk of WAR but mean ARGUMENT and, in doing so, aim to show that ARGUMENT is comparable to, but distinct from, WAR in some way; they compare *and* contrast ARGUMENT with WAR.

According to CMT, however, this explanation is wrong. CMT's Conceptual Claim, recall, is a claim about how abstract target concepts are structured; stating that the source concept is a necessary part of the target concept's structure. Thus, ARGUMENT, as an abstract concept, is not fully comprehensible except in terms of WAR. Therefore, the speaker who uses metaphorical language, it is hypothesized, does not simply understand ARGUMENT as being *comparable to WAR*, or *like WAR*, but, instead, that ARGUMENT *is* WAR. Hence, if CMT's Conceptual Claim is true, when one talks about arguments using war-like language, that language is literal, as they are not comparing and contrasting two disparate concepts, but are, rather, speaking directly, which Lakoff and Johnson point out themselves (1980). But if using war-based language to talk about arguments does not count as metaphorical or figurative, what does? Where is the line between literal and figurative to be drawn? CMT's Conceptual Claim, seems, at best, to blur the lines between literal and figurative, and, at worst, to destroy it all together.

Now, this is not problematic per se. Many theorists have argued against a clear-cut dichotomous distinction between literal and figurative (Gibbs and Colston 2012), but this *is* (ironically) a problem for CMT itself, as if the Conceptual Claim is correct, we might ask: How can CMT theorists conduct their work; how are literal uses to be distinguished from figurative ones? The consequence of the Conceptual Claim, then, is an immediate methodological conundrum: the isolation and analysis of linguistic metaphors in discovering conceptual metaphors, which is the standard and predominant way by which the presence of a particular conceptual metaphor is inferred, has been called into question, as it rests on the assumption that a distinction between metaphorical, and literal meaning can be easily drawn; an assumption which is undermined by CMT's foundational Conceptual Claim. Due to the fact that many CMT theorists continue to search for supposed linguistic metaphors as evidence for conceptual metaphors, it would seem that, although they hold onto the Conceptual Claim, they fail to realize its problematic consequences for the methods they use.

With that said, some have noted this problem. In recent work, for example, CMT scholar Zoltan Kövecses (2020) has perceptively acknowledged the possible troubles faced by the theory in relation to the literal-figurative distinction by noting that, in many cases, it is difficult to draw a clear line between figurative and literal uses of words:

It is reasonable to problematize even a basic assumption of CMT: the idea that there is such a thing as literal meaning. Is it indeed the case that we can and do base our understanding of figurative meaning on literal meaning? Even more radically: What is literal meaning? And, in the final analysis, how would the answers to these questions impact CMT as we know it today? (Kövecses 2020: p 19)

Based on these questions, Kövecses proposes a 'radical view of the extent of the literal, which greatly reduces or even eliminates the literal—at least in a certain sense' (Kövecses 2020: p 23). In his much-needed attempt to address known problems with CMT, Kövecses outlines a new Extended Conceptual Metaphor Theory, in his book of the same name, which claims that most words, even apparently uncontroversial

cases of literal language, are really figurative and gives support for this by showing that the etymological roots of many words have metaphorical origins, most of which have been forgotten by modern-day speakers:

It seems that a large number of expressions (that belong to abstract domains) are taken to be literal, but in synchrony their corresponding concepts are metaphorically comprehended and from a historical perspective they have metaphorical or metonymic origins. Their perception as literal seems to be based on the fact that their morphology does not reveal to contemporary speakers any metaphoric origins. (Kövecses 2020: p 25)

The view expressed here acknowledges the importance of the distinction between literal and figurative language which is vital to the study of conceptual metaphor. Faced with this issue, Kövecses (2020) concludes (contrary to Lakoff and Johnson) that, almost all language must be, to some degree, *figurative*.

Again, whether a clear-cut dichotomy between literal and figurative language can be maintained, or not, isn't the real issue here. So, the approach Kövecses advocates only keeps the problem at arm's length, because, whether it is the case that language turns out to be all literal or all figurative, the point still remains that the difference between the two is obscured, thus leaving theorists with no principled distinction by which to isolate and analyze metaphorical language. Hence, the study of conceptual metaphors cannot be accurately achieved via this method. A vast amount of the work on conceptual metaphor, however, employs this very technique of picking out what are assumed to be linguistic metaphors in discourse. This is perhaps exemplified by the MIP (Pragglejaz Group 2007) and MIPVU methods, both of which are prominent metaphor identifying procedures (Steen et al. 2010).²

The above considerations present an immediate challenge to CMT by showing how its own Conceptual Claim has the consequence of undermining the efficacy of the standard methodology it employs. The Conceptual Claim, however, might be held onto, so long as the method of inferring the existence of conceptual metaphors from the presence of linguistic metaphors is jettisoned and, along with it, the Linguistic Claim that licences such methods, as I will proceed to show.

4 Mounting the Attack

Although the Conceptual Claim is problematic, the Linguistic Claim is more so. The Linguistic Claim acts as the logical ground which licences and motivates the methodology of inferring the existence of conceptual metaphors through the analysis of linguistic metaphors. Although cognitive linguists often refrain from, and reel at, the use of the words 'necessary' and 'sufficient' conditions, the fact remains that in order to licence a valid inference from the presence of a linguistic metaphor to the

² See Gibbs (2017) where he devotes Chapter 3 to a discussion on metaphor identifying techniques and studies which employ them. He also covers the inferring of conceptual metaphors from linguistic metaphors in Chapter 4.

existence of a conceptual metaphor, the Linguistic Claim needs to be: *conceptual metaphors are necessary and sufficient for linguistic metaphors*, which, I will proceed to show, is not the case. Hence, the Linguistic Claim must be rejected.

Throwing out the Linguistic Claim is no mean feat for CMT theorists due to it underpinning and justifying large amounts of research done in the field. Indeed, despite the reliance on linguistic metaphors being criticized previously (Casasanto 2009), the Linguistic Claim is still held firmly as a foundational assumption of Conceptual Metaphor Theory, as Gibbs explains:

CMT firmly embraces the view that verbal metaphors [...] arise and are interpreted from the cross-domain mapping of knowledge from a source [...] to a target domain [...]. CMT sees these conventional expressions, along with other linguistic patterns, as “correlational” metaphors that provide evidence for the existence of cross-domain mappings, or conceptual metaphors. (Gibbs 2017)

For the above-outlined inference to be reliable, conceptual metaphors need to be both necessary and sufficient for linguistic metaphors. Hence, the legitimacy of this claim can be challenged by showing that conceptual metaphors (if real) are neither necessary nor sufficient for linguistic metaphors. In order to do this, it will suffice to show that (i) abstract target domains can be spoken of in literal terms, proving that conceptual metaphors are not *sufficient* for linguistic metaphors, and (ii) that ad hoc and creative linguistic metaphors can be produced which are clearly not the product of an underlying conceptual mapping, proving that conceptual metaphors are not *necessary* for their linguistic counterparts, both of which I will proceed to do.

I will start with the case for sufficiency. As an example, I will discuss the generally accepted (within CMT) conceptual mapping of MORALITY with UPRIGHTNESS. The cognitive importance of this conceptual metaphor is explained by Lakoff:

Without understanding morality as uprightness and evil as a force, we could not conceptualize moral failure as falling, remaining moral as standing up to evil, building moral strength as requiring discipline and self-denial. (Lakoff 1996)

This example has become accepted as an uncontroversial and standard case of a conceptual metaphor (Lakoff 1996; Gibbs 2017; Kövecses 2010; Yu 2016; Yu et al. 2016). The MORALITY IS UPRIGHTNESS mapping means that moral virtue is conceived of as being upright or high and, conversely, that moral depravity is fallen, down or low. The existence of this conceptual metaphor is normally inferred from such linguistic metaphors as the following:

VIRTUE IS UP; DEPRAVITY IS DOWN

He is *high*-minded.

She has *high* standards.

She is *upright*.

She is an *upstanding* citizen.

That was a *low* trick.

Don't be underhanded.

I wouldn't *stoop* to that.
 That would be *beneath* me.
 He *fell* into the *abyss* of depravity.
 That was a *low-down* thing to do.
 (Lakoff and Johnson 1980)

It should be noted that there is theoretical and experimental support for the embodied nature of moral concepts in general (Bartlett 2020, 2021; Liao 2016; Prinz 2007) and specifically to the existence of the MORALITY IS UPRIGHTNESS conceptual mapping (Lynott and Coventry 2014; Meier et al. 2007a, b) and so it is plausible that such a mapping does exist. I do not, however, wish to dispute that embodied concepts or cross-domain conceptual mappings are psychologically real, but only want to point out that, if they are, they are not sufficient for linguistic metaphors.

It is claimed, as seen above, by Lakoff (1996) that UPRIGHTNESS is essential to our conception of MORALITY; that we cannot understand this abstract target domain *par excellence* without the conceptual structuring from the concrete source domain of UPRIGHTNESS. Gibbs also makes this point:

[T]arget domains mostly refer to topics that are difficult to talk or think about without using metaphor, precisely because of their general abstract nature. For instance, it is almost impossible to speak of morality, thoughts, and relationships without metaphor rushing in to facilitate our understanding of these aspects of human life. (Gibbs 2017)

Stated in such a way, it, indeed, appears that CMT theorists assume conceptual metaphors to be both necessary and sufficient for linguistic metaphors. It is, however, possible to show that this is not the case. Take the following examples:

- (1) She is a *morally righteous* person.
- (2) She has the moral *high ground*.

Example (1) is an uncontroversial (assuming the literal-figurative distinction holds) case of literal language being used to talk about the abstract domain MORALITY and (2) is a classic case of metaphorical language. Assuming the psychological reality of the MORALITY IS UPRIGHTNESS mapping, the fact that the literal utterance in (1) is possible at all shows that conceptual metaphors are not sufficient for the production of linguistic metaphors. That is to say, if conceptual metaphors were sufficient for linguistic metaphors, speaking about an abstract domain in literal terms would be impossible. Hence, in light of this observation—that the same proposition about the abstract domain MORALITY *can* be expressed through both literal and figurative language—CMT theorists are faced with the following options:

- (a) Accept that conceptual metaphors are not sufficient for linguistic metaphors.
- (b) Reject the argument by claiming that the MORALITY IS UPRIGHTNESS mapping is not psychologically real.

- (c) Reject the argument by claiming that the apparent literal utterance in (1) is in fact metaphorical.

If option (a) is chosen, and the argument is accepted, the Linguistic Claim might possibly be rescued by showing that, although conceptual metaphors are not sufficient for linguistic metaphors, they *are* necessary. In this case, CMT would need to be enriched by positing another contributing factor which, together with conceptual metaphor, is jointly sufficient for the production of linguistic metaphor. Alternatively, it could be claimed (b) that the above argument is flawed due to the conceptual metaphor MORALITY IS UPRIGHTNESS not being a *bona fide* conceptual metaphor, but such a move would come dangerously close to undermining CMT itself, as it is a generally accepted and externally supported case of a cross-domain mapping. The last option is (c): to claim that example (1) is, in fact, a figurative utterance, thereby avoiding the challenge. However, this has problematic consequences for the literal-figurative distinction, and so the theory falls back onto the first argument above.

In addition to conceptual metaphors being insufficient for linguistic metaphors, it is also plausible that they are not *necessary* for them. Although CMT depends on a necessary connection between cross-domain conceptual mappings and metaphorical language, it is far from clear that conceptual metaphors are necessary for linguistic metaphors. Challenging this necessary connection can be shown by considering creative or ad hoc linguistic metaphors which appear not to have a corresponding conceptual metaphor. Consider the following well-worn examples:

- (3) Juliet is the sun
 (4) He kicked the bucket

Examples (3) and (4) are famous and uncontroversial cases of metaphorical language, the meanings of which are comprehensible to English speakers. It appears, upon analysis, that these examples of figurative language are not the result of an essential and necessary cross-domain mapping between source and target concepts in the mind. That is, to take example (3), we can see from introspection that the concept LIGHT SOURCE (sun) is neither intrinsic to, nor necessary for, our understanding of the concept LOVER (Juliet); that the source domain LIGHT SOURCE is not a conceptual prerequisite for grasping the target domain LOVER. Yet, we can still use and understand the linguistic metaphor in (3). CMT's Linguistic Claim would predict that one could not produce or comprehend (3) unless one's mind contained a corresponding conceptual metaphor. This, however, is not the case for myself, and I appeal to the reader (assuming you understand Shakespeare's metaphor) to consider whether LIGHT SOURCE is a concept which is necessary to your understanding of LOVER. And the same, of course, applies to example (4). It is intuitively implausible that our understanding of KICKING A BUCKET/PHYSICAL MOVEMENT is a conceptual prerequisite for being able to conceive of DYING, yet we still grasp and use the idiom.

In addition to the above poetic and idiomatic examples, some extremely creative and often humorous metaphors can be seen in euphemisms and slang phrases. Such examples give further reason to believe that conceptual metaphors are not necessary for linguistic metaphors. Take, for instance, those often-vulgar phrases referring to masturbation:

- (5) Bashing the bishop
- (6) Polishing the banister
- (7) Spanking the monkey
- (8) Flicking the bean
- (9) Patting the bunny
- (10) Paddling the pink canoe

Although (5–10) all refer to the same act of self-love, and, therefore, all share the same target domain of MASTURBATION, from both male (5–7) and female (8–10) perspectives, the source domains evoked are extremely diverse. In examples (5–10) there is a two-fold mapping where the action and the object are both expressed metaphorically. Take example (5), for instance, which invokes a conceptual mapping between PENIS (target domain) and BISHOP/PERSON (source domain) and another separate mapping between MASTURBATION (target domain) and VIOLENT ACTION (source domain). Upon introspection, although I understand, and have used, such phrases, it appears extremely implausible that the concept PENIS is necessarily structured by the concept PERSON (bishop); that PERSON is a conceptual prerequisite to PENIS. Or, indeed, that I would not be able to comprehend MASTURBATION without it being structured by VIOLENT ACTION (bashing). This is even less plausible in light of examples (6), (9) and (10) which map MASTURBATION with non-violent actions, for the same target concept MASTURBATION, in the mind of a speaker who comprehends and produces utterances (5–10), would seem to be paradoxically structured by the incongruent source domains of VIOLENT ACTION *and* NON-VIOLENT ACTION. It seems extremely implausible that PERSON/BISHOP would be a conceptual prerequisite for understanding MASTURBATION. Indeed, an argument could be made that the very reason euphemisms such as the above are humorous is due to the juxtaposition and jarring of the concepts which the metaphors invite us to compare.

When faced with the above, the CMT theorist could claim that, if analysed at the most schematic level, (5–10) all make manifest a single conceptual metaphor MASTURBATION IS DOING SOMETHING TO SOMETHING, and that this conceptual mapping is necessary for the creation and comprehension of these linguistic metaphors. Indeed, the obvious similarity between the examples above seems to be the grammatical frame ‘X-ing the Y’ where X is a verb and Y is a singular noun. But when formulated at this maximally schematic level, it tells us nothing interesting about the way MASTURBATION is conceptually structured; only that it is an action of *doing something to something* which is, of course, a literally true and trivial analysis. What *can* be safely concluded, however, is that it is extremely implausible that comprehension of bashing bishops, patting bunnies or flicking beans is necessary for

our understanding of masturbation. Such phrases, therefore, give reason to doubt the necessary role of conceptual metaphor in figurative language.

In sum, I have argued that conceptual metaphors are neither necessary nor sufficient for linguistic metaphors. It is possible to talk of abstract domains such as MORALITY—for which there is experimental data supporting the psychological reality of the cross-domain conceptual structure—with literal language, hence conceptual metaphors are not *sufficient* for figurative language. Moreover, it is not plausibly the case that conceptual metaphors are *necessary* for linguistic metaphors, due to the possibility and occurrence of metaphors which can be easily comprehended, but which are unlikely to be the product of a cross-domain conceptual mapping. Hence, if we can produce the linguistic metaphor without the conceptual metaphor, the latter is not necessary for the former, thus undermining the Linguistic Claim.

4.1 A Strategic Victory

If the above holds, the problem now faced by CMT is that inferring the existence of conceptual metaphors from the presence of linguistic metaphors represents fallacious reasoning of the following formalized structure:

P1: $\neg(A \Leftrightarrow B)$
 P2: B
 C: $\therefore A$

It must also be noted that even if A was sufficient for B , but not necessary, this argument form would still be invalid—a case of affirming the consequent. Only if the link between conceptual metaphor and linguistic metaphor were necessary *and* sufficient, would one be justified in inferring the existence of a particular cross-domain mapping or conceptual metaphor from observations of linguistic metaphors. Hence, the CMT theorist needs to reconsider the nature of the hypothesized link claimed to exist between thought and language stated in the Linguistic Claim along with their methodology.

5 Counterstrike

The argument presented here might be objected to on the grounds that the current reading of CMT's foundational claims is deliberately too strong; that I have mischaracterized CMT by asserting that the theory understands concrete source domains as the *necessary* building blocks of abstract target domains. It might be claimed that the predominantly held, and correct, version of CMT is weaker—let's call it *Weak CMT*. Weak CMT would be the claim that abstract target domains are *sometimes* structured by concrete source domains in order to aid conceptualization, but that such structuring is neither universal nor necessary. Directed at Weak CMT, the challenges presented here would, indeed, have little force. It appears, however, that CMT is, in fact, intended as a strong claim, as I will proceed to show. Furthermore, a weak

interpretation of CMT should, I caution, be avoided, as it is philosophically and psychologically uninteresting and risks reducing the theory to triviality.

Now, as just outlined, it might be claimed, in order to avoid the arguments presented above, that conceptual metaphor theorists, in fact, make a weaker claim than the one I have argued against here, thus espousing Weak CMT which could be characterized as holding onto the following ‘weak’ claims:

Weak Conceptual Claim Abstract concepts are, sometimes, but not necessarily, structured by concrete concepts.

Weak Linguistic Claim Conceptual metaphors sometimes, though not always, give rise to linguistic metaphors and figurative language.

This weak version of CMT states that it is possible, but not necessary, for abstract concepts to be understood in terms of more concrete ones and that the connections between mind and language are contingent. The weaker claim, however, amounts to little more than stating that some concepts can become associated with others; that we have the cognitive capacity of being able to compare and notice similarities held between divergent conceptual domains, and that sometimes this leads us to speak figuratively. But this, I claim, is not at all a compelling theory about human conceptualization, as the theory is now about association and recognition, as opposed to conceptual structure.

Moreover, it is clear that CMT theorists *do* hold onto a strong version of CMT. Recall the sections quoted above:

The metaphor is not merely in the words we use *it is in our very concept of an argument*. [...] We talk about arguments that way because we conceive of them that way. (Lakoff and Johnson 1980; emphasis added)

And Raymond Gibbs:

CMT firmly embraces the view that verbal metaphors [...] *arise and are interpreted from the cross-domain mapping* of knowledge from a source [...] to a target domain. (Gibbs 2017; emphasis added)

The Strong view is still evident in Lakoff and Johnson’s later writings. On the abstract domain MORALITY they write:

Virtually all of our abstract moral concepts are structured metaphorically. (Lakoff and Johnson 1999; p 290)

This makes clear the Conceptual Claim of CMT: abstract concepts are ‘structured’ metaphorically by source concepts. It is not simply that target domains are ‘associated’ with source domains, but that they are ‘structured’ by them, making the source domains necessary for the understanding of the target. If what is actually meant by ‘structured’ is ‘associated with’, please forgive my misunderstanding, which is, I hold, through no fault of my own.

If it is rebuked that Strong CMT was an earlier version and that contemporary theorists no longer hold the ‘original’ strong view, this can be rebutted by showing

that this reading is, indeed, still endorsed in contemporary work. Take, for instance, the following work published in *Cognitive Linguistics* by Benczes and Ságvári (2018) who preface their CMT-based study in the following way:

As laid out by Lakoff and Johnson (1980) in what has become known as Conceptual Metaphor Theory (CMT), abstract concepts, such as life, *can only be understood or made sense of by relying on more concrete concepts*, resulting in conceptual metaphors – such as LIFE IS A JOURNEY – that serve as the “principal vehicles for understanding” (Lakoff and Johnson 1980: 133). (Benczes and Ságvári 2018: emphasis added)

The Strong CMT view can also be seen in Ning Yu’s work:

[O]ur moral thinking is imaginative in nature, *depending fundamentally on our metaphorical understanding*, and it is through metaphor that many of our ethical values and principles emerge from our embodied and socioculturally situated habitation of the world. (Yu 2015; emphasis added)

Yu explains that moral thinking ‘depends fundamentally’ on our ‘metaphorical understanding’, which suggests, again, that source concepts are seen as *necessary* to the understanding of target concepts. If, again, what is actually meant by ‘depending fundamentally on our metaphorical understanding’ (Yu 2015) is ‘sometimes having contingent conceptual associations’, then I am sure the reader will forgive me, once more, for being misled by the language used in the literature.

Lastly, the Strong view is also evident in the recent writings of George Lakoff on his Neural Theory of Metaphor where he explains what he considers to be the neural basis of conceptual metaphors:

Complex concepts are formed by neural binding circuits, which bind together schemas in different parts of the brain. [...] Binding circuits are the primary mechanism of neural composition forming complex concepts by binding nodes across diverse brain regions. (Lakoff 2014)

It is evident from the literature that Strong CMT is the version standardly proposed and supported. Thus, theorists are forced to either continue with Strong CMT and amend it in light of the arguments put forward here, or adopt Weak CMT, which might avoid some of my challenges, but risks reducing the theory to triviality and still leaves the CMT theorist with the methodological conundrum of not being able to validly infer the existence of a conceptual metaphor from the presence of a linguistic metaphor.

6 Conclusion: Ceasefire

I have presented two challenges to the theoretical framework of Conceptual Metaphor Theory which need to be addressed by scholars working within the field. Firstly, I have argued that the Conceptual Claim leads to the blurring of the literal-figurative distinction and, therefore, casts doubt on the standard methodology used

to discover conceptual metaphors by identifying linguistic metaphors, which relies on a clear distinction between metaphorical and literal utterances.

Secondly, I have challenged the Linguistic Claim by showing that conceptual metaphors—if psychologically real—are neither necessary nor sufficient for linguistic metaphors. The CMT theorist, therefore, needs to provide a counterargument falsifying the claims made here by proving the necessary role of conceptual metaphors in the construction of figurative language or must reject the Linguistic Claim.

Crucially, I have argued, the CMT theorist needs to prove that conceptual metaphors are necessary *and* sufficient for linguistic metaphors in order to licence a valid inference from observations of linguistic metaphors to the existence of conceptual metaphors. Alternatively, the CMT theorist could accept the argument here and hold on to the Conceptual Claim, but discard the Linguistic Claim. In doing this, they are forced to abandon standard linguistic methods in the study of conceptual metaphor. Hence, in light of the arguments outlined here, the following four options present themselves to CMT theorists with respect to the two foundational tenets of the theory:

- A: Hold on to both the Conceptual Claim and Linguistic Claim, by showing that conceptual metaphors are indeed necessary and sufficient for linguistic metaphors in order to avoid the charge of fallacious reasoning.
- B: Discard the Conceptual Claim, and give up the theory.
- C: Discard both claims, and give up the theory.
- D: Hold on to the Conceptual Claim, but reject the Linguistic Claim, and abandon the methods of discovering conceptual metaphors through analysis of metaphorical and figurative language.

I argue that D is the only tenable option.

Acknowledgements I would like to thank all anonymous reviewers for their helpful critiques on earlier versions of this paper which have ultimately helped to improve the arguments presented herein. This work was conducted at Srinakharinwirot University, Bangkok as part of a larger Ph.D. thesis under the supervision of Sugunya Ruangjaroon.

Declarations

Conflict of interest The authors have no conflicts of interest to disclose.

References

- Aristotle., & Heath, Malcolm (1996) *Poetics*. Penguin Books
- Bartlett JJ (2020) An expected error: an essay in defence of moral emotionism. *Axiomathes*. <https://doi.org/10.1007/s10516-020-09524-5>
- Bartlett JJ (2021) Motivating emotions: emotionism and the internalist connection. *Axiomathes*. <https://doi.org/10.1007/s10516-021-09550-x>
- Benczes R, Ságvári B (2018) Where metaphors really come from: social factors as contextual influence in Hungarian teenagers' metaphorical conceptualizations of life. *Cogn Linguist* 29(1):121–154. <https://doi.org/10.1515/COG-2016-0139>

- Bundgaard PF (2019) The structure of our concepts: a critical assessment of Conceptual Metaphor Theory as a theory of concepts. *Cogn Semiot*. <https://doi.org/10.1515/COGSEM-2019-2010>
- Casasanto D (2009) When is a linguistic metaphor conceptual metaphor?, pp 127–145. <https://doi.org/10.1075/hcp.24.11.cas>
- Crawford LE (2009) Conceptual metaphors of affect. *Emot Rev* 1(2):129–139. <https://doi.org/10.1177/1754073908100438>
- Evans V (2004) *The structure of time: language, meaning and temporal cognition*. John Benjamins Publishing Company
- Fogelin RJ (2011) *Figuratively speaking: revised edition*. Oxford University Press, Oxford, pp 1–144. <https://doi.org/10.1093/ACPROF:OSO/9780199739998.001.0001>
- Gibbs R (1994) *The poetics of mind: figurative thought language and understanding*. Cambridge University Press
- Gibbs R (2014) Why do some people dislike conceptual metaphor theory? *Cogn Semiot* 5(1–2):14–36. <https://doi.org/10.1515/cogsem.2013.5.12.14>
- Gibbs R (2017) *Metaphor wars*. Cambridge University Press. <https://doi.org/10.1017/9781107762350>
- Gibbs J, Colston HL (2012) *Interpreting figurative meaning*. Cambridge University Press. <https://doi.org/10.1017/cbo9781139168779>
- Grady J (1997) *Foundations of meaning: primary metaphors and primary scenes*. University of California
- Haser V (2005) Metaphor, metonymy, and experientialist philosophy. *DE GRUYTER*. <https://doi.org/10.1515/9783110918243>
- Hills D (xxxx) Metaphor. *Stanford encyclopaedia of philosophy*. Retrieved from <https://plato.stanford.edu/entries/metaphor/>
- Horstmann G, Ansorge U (2011) Compatibility between tones, head movements, and facial expressions. *Emotion* 11(4):975–980. <https://doi.org/10.1037/a0023468>
- Keysar B, Shen Y, Glucksberg S, Horton WS (2000) Conventional language: How metaphorical is it? *J Mem Lang* 43(4):576–593. <https://doi.org/10.1006/jmla.2000.2711>
- Kövecses Z (2010) *Metaphor: a practical introduction*, Second. Oxford University Press
- Kövecses Z (2020) *Extended conceptual metaphor theory*. Cambridge University Press. <https://doi.org/10.1017/9781108859127>
- Lakoff G (1996) *Moral politics: how liberals and conservatives think*. University of Chicago Press
- Lakoff G (2014) Mapping the brain's metaphor circuitry: metaphorical thought in everyday reason. *Front Hum Neurosci* 8(1):958. <https://doi.org/10.3389/fnhum.2014.00958>
- Lakoff G, Johnson M (1980) *Metaphors we live by*. University of Chicago Press
- Lakoff G, Johnson M (1999) *Philosophy in the flesh: the embodied mind and its challenge to Western thought*. Basic Books
- Leezenberg M (2014) From cognitive linguistics to social science: thirty years after metaphors we live by. *Cogn Semiot* 5(1–2):140–152. <https://doi.org/10.1515/cogsem.2013.5.12.140>
- Liao SM (2016) *Moral brains: the neuroscience of morality*—Google Books. Oxford University Press
- Lycan WG (2008) *Philosophy of language: a contemporary introduction*. Routledge
- Lynott D, Coventry K (2014) On the ups and downs of emotion: Testing between conceptual-metaphor and polarity accounts of emotional valence-spatial location interactions. *Psychon Bull Rev* 21(1):218–226. <https://doi.org/10.3758/s13423-013-0481-5>
- McGlone MS (2007) What is the explanatory value of a conceptual metaphor? *Lang Commun* 27(2):109–126. <https://doi.org/10.1016/j.langcom.2006.02.016>
- McGlone MS (2011) Hyperbole, homunculi, and hindsight bias: an alternative evaluation of conceptual metaphor theory. *Discourse Process* 48(8):563–574. <https://doi.org/10.1080/0163853X.2011.606104>
- Meier BF, Robinson MD (2006) Does “feeling down” mean seeing down? Depressive symptoms and vertical selective attention. *J Res Pers* 40(4):451–461. <https://doi.org/10.1016/j.jrp.2005.03.001>
- Meier BF, Hauser DJ, Robinson MD, Friesen CK, Schjeldahl K (2007a) What’s “up” with God? Vertical space as a representation of the Divine. *J Pers Soc Psychol* 93(5):699–710. <https://doi.org/10.1037/0022-3514.93.5.699>
- Meier BF, Sellbom M, Wygant DB (2007b) Failing to take the moral high ground: psychopathy and the vertical representation of morality. *Personal Individ Differ* 43(4):757–767. <https://doi.org/10.1016/j.paid.2007.02.001>
- Miller KA, Raney GE, Demos AP (2020) Time to throw in the towel? No evidence for automatic conceptual metaphor access in idiom processing. *J Psycholinguist Res* 49(5):885–913. <https://doi.org/10.1007/s10936-020-09728-1>

Axiomathes

- Pragglejaz Group (2007) MIP: a method for identifying metaphorically used words in discourse. *Metaphor Symb* 22(1):1–39. <https://doi.org/10.1080/10926480709336752>
- Prinz JJ (2007) *The emotional construction of morals*. Oxford University Press
- Rakova M (2002) The philosophy of embodied realism: a high price to pay? *Cogn Linguist* 13(3):215–244. <https://doi.org/10.1515/cogl.2002.015>
- Schubert TW (2005) Your highness: vertical positions as perceptual symbols of power. *J Personal Soc Psychol* 89(1):1–21. <https://doi.org/10.1037/0022-3514.89.1.1>
- Steen GJ, Dorst AG, Herrmann JB, Kaal A, Krennmayr T, Pasma T (2010) A method for linguistic metaphor identification. In: *Converging evidence in language and communication research*. John Benjamins. <https://doi.org/10.1075/celcr.14>
- Vervaeke J, Kennedy JM (1996) Metaphors in language and thought: falsification and multiple meanings. *Metaphor Symb Act* 11(4):273–284. https://doi.org/10.1207/s15327868ms1104_3
- Wierzbicka A (1986) Metaphors linguists live by. *Pap Linguist* 19(2):287–313. <https://doi.org/10.1080/08351818609389260>
- Yu N (2015) Metaphorical character of moral cognition: a comparative and decompositional analysis. *Metaphor Symb* 30(3):163–183. <https://doi.org/10.1080/10926488.2015.1049500>
- Yu N (2016) Spatial metaphors for morality: a perspective from Chinese. *Metaphor Symb* 31(2):108–125. <https://doi.org/10.1080/10926488.2016.1150763>
- Yu N, Wang T, He Y (2016) Spatial subsystem of moral metaphors: a cognitive semantic study. *Metaphor Symb* 31(4):195–211. <https://doi.org/10.1080/10926488.2016.1223470>

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

VITA

NAME Justin James Bartlett

DATE OF BIRTH 05 April 1988

PLACE OF BIRTH UK

INSTITUTIONS ATTENDED University of Glamorgan
Bangor University
Ph.D. of Srinakharinwirot University

HOME ADDRESS 282/502 Condo Centric, Ratchadapisek Road, Huai Kwang,
Huai Kwang, Bangkok, Thailand

